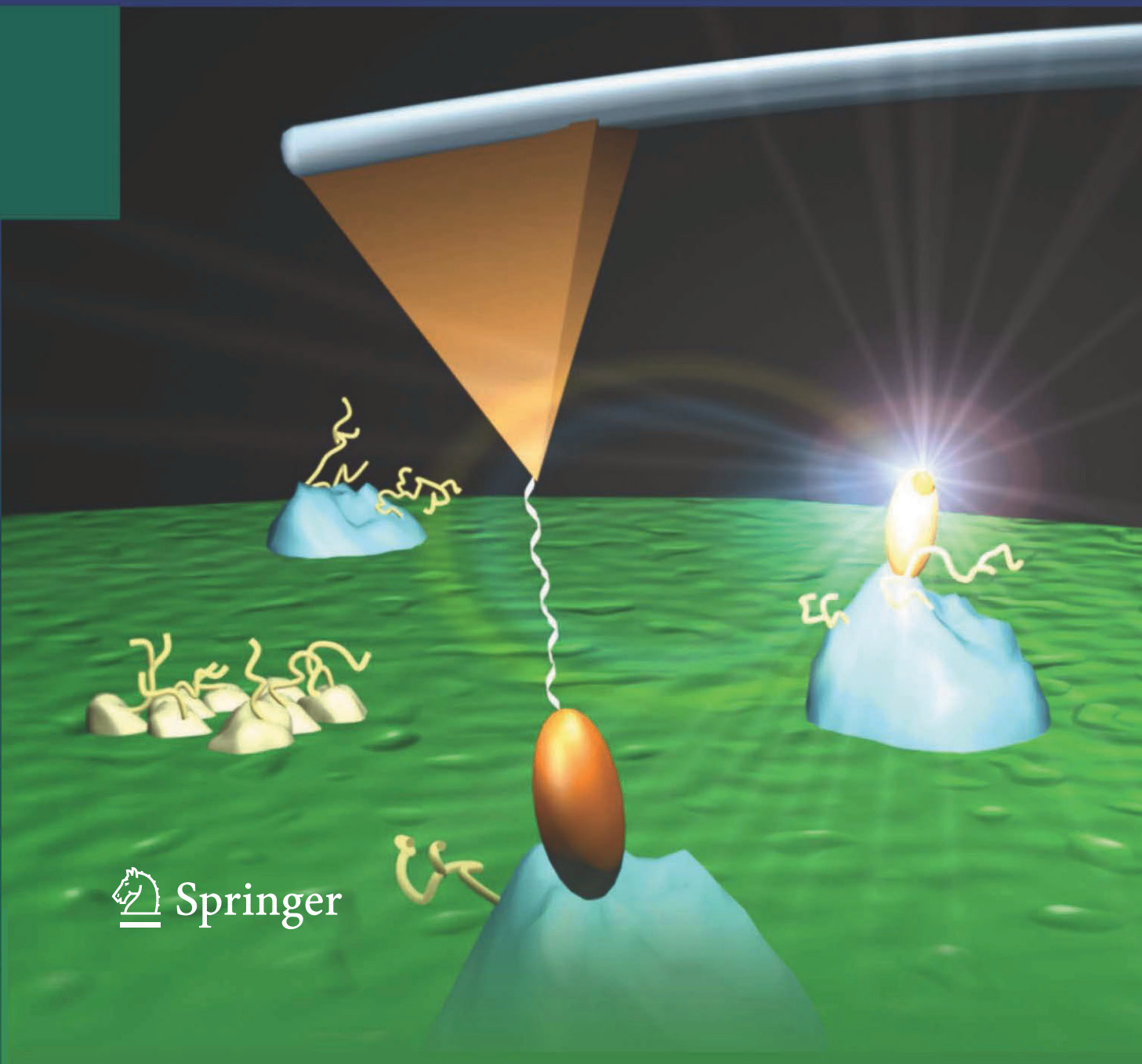
HANDBOOK OF
Single-Molecule Biophysics

Peter Hinterdorfer ▾ Antoine van Oijen
Editors



Springer

Handbook of Single-Molecule Biophysics

Peter Hinterdorfer, Antoine van Oijen
Editors

Handbook of Single-Molecule Biophysics

Editorial Advisory Board

Springer

Peter Hinterdorfer
Institute for Biophysics
Johannes Kepler University Linz
Altenberger Straße 69
4040 Linz
Austria
peter.hinterdorfer@jku.at

Antoine van Oijen
Department of Biological Chemistry
& Molecular Pharmacology
Harvard University
Harvard Medical School
250 Longwood Ave.
Boston MA 02115
Bldg. C
USA
antoine_van_oijen@hms.harvard.edu

ISBN 978-0-387-76496-2 e-ISBN 978-0-387-76497-9
DOI 10.1007/978-0-387-76497-9
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009921359

Preface

The history of single-molecule sciences can be best captured in three short quotations. Erwin Schrödinger stated in *Br. J. Phil. Sci.* 3:109 (1952) that "In the first place it is fair to say that we are not experimenting with single particles any more than we can raise Ichthyosauria in the zoo", demonstrating that the observation of individual molecules was beyond imagination at the time. Subsequently, Richard Feynman's talk at the December 29,1959, annual meeting of the American Physical Society at the California Institute of Technology titled "There's plenty of room at the bottom" represents the initial exploration of the microcosmos and the nanoworld. Finally, Joseph M. Beechem wrote in *Biophys. J.* 67:2133 (1994) "it could very well occur that biophysics (in the next decade) could become dominated by single molecule techniques", anticipating that with the potential of observing single biomolecules a new, exciting scientific area will be opened.

Indeed, during the last decade, a number of novel biophysical methods have been developed that allow the manipulation and study of individual biomolecules. The ability to monitor biological processes at the fundamental level of sensitivity, that of a single molecule, has given rise to an improved understanding of the underlying molecular mechanisms. Through the removal of ensemble averaging, distributions and fluctuations of molecular properties can be characterized, transient intermediates identified, and catalytic mechanisms elucidated. Compared to conventional ensemble methods, single-molecule experiments offer several advantages. By conducting many sequential measurements, they allow the distribution of molecular properties of inhomogeneous systems to be determined. Being direct records of the system's fluctuations, single-molecule trajectories provide dynamic and statistical information that is often hidden in ensemble-averaged results. They also permit real-time observation of rarely populated transients, which are difficult or impossible to capture using conventional methods.

This book will introduce the reader to the different classes of single-molecule approaches developed over recent years and how these methods can be applied to understand biological systems. Broadly, single-molecule biophysical techniques can be separated into two classes of methods: one that allows us to visualize single molecules, the other to manipulate them. The potential of imaging the emission of single fluorophores has led to many exciting developments that allow us to study biomacromolecules, their movement, their interactions, and their cellular context. Biochemical and genetic methods have been developed that allow the very specific and precise fluorescent labeling of biomacromolecules. Optical excitation and detection methods have been introduced that allow the visualization of individual, fluorescently labeled molecules with high spatial and temporal resolution.

Using fluorescence as read-out, many molecular properties become accessible and a wealth of mechanistic detail can be obtained. Fluctuations in the height of fluorescence signal are used in fluorescence correlation spectroscopy to extract information on concentrations and diffusion kinetics. In the proximity of a second fluorophore, fluorescence resonance energy transfer can be used to correlate changes in fluorescence intensity to physical movements

of two biomacromolecules. Besides fluorescence intensity, also wavelength, polarization, and fluorescence lifetime can be used to probe molecular interactions and conformational dynamics.

Visualization of fluorescence in a spatial context (imaging) can be used to track the movement of fluorescently labeled molecules through cells, across membranes, or along polymer tracks. The ability to pinpoint the exact position of a single fluorophore with a precision much better than the diffraction limit has allowed researchers to extract nanometer-scale information on the mechanisms of translocation and rotation. Recently, this approach has been used to achieve the imaging of cellular structures at unprecedented resolutions, filling the gap between traditional optical imaging and high-resolution techniques such as electron microscopy and X-ray crystallography. This book not only focuses on the various fluorescence-based single-molecule methods, but also discusses the many different biological systems that can be studied using these techniques. Examples are presented that vary from bare-bones, *in vitro* single-molecule biochemistry and biophysics to single-molecule studies in the context of living cells.

The second class of single-molecule tools, those that enable mechanical manipulation, allow us to perturb molecular systems at the molecular scale and enable the study of energetics of molecular interactions, mechanical properties of biopolymers, and kinetics of biochemical reactions, all while having complete control of physical experimental parameters such as force and distance. In the last decade, the atomic force microscope (AFM) has opened the door for structural and functional studies of native biological samples such as proteins, nucleic acids, membranes, and cells in their physiological environment. In some cases, topographical images with lateral and vertical resolutions in the angström range have been achieved. In addition, functionalizing AFM sensors with ligands allows for localizing cognate receptors on cell and membrane surfaces. Known for a long time as a relatively slow scanning technique, it has been developed (although not commercially available) into a video-imaging technique. This latter achievement renders possible the imaging of dynamic biological processes at the nanoscale.

The high sensitivity of force measurements utilizing AFM, optical and magnetic tweezers, or the biomembrane force apparatus allows for determination of inter- and intramolecular forces on the single-molecule level. The proof-of-principle stage of these pioneering experiments has evolved into established methods for exploring kinetic and structural details of molecular interactions, unfolding, and recognition processes. Data obtained from force spectroscopy include physical parameters not measurable by other methods and open new perspectives for exploring the regulation of the dynamics of biomolecular processes. With force spectroscopy, molecular interaction and unfolding can be analyzed in terms of affinity and rate constants. Moreover, force spectroscopy experiments yield details on structural parameters of the binding pocket, on the molecular dynamics of unfolding and recognition process, and on the energylandscape of the inter- and intramolecular interactions.

By applying forces on biomolecules while monitoring their activity, important information can be obtained on how proteins couple function to structure. The *Handbook of Single-Molecule Biophysics* provides an introduction to these techniques and presents an extensive discussion of the new biological insights obtained from them. It discusses the following areas:

- Experimental techniques to monitor and manipulate individual biomolecules (fluorescence detection, atomic force microscopy, optical and magnetic trapping)
- The use of single-molecule techniques in superresolution and functional imaging

- Single-molecule studies of physical properties of biomolecules (folding, polymer physics of proteins and DNA)
- Single-molecule enzymology and biochemistry
- Single molecules in the membrane
- Single-molecule techniques in living cells
- Integration of single-molecule biophysics and nanoscience

The *Handbook* provides a comprehensive overview for a broad audience of researchers in biophysics, molecular biology, chemistry, and related areas, as well as for advanced undergraduate and graduate students. We hope that it will stimulate further activities in this research area and also bring young researchers into this exciting field.

P. Hinterdorfer
A. van Oijen

Contents

List of Contributors

D. Alsteens, Unité de chimie des interfaces, Université catholique de Louvain, Croix du Sud 2/18, B-1348 Louvain-la-Neuve, Belgium
T. Ando, Department of Physics, Kanazawa University, Kakuma-machi, Kanazawa 920-1192 Japan, and CREST, JST, Sanban-cho, Chiyoda-ku, Tokyo 102-0075, Japan
R. M. Berry, Department of Physics, Clarendon Lab, University of Oxford, Oxford OX1 3PU, UK
W. Chen, Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
L. Chtcheglova, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
N. H. Dekker, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
G. T. Dempsey, Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA
Y. F. Dufrêne, Unité de chimie des interfaces, Université catholique de Louvain, Croix du Sud 2/18, B-1348 Louvain-la-Neuve, Belgium
V. Dupres, Unité de chimie des interfaces, Université catholique de Louvain, Croix du Sud 2/18, B-1348 Louvain-la-Neuve, Belgium
A. Ebner, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
J. Enderlein, III. Drittes Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
E. Evans, Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA, and Physics and Pathology, University of British Columbia, Vancouver, BC, Canada V6T 2A6
H. Gruber, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
K. Halvorsen, The Rowland Institute at Harvard, Harvard University, Cambridge MA 02142, USA
P. Hinterdorfer, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
J. Hohlbein, Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, OX1 3PU, UK
S. J. Holden, Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, OX1 3PU, UK
S. Howorka, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK
L. C. Hwang, Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, OX1 3PU, UK

A. N. Kapanidis, Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, OX1 3PU, UK
R. Kapon, Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel, 76100
K. Kinoshita, Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA, and Physics and Pathology, University of British Columbia, Vancouver, BC, Canada V6T 2A6
D. A. Koster, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
S. G. Lemay, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
J. Lipfert, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
J. J. Loparo, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
A. Meller, Department of Biomedical Engineering and Department of Physics, Boston University, Boston, MA 02215, USA
D. J. Müller, Institut Curie, UMR168-CNRS, 26 Rue d'Ulm, 75248 Paris, France, and Biotechnology Center, Technische Universität Dresden, 01307 Dresden, Germany
R. Nevo Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel, 76100
T. Pilizota, Department of Physics, Clarendon Lab, University of Oxford, Oxford OX1 3PU, UK
J. Preiner, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
C. Rankl, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
Z. Reich, Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel, 76100
F. Rico, Department of Physiology and Biophysics, University of Miami Miller School of Medicine, 1600 NW 10th Avenue, Miami, FL 33136, USA
M. Rief, Physikdepartment E22, Technische Universität München, James-Franck-Strasse, 85748 Garching, Germany
K. T. Sapra, Institut Curie, UMR168-CNRS, 26 Rue d'Ulm, 75248 Paris, France, and Biotechnology Center, Technische Universität Dresden, 01307 Dresden, Germany
K. Sarangapani, Coulter Department of Biomedical Engineering and Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
S. Scheuring, Institut Curie, UMR168-CNRS, 26 Rue d'Ulm, 75248 Paris, France, and Biotechnology Center, Technische Universität Dresden, 01307 Dresden, Germany
M. Schlierf, Physikdepartment E22, Technische Universität München, James-Franck-Strasse, 85748 Garching, Germany
T. Schmidt, Physics of Life Processes, Leiden Institute of Physics, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
G. J. Schütz, Institute of Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
P. Schwille, BIOTEC, Biophysics, Technische Universität Dresden, Tatzberg 47-51, 01307 Dresden, Germany

Z. Siwy, Department of Physics and Astronomy, University of California at Irvine, 210G Rowland Hall, Mail Code: 4575, Irvine, CA 92697, USA
G. V. Soni, Department of Biomedical Engineering and Department of Physics, Boston University, Boston, MA 02215, USA
Y. Sowa, Department of Physics, Clarendon Lab, University of Oxford, Oxford OX1 3PU, UK
J. Tang, Institute for Biophysics, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria
T. Uchihashi, Department of Physics, Kanazawa University, Kakuma-machi, Kanazawa 920-1192 Japan, and CREST, JST, Sanban-cho, Chiyoda-ku, Tokyo 102-0075, Japan
M. T. Valentine, Department of Mechanical Engineering, University of California, Santa Barbara, CA, USA
A. van Oijen, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
I. D. Vilfan, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands
W. Wang, Department of Physics, Harvard University, Cambridge, MA 02138, USA
M. Wanunu, Department of Biomedical Engineering and Department of Physics, Boston University, Boston, MA 02215, USA
T. Weidemann, BIOTEC, Biophysics, Technische Universität Dresden, Tatzberg 47–51, 01307 Dresden, Germany
W. P. Wong, The Rowland Institute at Harvard, Harvard University, Cambridge MA 02142, USA
M. T. Woodside, National Institute for Nanotechnology, National Research Council of Canada, and Department of Physics, University of Alberta, Alberta, Canada
J. Xiao, Johns Hopkins University School of Medicine, 708 WBSB, Baltimore, MD 21205, USA
A. J. Xu, MD/PhD program, HMS-MIT Division of Health Sciences and Technology, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA
A. Yildiz, Departments of Physics and Molecular and Cell Biology, University of California, Berkeley, CA, USA
V. I. Zarnitsyna, Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
X. Zhang, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China
C. Zhu, Coulter Department of Biomedical Engineering and Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
X. Zhuang, Departments of Physics, Chemistry and Chemical Biology, Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138, USA

1

Single-Molecule Fluorescent Particle Tracking

Ahmet Yildiz

Abstract One of the most fascinating processes in biology is the directed movement of organisms, subcellular compartments, and single proteins. Tracking the cellular motion is of great interest to single-molecule biophysicists to understand the mechanism of wide variety of biological processes, from basic mechanism of molecular machines to protein–protein interactions. In the last two decades, random diffusion of proteins and lipids has been tracked under the fluorescence microscope to understand how they associate with their targeted molecules. However, cellular motility is not limited to diffusion of small particles. Many fundamental processes occur by discrete physical movements upon each enzymatic cycle. For example, motor proteins of cytoskeleton can transport intracellular cargoes by taking nanometer-sized steps along the linear tracks within the cell. Several high precision techniques have been developed to understand the working principles and kinetics of motors in a detailed manner. This chapter summarizes the recent advances in fluorescence microscopy techniques that allow high precision tracking of biological molecules.

1.1. The History of Single-Particle Tracking

Various microscopy techniques were developed to track the movement of single organelles and proteins to understand important biological function. Video microscopy [1] allowed acquisition of consecutive images by a camera to observe the dynamic movement in time series. Optical density of the material and light scattering were utilized in bright-field and dark-field microscopy, respectively, to visualize large organelles (mitochondria) or single macromolecules (microtubule filaments). Differential interference contrast (DIC) microscopy can obtain finer detail by utilizing the difference in index of refraction of the molecule from that of its environment. The interference of two polarized beams that take slightly different paths was used to generate contrast within the specimen.

Although conventional microscopy allows imaging of a "large object," single proteins and many organelles that are smaller than the diffraction limit of the light are still invisible since they do not have sufficient light scattering or optical density. One way to track single

A. Yildiz • Departments of Physics and Molecular and Cell Biology, University of California, Berkeley, CA, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_1,

molecules is to attach them to a larger probe, for example, a polystyrene bead with a high index of refraction. Gelles et al. tracked the movement of a single kinesin motor bound to 190-nm-size beads with a few-nanometer precision by using video-enhanced DIC microscopy [2]. Svoboda et al. trapped a micron-sized bead with a focused laser beam to minimize its Brownian motion and revealed individual 8-nm steps of kinesin motors along microtubules [3]. To achieve high precision in bead tracking, the diameter of the bead needs to be larger than ~100 nm to obtain high precision. The size of the bead can potentially impose perturbations on the molecule of interest when studying single proteins, especially *in vivo*. Therefore most single molecule applications require usage of smaller probe size to monitor protein movement without affecting its function.

Development of fluorescent probes has opened a new era in particle tracking. Fluorescence imaging is achieved by exciting the fluorophores with high-energy photons (lower wavelength) and collecting their emission of low-energy photons (higher wavelength) with a detector. The size of the fluorophores is usually small (~1 nm across), and reactive forms are commercially available for labeling proteins and nucleic acids. A fluorescence signal can make "invisible" single molecules traceable *in vitro* and in living cells. However, single dyes are challenging to observe because the signal-to-noise ratio (SNR) is usually poor when excitation and imaging are done with an epifluorescence microscope (Figure 1.1).

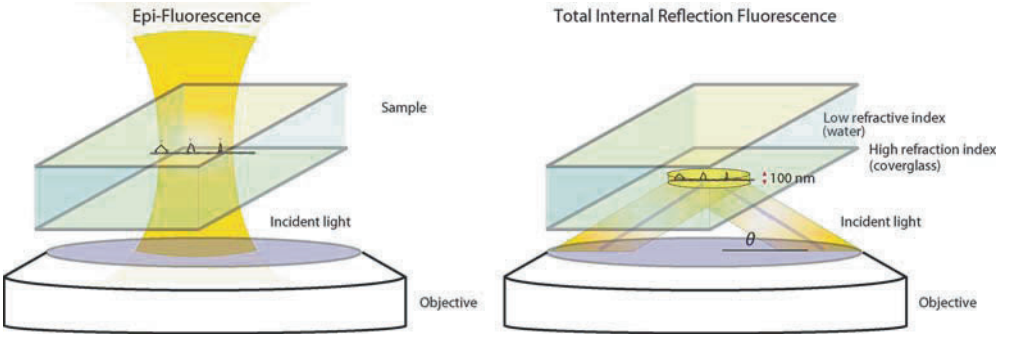


Figure 1.1. Epifluorescence versus total internal reflection (TIR) illumination. In epifluorescence illumination **(left)** the laser beam (*light green*) penetrates into the sample and excites the fluorescence molecules within the whole depth of the water. In TIR **(right)**, a parallel laser beam is sent through the sample with an incident angle θ greater than the critical angle θ_c [$\theta_c = \sin^{-1}(n_2/n_1)$, where n_1 and n_2 refer to the high– and low–index-of-refraction media, respectively]. Total internal reflection from the n_1 to the n_2 medium creates an evanescent field (*yellow disk*) that decays exponentially from the interface into the water (n_2) within ~150 nm. Only the molecules located near the surface get excited. In both microscopes, fluorescence emission from the sample is collected with the same objective and focused onto a charge-coupled device camera (not shown).

Funatsu and coworkers showed that single dye molecules can be detected by a camera when the background fluorescence from out-of-focus dyes or stray objects is eliminated [4]. They utilized total internal reflection (TIR) of light to minimize the background noise (Figure 1.1). In TIR, a collimated laser beam is sent through the sample at high incidence angle and totally internally reflected at the glass–water interface. This creates an evanescent field that has an intensity decaying exponentially as the distance from the glass–water surface increases (~150 nm deep). TIR illuminates a wide (on the order of 100 μm^2) disk inside the water and allows simultaneous detection of hundreds of molecules with a camera without the need for an image scan. Combination of wield-field excitation with low background makes TIR superior to the confocal microscope for time-resolution-demanding applications.

By only illuminating 150 nm thickness of water, TIR minimizes the autofluorescence from the stray objects and offers high SNR imaging at low light levels. These advances in single-molecule fluorescence have allowed the observations of a single substrate (adenosine triphosphate ([ATP]) and a single kinesin motor in a TIR microscope. Recently, Sakamoto et al. detected the binding of single nucleotide to myosin V and correlated it with the movement of the motor [5]. The remaining question for this chapter is whether it is possible to reach 1-nm precision with fluorescence to track single enzymatic events.

1.2. Localization in Fluorescence Microscopy

The spatial resolution of conventional fluorescence imaging has been limited to the diffraction limit of the light. The diffraction pattern of a small light-emitting object is often referred to as a point spread function (PSF). For a well-optimized microscope, the width of the PSF is approximately $\lambda/[2(\mathrm{NA})]$, where λ is the wavelength of light and NA is the numerical aperture of the collection lens. By collecting the light with high NA (1.40–1.49) oil objectives, this width is typically 200 nm in the visible region of light. Therefore, two objects separated less than this distance are irresolvable if they emit fluorescence at the same wavelength. However, if there is only a single light-emitting object within the PSF, it can be localized with better precision simply by determining the center of the PSF, which represents the position of the molecule under favorable conditions.

The shape of the PSF on a two-dimensional detection plane through a circular aperture is called the Airy pattern (Figure 1.2A). If a is the aperture radius, q is the distance from the emission center, and R is the distance from the aperture to the detection plane, then $q/R = \sin(\theta)$ and the intensity (I) can be written as a function of θ:

$$I(\theta) = I(0)\left[\frac{2J_1(ka\sin(\theta)}{ka\sin(\theta)}\right]^2 \tag{1.1}$$

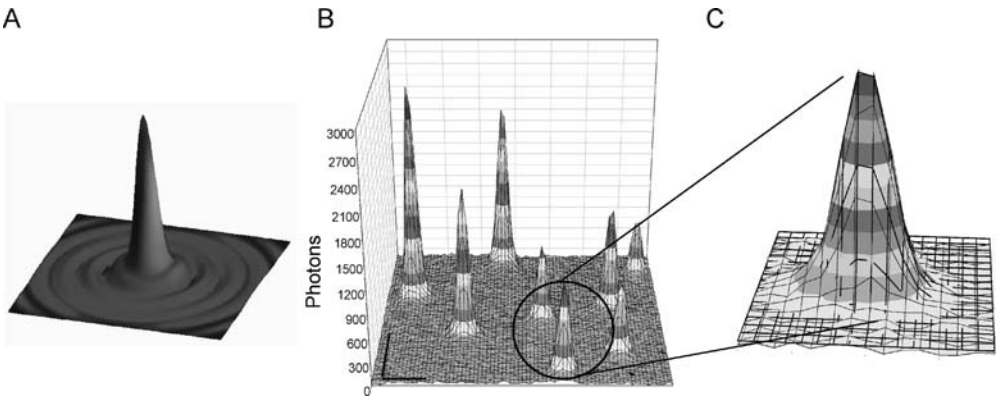


Figure 1.2. A. The Airy pattern of a diffraction-limited spot in two dimensions. **B.** Fluorescence images of several single Cy3-DNA molecules immobilized on a glass surface. The data were taken with a total internal reflection fluorescence microscope in 0.5 sec. **C.** Expanded view of one point spread function (PSF) with two-dimensional elliptical Gaussian curve fit (*solid lines*). The center of this PSF can be located to within 1.5 nm (σ_μ). (From Yildiz et al. [14].)

where J_1 is a Bessel function of the first kind and k is a propagation vector [6]. The central maximum contains 85% of the signal. In single-molecule fluorescence microscopy, the other rings of the pattern are hard to observe since they are dominated by the background noise. To determine the position of the probe, a two-dimensional Gaussian was shown to be a better estimate than Airy function or centroid analysis to fit images of single fluorescent dyes [7].

The standard error of the mean (s.e.m., σ_μ) of the distribution determines the ability to localize a single molecule. In the photon-noise-limiting case (no background fluorescence),

$$\sigma_{\mu_i} = s_i/\sqrt{N} \tag{1.2}$$

where s is the standard deviation of a Gaussian distribution and N is the total number of collected photons from the object. Background fluorescence, dark current, and readout noise of the camera add noise to the detection system. Moreover, "pixelation" of the PSF adds uncertainty of where the photons arrived within the pixel. Thompson et al. derived the s.e.m. for a two-dimensional Gaussian:

$$\sigma_{u_i} = \sqrt{\left(\frac{s_i^2}{N} + \frac{a^2/12}{N} + \frac{8\pi s_i^4 b^2}{a^2 N^2}\right)} \tag{1.3}$$

where a is the pixel size and b^2 is the background [8]. To obtain the best precision for any given b and N, one can calculate the ideal pixel size a, which is usually in the vicinity of s ($\sim$100 nm for a typical PSF). The authors also experimentally showed that the PSF center can be localized within 2 nm by using 30-nm-sized fluorescent beads. The precision of localization can be improved by increasing the signal and decreasing the noise, and the s.e.m. can theoretically be made arbitrarily small without a theoretical limit.

1.3. Higher Signal, Lower Noise

Achieving high-SNR images of single fluorophores is not trivial. Organic dyes are ideal for fluorescent tagging of proteins because of their small (<1 nm) size and specific labeling ability. However, they are not bright enough to generate the number of photons required for nanometer-precision images. Moreover, most fluorophores photobleach within a few seconds, which is not long enough for most tracking applications. It is therefore critical to choose the right probe for applications that demand high signal levels and photostability. Cyanine and rhodamine dyes actively emit fluorescence for a reasonably longer period ($\sim$5–10 sec) with a high quantum yield (the ratio of emitted photons to absorbed photons). Quantum dots are significantly brighter and more photostable than organic dyes, but commercial dots for biological applications are as big as 10–40 nm. Their size can potentially hinder a protein's function, and it is not trivial to deliver quantum dots inside the cell. Here, I focus on the recent work related to tracking of organic dyes, and the advances in probe development will be discussed at the end of this chapter.

In the last decade, the diffusion of individual rhodamine dyes in a lipid bilayer [9] and green fluorescent proteins (GFP) in a viscous solution [10] was tracked with 30-nm precision by epifluorescence microscopy. To further improve the precision of localization, one needs to extend the photostability and brightness of dyes. For this purpose, oxygen-scavenging enzymes (i.e., catalase and glucose oxidase mixture) have been used to remove free oxygen

from the solution [11]. Oxygen collides with the molecule in its excited state and undergoes a chemical reaction with the dye, resulting in permanent photobleaching. Oxygen scavengers make it possible to watch a single organic dye on the order of 10 sec. Removal of oxygen, however, comes with a consequence. Dye molecules can sometimes be trapped in triplet state upon spin conversion of the excited electron, and relaxation from the triplet state to the ground state occurs on the order of 1 sec, resulting in a temporary blinking of the dye [12]. Since oxygen's ground state is triplet, it can readily interact with the dye to quench the triplet state. To compensate the positive role of oxygen, β-Mercaptoethanol (BME) and dithiothreitol (DTT) are used as alternative quenchers to achieve steady light emission without blinking [13]. In addition, various salt concentrations, the pH, the buffer, the dye's local environment, and conjugation to a DNA or to a protein (specific to each dye) can alter the photostability. By optimizing these conditions for Cy3-labeled DNA molecules, the Selvin's laboratory detected ∼2 million photons from a single Cy3 before photobleaching [14]. Rasnik et al. used Trolox (6-hydroxy-2,5,7,8- tetramethylchroman-2-carboxylic acid) as an alternative quencher and extended the stability of Cy5 dyes tenfold [15].

To sensitively detect fluorescence with minimal background, the sample is excited with a TIR fluorescence (TIRF) microscope and photons are collected with high-NA (1.45–1.65) objectives. Selection of the right bandwidth of dichroic mirrors and emission filters is essential to efficiently collect fluorescence and remove background noise, and >95% transmitting band pass filters are commercially available. To achieve high gain without spreading the image, Electron-multiplied (EM)-CCD cameras are developed. Back-illuminated charge-coupled device (CCD) cameras can collect ∼92% of incoming photons with 8-e^- rms noise at slow readout speeds. To minimize the detector noise, the signal is intensified before photons reach the CCD chip. While intensified CCDs can yield images with an extremely low amount of light, they have low spatial resolution because the photons are spread in the image intensifier. Electron-multiplied (EM)-CCD cameras can multiply the signal after photons reach the camera pixels. Charges are multiplied in EM registry up to 1,000-fold by applying high clock voltages, similar to the design of photomultiplier tubes. Readout noise becomes negligible even at higher acquisition speeds. Therefore, the combination of a TIRF system with EM-CCDs can perform at nearly the fundamental photon-noise limit of detection even at extremely low light emissions.

1.4. Fluorescence Imaging with One-Nanometer Accuracy

By combining TIRF excitation and a highly sensitive fluorescence detection system, Yildiz et al. localized the position of surface-immobilized Cy3-DNA molecules with 1.5-nm precision [14]. Figure 1.2B shows an image of individual diffraction-limited spots taken by a CCD camera. Each spot corresponds to a single dye since the sample is sparsely coated with Cy3-DNA molecules, and their fluorescence intensity shows a single-step photobleaching. Figure 1.2C shows an expanded view of one PSF overlaid with a curve-fit to a two-dimensional Gaussian (solid lines). Altogether, 14,000 photons were collected within 0.5 sec, and the Gaussian fit estimated 1.3-nm localization accuracy. The standard deviation calculated from 100 consecutive images was slightly higher (1.4 nm) than the estimated error, which may correspond to vibrational noise or thermal fluctuations of the microscope. Cy3 emission was highly stable and lasted up to 4 minutes at a rate of ∼20,000 photons/sec. The technique was named fluorescence imaging with 1-nm accuracy (FIONA) and improved the

localization accuracy of single fluorophores ~20 fold [9,10] and provided longer tracking time than earlier studies [16,17].

FIONA is complementary to other high-precision single-molecule techniques, such as fluorescence resonance energy transfer (FRET) and optical trapping. The most precise optical traps can measure down to angström-scale motion of surface-bound molecules attached to a large bead with ~100-μs intervals, significantly better than what FIONA can achieve. On the other hand, optical traps can track one molecule at a time and they need to apply forces to decrease the compliance between the bead and the protein. In addition, the size of the large bead used limits their application to certain enzymes, and the motion of small subdomains of a protein cannot be readily detected. Single-molecule FRET measures neither the absolute distance nor distances larger than 10 nm. It is a more suitable technique than FIONA for studying small conformational changes within a single protein or protein–protein interactions with subnanometer precision and up to millisecond time resolution. The details of optical traps and FRET can be found in Chapters 12 and 5, respectively.

1.5. Tracking the Movement of Molecular Motors

Myosin, kinesin, and dynein motor proteins transport organelles and receptor molecules along the linear tracks of the cytoskeleton by hydrolyzing ATP [18]. Although these motors are made of different mechanical elements and move along different tracks and directions, they represent remarkably similar features. By oversimplifying their structure, one can say that these motors are "world's tiniest bipeds" [19], having two feet connected by two tiny legs to a common body. The motors use their feet (referred as motor heads) to bind actin/microtubule filaments and hold their cargo with their arms located at the tip of their body. Single motors can move a cargo along several microns, and they advance along their track by taking 8- to 36-nm steps.

Although optical traps were highly useful in the characterization of motor stepping, the question of how motors take a step remained unclear. Specifically, it was not clear how the two heads are coordinated to allow the motor to move processively without falling off its track. There were several models proposed to explain their motion (Figure 1.3A). In *the hand-over-hand model* [20], ATP hydrolysis creates a conformational change in the leading head that pulls the rear head forward. In the next step, the heads swapped the lead, and the other head undergoes the same cycle to pull its partner forward. *The inchworm model* [21] proposed that only one head catalyzes ATP and always leads while the other head follows. The most direct way to distinguish among these models is to track how the heads move. The hand-over-hand model predicts that a head alternately moves twice the stalk displacement and stays stationary in the next step while the other head takes a step. In contrast, the inchworm model predicts that both heads move forward as much as the stalk moves.

By labeling the individual heads of the myosin V, which has a 37-nm center-of-mass step size [22], Yildiz et al. observed that the myosin V head takes alternating 74-nm/0-nm steps (Figure 1.3B) [14]. This result demonstrated that myosin V walks hand over hand with its rear head taking twice the step size of the motor as the front head stays stationary. Recent experiments using FIONA, optical trapping, and mutational assays showed that myosin VI [23, 24], kinesin [25–27] and cytoplasmic dynein [28] motors also utilize a hand-over-hand mechanism to move along the cytoskeleton.

FIONA has also been applied to other systems. For example, Xiao et al. labeled multiple sequence motifs in a lambda DNA, stretched it on a glass coverslip, and determined the

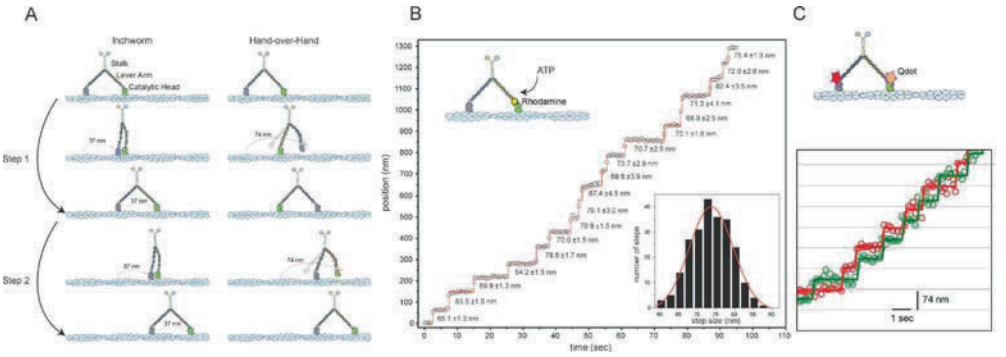


Figure 1.3. A. Hand-over-hand versus inchworm model for myosin V. The hand-over-hand and inchworm models make different, testable predictions for the motions of each individual head. In the inchworm model, one of the heads always leads and the other follows. The step size of each catalytic domain is therefore equal to the step size of the stalk (∼37 nm). However, in the hand-over-hand model, the trailing head (*blue*) takes a step that is twice the step size of the stalk (∼74 nm), whereas the leading head does not move. In the second step of motor, the blue head stays bound to actin as the green head takes a large step and rescues the leading position. **B.** The myosin V light chain domain was labeled with a single rhodamine dye (*yellow circle*) near the catalytic head. The movement of labeled motors on a surface-immobilized actin was tracked with fluorescence imaging with 1-nm accuracy (FIONA). Individual heads of myosin V take 74-nm steps, in agreement with the hand-over-hand model. (Figure from Yildiz et al. [14].) **C.** Both catalytic heads can simultaneously be tracked with single-molecule high-resolution colocalization (SHREC). The positions of the green and red colors indicate that the two heads alternately take a step and swap the lead. (From Warshaw et al. [35].)

position of these sites by localizing the position of labeled fluorophores [29]. They proposed that a similar approach can be used as a barcode reader for specific sites to recognize human pathogenic DNA.

1.6. Multicolor Fluorescent Tracking

Fluorescent probes can also be used to measure the distance between two specific sites on a DNA or a protein complex. FRET is widely used as a spectral "ruler" by measuring the distance between the acceptor and donor probes [30]. However, FRET does not measure the absolute distances and cannot measure distances larger than 10 nm. Similarly, single-color FIONA cannot resolve two identical fluorophores that are less than $\sim\lambda/2$, or ∼250 nm, apart. However, if one of the dyes photobleaches permanently, the change in image position is used to determine the projection of their separation onto the image (*xy*) plane. Single-molecule high-resolution imaging with photobleaching (SHRImP) is capable of measuring the separation between two (or more) identical dyes with <10 nm precision [31,32]. However, its reliance on photobleaching of one of the two dyes has limited its use to static measurements.

Multicolor imaging can be used to track two proteins (or two sites in a single protein). Lacoste et al. showed that the distance between two nanocrystals of different emission wavelengths placed on a coverslip can be measured with a precision of ±6 nm using a total integration time of 20 sec [33]. Two-color FIONA, referred as single-molecule high-resolution colocalization (SHREC) [34], delivers all of the advantages of FIONA, FRET, and SHRImP because the translational motion of both probes and their separation on the *xy* plane are measured simultaneously. To simultaneously image two spectrally separable probes (i.e., Cy3 and Cy5), the sample is excited with green and red laser beams in an objective-type TIRF

microscope. The fluorescence is then separated by a dichroic mirror that reflects Cy3 fluorescence (orange) and transmits Cy5 fluorescence (red). The two fluorescence channels on a CCD chip are then registered to one another by using the coordinates of surface-immobilized beads as reference points, with ~10-nm agreement. Churchman et al. used SHREC to measure the distance between the ends of a DNA segment and to track the movement of the two heads of myosin V dimers containing Cy3- and Cy5-labeled calmodulin [34]. They observed that Cy3 and Cy5 are separated by 37 nm, and the rear head takes 74 nm (twice the center-of-mass step) and becomes the leading head as the front head stays bound to actin. Warshaw et al. [35] reported similar results in experiments with myosin V labeled on the heads with two spectrally separable quantum dots with better time resolution (Figure 1.3C).

1.7. Tracking Fluorophores inside Living Cells

Most single-molecule experiments have been carried out *in vitro* with purified proteins on surface-modified coverslips and controlled substrate concentrations. Isolation of the assay from the crowded environment of the cell is essential to obtain utmost sensitivity. However, it is of great interest to researchers to figure out how these proteins function inside living cells. Specifically, *in vivo* studies are critical to understand how the collective behavior of identical proteins affect a cell's function, how different proteins interact with each other, and how their activity is controlled by the cell-cycle and regulatory enzymes. Tracking single molecules in live cells remains a challenge for multiple reasons. First, the autofluorescence of the cell is usually well above the signal of the brightest organic dyes, which prevents visualization of single molecules. TIRF microscopy is commonly used to limit the excitation volume, and cells are immobilized on a glass surface treated with concavalin A or other adhesion proteins to keep the cells as flat as possible. The spinning disk confocal approach is an alternative low-background imaging technique for visualizing the fluorescent signal in the cell with a moderately fast acquisition. However, as a significant part of the fluorescence signal is lost when passing through the spinning disk, this technique is not ideal for single-molecule imaging at low light levels.

Another challenge is to specifically label the protein of interest in a crowded environment of a cell without affecting cell's function. This can be readily achieved by fusing a green fluorescent protein (GFP) sequence into the gene that encodes the protein in fly and yeast cells, but the GFP signal is even lower than that of organic dyes like rhodamines and cyanines. For specifically labeling the proteins inside the cell with functional probes, genetic tags (the HaloTag and the SNAP tag) have recently become commercially available, and these tags can specifically be labeled by membrane-permeable reactive dyes. Monoclonal antibodies attached to multiple dyes seem ideal for specific labeling in higher organisms. Other strategies can be microinjection, electroporation, or endocytosis of fluorescent probes into the cell.

Earlier tracking of molecules in living cells was limited to measuring two-dimensional diffusion in a lipid membrane. By measuring the intensity and diffusion constant of GFP-tagged transmembrane proteins at 10- to 50-msec image acquisition rates, cluster formation of adhesion molecules [36], Ca^{2+} channels [37], and anchoring proteins [38] in a plasma membrane have been shown experimentally. Signaling proteins in T cells were visualized by multicolor imaging (Figure 1.4), and microdomain formation of these proteins has been proposed to facilitate T cell signaling [39]. Counting the number of photobleaching steps helped Isacoff and coworkers to demonstrate the oligomeric states of active ion channels [40,41]. Yanagida and coworkers tracked Cy3-labeled receptor molecules and showed that the

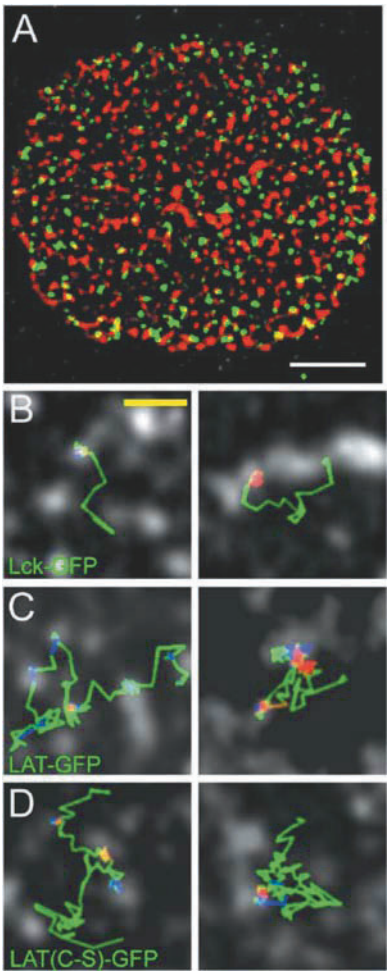


Figure 1.4. Multicolor tracking of single molecules in a cell membrane. **A.** Superimposed images of fluorescently tagged T cell receptors in Jurkat cells. **B–D.** Diffusion of green fluorescent protein (GFP)–tagged individual receptors Lck, LAT, and LAT(C-S) is tracked by a total internal reflection fluorescence microscope. (From Douglass and Vale [39].)

activation of epidermal growth factor (EGF) receptor requires dimer formation and binding of a single EGF molecule to each receptor [42]. To achieve higher signal-to-noise ratio, Dahan et al. imaged highly bright single quantum dots (15–30 nm in size) labeled to individual glycine receptors when they diffused on the neuronal membrane in a living cell [43]. They obtained ∼5-nm localization with continuous images of 75-msec integration in each 60-sec duration and tracked the receptors over 20 min.

In addition to the observation of receptor diffusion in a lipid membrane, proteins, RNA transcripts, and viruses were tracked inside the cytoplasm. For example, Zhuang and coworkers tracked the movement of an influenza virus labeled with organic dyes [44,45]. The movement of RNA transcripts was tracked by attaching a GFP fusion protein MS2 to mRNA molecules [46]. Golding et al. detected the transcriptional bursts of mRNA signal in living

cells [47]. Yu et al. probed the expression of single fluorescently labeled proteins real time in bacteria [48]; details of these experiments are explained in Chapter 3.

Most *in vitro* tracking experiments have been carried out at limiting substrate concentrations (i.e., ATP) to slow down the speed of the enzymatic cycle. However, it is not trivial to control the ATP concentration within a cell. Moreover, the rapid diffusion of molecules in three dimensions and the flexibility and dynamic instability of cellular structures are other challenges related to the need for faster image acquisition. The collection of more than 10,000 photons is needed for 1-nm localization, and that usually takes hundreds of milliseconds for cyanine and rhodamine dyes. To push the limits of time resolution *in vivo*, brighter probes should be used. Kural et al. labeled peroxisomes, a cargo carried by motor proteins, with thousands of GFP molecules in *Drosophila* S2 cells and tracked the movement of this single macromolecule with 1-msec time resolution along microtubules [49]. The authors demonstrated that cargoes are moved by kinesin and dynein motors in forward and backward directions on microtubules with 8-nm increments. In a subsequent study, same authors tracked dark particles (i.e., pigment granules) in a cell with 1-nm precision by using a bright-field microscope [50]. This experiment did not require any specific fluorescent probe attachment, but the particles needed to be large. Kural et al. directly observed, by measuring the stepwise trajectory of the cargo transport, that pigment granules switch from microtubule motors to actin motors (Figure 1.5A). A Gaussian fit gives reliable results if the object is markedly smaller than the diffraction limit of light. For larger objects (such as mitochondria), the shape and size of the object need to be taken into account [51]. The groups of Xie [52] and Higuchi [53] tracked endocytosed vesicles containing quantum dot aggregates and measured stepping

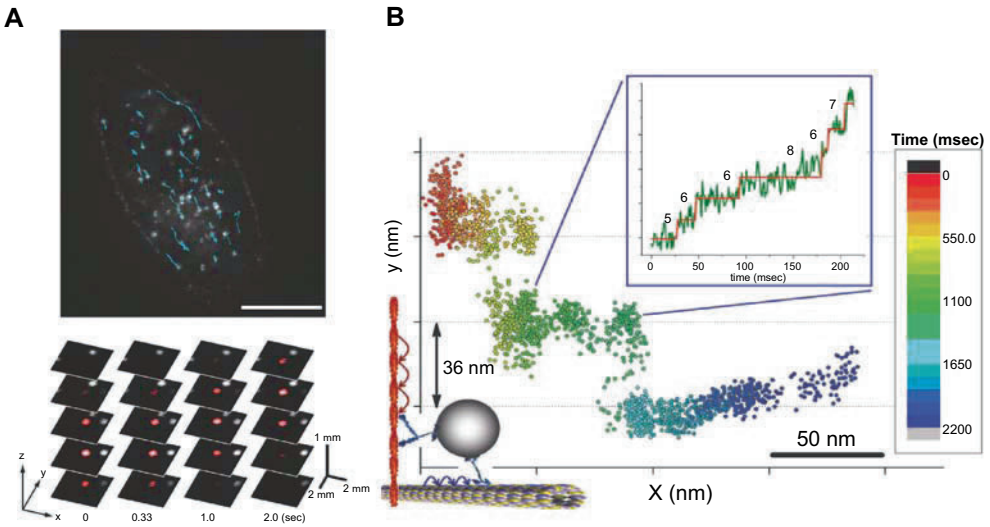


Figure 1.5. Tracking motor movement inside the cytoplasm. **A.** Vesicles containing multiple quantum dots (*bright spots*) are endocytosed into a cell. Transport of these vesicles can be tracked with a few-nanometer precision. Vesicles can be tracked in three dimensions by moving the objective up and down along the optical axis with a piezoelectric stage. (Figure from Watanabe and Higuchi [53].) **B.** Transition between actin- and microtubule-based transport can be observed directly by tracking pigmented granules in a bright-field microscope. As the cargo is moved along the *x* direction by microtubule motors, myosin motors pull the cargo along an actin filament in the *y* direction. (From Kural et al. [50].)

movement of motors with 300-μs time resolution (Figure 1.5B). The speed of these measurements was limited by the readout of the CCD camera. Later, Xie's lab improved their temporal (25 μs) resolution by detecting the scattered light from 100- to 150-nm-diameter gold nanoparticles with a quadrant photodiode [54]. Microsecond time resolution allowed clear detection of individual stepping of kinesin and dynein motors moving with speeds up to 8 μm/sec.

1.8. Rotational Movement

Single-fluorescent-particle tracking assumes that the dye acts as a point source and yields a symmetric PSF in far field. In reality, fluorophores have an excitation and emission dipole moment (Figure 1.6A) and they usually prefer fluorescence emission polarized along their long axis of aromatic chains. The probability that incoming photons will be absorbed by a fluorophore is given by $\cos^2\phi$, where ϕ is the angle between photon polarization and the absorption dipole. If the probe has high rotational mobility tethered around the covalent bond, the orientation effect averages out during each image acquisition and the dye can be accepted as an isotropic emitter. For a fixed dipole, anisotropic dipole radiation can cause up to 10-nm error in position [55]. For localization experiments, it is therefore essential to ensure that the dye emission is unpolarized during the time course of image acquisition.

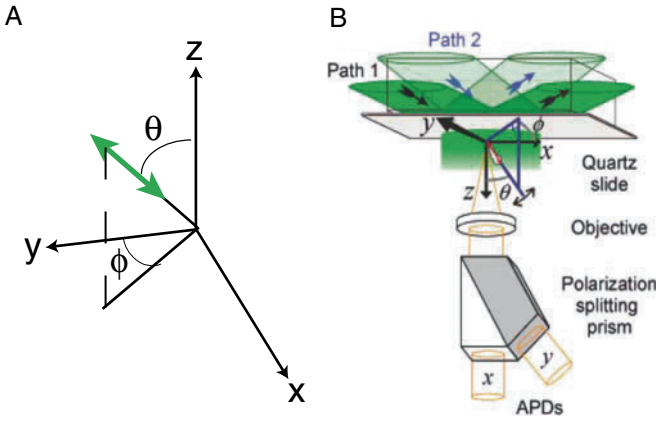


Figure 1.6. A. In-plane (ϕ) and out-of-plane (θ) angles of a fluorophore dipole moment (*green arrow*) in *xyz* coordinates. **B.** Single-molecule fluorescence polarization (SMFP) microscope. The sample is excited with two orthogonal laser beams (paths 1 and 2), each of which can switch between *s*- and *p*-polarized excitation. The fluorescence emission is separated onto two avalanche photodiodes by a polarizing beam splitter cube. (From Rosenberg et al. [78].)

As a fixed dipole constitute a problem for fluorophore localization, its orientation can be tracked to study the rotational movement of single enzymes. Nonprocessive motors, DNA enzymes, and ion channels present rotational movement as they undergo conformational changes to function in a cell. For example, F_0F_1 ATPase was proposed to be a rotary motor. To test this idea, Kinosita and coworkers imaged a gold nanoparticle attached to F_1 and demonstrated that its γ subunit rotates 120° counterclockwise on each ATP synthesis [56].

To monitor the rotational dynamics of fluorescently labeled macromolecules, samples can be excited with linearly polarized light. The changes in fluorescent intensity would represent the reorientation of an enzyme (different ϕ) if the fluorophore did not undergo rapid

rotational movement independent of the targeted molecule. The polarization of single dye molecules was first measured when they were adsorbed into glass [57]. Ha et al. studied the rotational dynamics of short DNA molecules tethered to a glass surface [58]. The fluorescence intensity as a function of excitation polarization was recorded to determine the in-plane angle of the fluorophore, and rapid changes in the intensity were observed real time in response to the change of the dye's orientation. The signal was collected with a single detector, and the fluctuations of fluorescence intensity due to dye blinking could not be easily distinguished from its rotation. An alternative strategy is to excite the sample with circularly polarized light and to record the polarized emission with two detectors [59]. The ratio of the intensities of the two polarized emissions gives the in-plane angle, and it is insensitive to dye blinking. However, these methods cannot distinguish the ϕ from $-\phi$ because the cosine squared would yield the same value.

To circumvent the ambiguity of angular degeneracy, Sosa et al. [60] alternated the linear polarization of the laser light excitation to four axes (0°, 45°, 90°, 135°). In this setup, ϕ and $-\phi$ were separated from each other because they yield different intensities at 0° and 45°. Fixing the polarization of the laser excitation to well-defined orientations increased the time resolution of the measurements. The authors attached a single dye to the kinesin head domain and monitored its mobility as kinesin was bound to a microtubule. To minimize the rotational mobility of a fluorophore independent of the head orientation, they used a bifunctional rhodamine (Br) attached to two closely spaced ($\sim$1 nm apart) cysteines. Based on its nucleotide state, kinesin either displayed a high fluctuation of fluorescence intensity, which was interpreted as the labeled head is detached from the microtubule, or nearly constant fluorescence when the head is attached [60]. Subsequent studies demonstrated the orientation of kinesin's two heads [61] and the neck-linker domain [62] that interconnects the two heads together.

These optical setups are insensitive to the out-of-plane angle (θ) of the dye because both excitation and detector polarization are in the *xy* plane. To determine the three-dimensional orientation of a fluorophore, Goldman and coworkers (Figure 1.6B) excited the sample with two orthogonal incident beams in a prism-type TIRF microscope [63]. Each beam can switch between horizontal (*s*) and vertical (*p*) polarizations. The *s*-polarized excitation creates *x* and *y* evanescent waves, depending on which beam was used to excite the sample. The *p*-polarized excitation beams produce *z*-polarized evanescent waves with weak component on the *x* or *y* axis (along the direction of the excitation beam). The emission pathway was also split into two channels (onto avalanche photodiodes) by a polarizing beam splitter cube. The in-plane (ϕ) and out-of-plane (θ) angles of the probe were tracked with 40-msec time resolution. Using a single-molecule fluorescence polarization (SMFP) microscope, the authors observed that myosin V light chain alternates between two well-defined orientations as the motor moves along actin. The results provided evidence for hand-over-hand movement of myosin V. To track many dyes at a time with SMFP, Syed et al. used a multichannel CCD. The changes in fluorescence intensity were recorded to monitor rotational movement in the *xy* plane, and fluorescent images were fitted with a two-dimensional Gaussian to relate the tilting of myosin V light chain and the stepping of the motor [64]. The authors observed that tilting of the Br-labeled lever arm coincided with the stepping of the myosin V rear head.

The three-dimensional (3D) orientation of the probes without angular degeneracy can readily be monitored by imaging the fluorescence intensity away from the focal plane. Bartko et al. observed that frozen DiI molecules result in a combination of lobes and fringes in defocused images [65]. Enderlein and coworkers calculated the position-dependent intensity distribution of the fluorophore based on its orientation (θ, ϕ) and its distance from the focal plane, *z* [66]. Images of single rhodamine molecules can yield 10°–15° of precision in orien-

tation and 15 nm in lateral positional accuracy. Toprak et al. combined defocused orientation with position imaging (DOPI) to observe true 3D orientation of Br molecules on myosin V light chain (Figure 1.7) [67]. The sample was excited with a circularly polarized laser, and the objective lens was moved 500 nm away from the focus to find (θ, ϕ). The objective was then switched back to the focal plane to achieve a nanometer position accuracy and find the

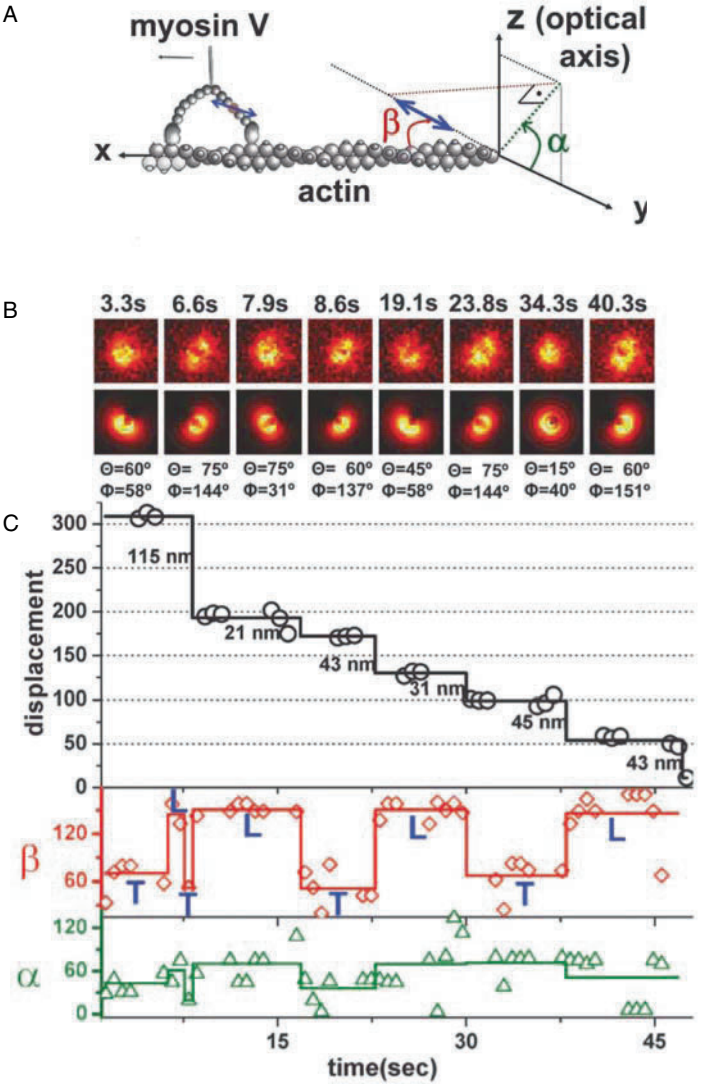


Figure 1.7. Defocused orientation with position information. **A.** Myosin V light chain was labeled with bifunctional rhodamine (Br), which has a fixed dipole moment. The orientation of a dipole (α, β) is shown with respect to the long axis of an actin filament. **B.** Defocused images of Br in time series as myosin V moves along actin **(above)**. Fluorescence images were fitted to θ and ϕ **(below)**. **C.** Objective is moved to in-focus (*black trace*) precisely to measure the dye position and to out-of-focus to obtain the dye orientation (α, β). Discrete jumps in displacement coincide with the angular orientation. The data show that the myosin V lever arm rotates as the motor takes a step. (From Toprak et al. [67].)

coordinates (x, y) of the molecule. Therefore, DOPI combined with FIONA can measure 3D rotational dynamics without an angular degeneracy and track the position of the probe with nanometer precision.

1.9. Future Directions

Recent advances in single-molecule fluorescence microscopy have had a huge effect on the understanding of structural and molecular biology. The precision of the setups is reaching near-atomic detail, but there is still much to do. The techniques described in this chapter all have practical limitations in spectral and temporal resolution, although theoretically there is no limit on how precisely the particles can be tracked.

1.9.1. Probe Development

The first limitation of these applications is the performance of fluorescent probes. Genetically encoded probes like GFP and its variants (cyan, yellow, and red fluorescent proteins [CFP, YFP, and RFP, respectively]) are ideal for immunofluorescence [68]; they are the most commonly used probes in cell biology. However, their performance is poor for high-precision tracking. Enhanced GFP (eGFP) offers a considerable increase in the quantum yield of GFP [69]. Future studies are needed to improve the performance of fluorescent proteins. As discussed earlier, organic dyes are ideal because of their small size, but producing high-SNR images takes on the order of hundreds of milliseconds. Cy3 and TMR show the best performance in light emission and photostability. However, these two dyes emit in the same wavelength region, and high-performance probes are needed that emit at shorter and longer wavelengths, especially for FRET and multicolor FIONA experiments. The total number of emitted photons from Cy5 is typically tenfold less than that from Cy3, and its performance recently been improved by using Trolox [15]. Atto647N seems a highly promising Cy5 alternative, yielding a similar number of photons to Cy3. There are multiple dyes coming onto the market every year. Based on trial and error, better-performing probes with high quantum yield can be expected in the future.

For *in vivo* tracking experiments, faster image acquisition is needed to temporally resolve single enzymatic events, and researchers have used clusters of quantum dots or organic dyes to collect a sufficient number of photons on a shorter time scale. Therefore, tracking experiments again are limited to larger particles. Quantum-dot nanocrystals are highly promising, in that they are considerably smaller (5–10 nm) and ~50-fold brighter than organic dyes. However, the semiconductor core needs to be coated with poly(ethylene glycol) to make these dots water soluble, yielding quantum dots with 30–50 nm diameter. Moreover, heavy metals in their semiconductor core are toxic to cells, and multiple reactive sites (at least ten) on their surface cross-link the enzymes to a single dot. Recently, the diameter of water-soluble quantum dots has been reduced to less than 10 nm [70] by tuning the size of the core and using smaller coating material. The surface can also be treated with the right number of reactive sites to achieve monofunctional quantum dots. The performance of these laboratory-made probes is already significantly better than that of commercial dots. Different semiconductor materials are being tried to make them even brighter, and they have been covered with alternative polymers to reach 2- to 5-nm-size dots [71]. Besides the brightness and size of quantum dots, there is room for improving shelf life, water solubility, resistance to blinking, and toxicity of these nanocrystals.

An alternative approach to obtaining higher signal is to revisit dark-field microscopy. Excitation light can be scattered from the surface of small objects with a high index of refraction. The particle acts like a point source, and scattered light forms a diffraction-limited image similar to that with fluorescent probes. These particles can be treated with biological material to conjugate them to the protein of interest. Dunn and Spudich used light scattering from small (40 nm) gold nanoparticles and observed the myosin V lever arm movement with 40-μs time resolution [72]. In general, the amount of signal that can be obtained from light scattering depends on the size and the power of the excitation laser, and there is no problem of photobleaching. Smaller crystals and more powerful lasers can be used to improve the performance of the dark-field microscope.

1.9.2. Instrumentation

More light can be detected from fluorophores by improving the efficiency of microscopes. Inverted microscopes collect only a certain fraction of light. Hess's laboratory collected fluorescence with two objectives facing toward each other in the optical axis and combined their pathway on a single detector (4Pi microscope) [73]. The result was a doubling in efficiency of light collection and sharper images ($\lambda/4NA$) on the z axis. The numerical aperture of objectives can be further improved to collect light at higher angles and obtain sharper images. Objectives with 1.65 NA are commercially available, but they require high–index-of-refraction (n) sapphire coverslips, significantly increasing the cost of each sample. Polymer fluid that matches the n of sapphire should be used instead of oil between the objective and the coverslip, and commercial fluids dry out quickly and expel carcinogenic evaporates. Affordable high-n coverslips and more user-friendly fluids are necessary for practical use of these objectives in biology laboratories.

EM-CCDs dominate the market for low-light-level collections at visible wavelengths (400–700 nm). These multichannel detectors can provide >90% photon collection efficiency with minimal readout noise. However, electron multiplication is a probabilistic event that adds an uncertainty to the number of counts per pixel. This is called an *excess noise factor*, which multiplies photon shot noise by 1.4. In addition, high clock voltage produces secondary photons that can be observed as sharp spikes in the camera. Therefore, the current state of EM is beneficial for low-SNR imaging, but at high SNR (>10), EM increases the noise in the system. The time resolution of CCD cameras is now much faster than video rates, but this speed is no longer considered fast, even for single-molecule imaging. EM cameras with higher readout rates and lower noise are highly anticipated in the near future.

1.9.3. Beyond the Diffraction Limit

The resolution of optical microscopy is still limited by the diffraction limit. Localization ability can be very high (1 nm), but two fluorophores cannot be resolved if they are closer than 200 nm. The resolution can be improved by using different colors (SHREC) or permanent photobleaching (SHRImP). Recently several novel methods have been developed to resolve objects closer than the diffraction limit. To obtain ~10-fold sharper images than the diffraction limit in far field, the stimulated emission depletion microscope (STED) has been developed [74] by saturating the optical transitions of the fluorophore. Structure illumination microscopy can introduce the highest possible frequencies of a sinusoidal excitation pattern and allows double the resolution of conventional microscopes [75]. Finally, randomly

switching fluorophores on and off within the diffraction volume can be used to resolve the position of each molecule when only one of the fluorophores is on [76,77]. The details of superresolution imaging are discussed in Chapter 4.

References

1. Horio, T., and H. Hotani, Visualization of the dynamic instability of individual microtubules by dark-field microscopy. *Nature*, 1986;**321**(6070):605–7.
2. Gelles, J., B. J. Schnapp, and M. P. Sheetz, Tracking kinesin-driven movements with nanometre-scale precision. *Nature*, 1988;**331**(6155):450–3.
3. Svoboda, K., et al., Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, 1993;**365**(6448):721–7.
4. Funatsu, T., et al., Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, 1995;**374**:555–9.
5. Sakamoto, T., et al., Direct observation of the mechanochemical coupling in myosin Va during processive movement. *Nature*, 2008;**455**(7209):128–32.
6. Hecht, E., *Optics*. 4th ed. San Francisco: Addison-Wesley, 2002.
7. Cheezum, M. K., W. F. Walker, and W. H. Guilford, Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys J*, 2001;**81**(4):2378–88.
8. Thompson, R. E., D. R. Larson, and W. W. Webb, Precise nanometer localization analysis for individual fluorescent probes. *Biophys J*, 2002;**82**(5):2775–83.
9. Schmidt, T., et al., Imaging of single molecule diffusion. Proc Natl Acad Sci USA, 1996;**93**(7):2926–9.
10. Kubitscheck, U., et al., Imaging and tracking of single GFP molecules in solution. *Biophys J*, 2000;**78**(4): 2170–9.
11. Harada, Y., et al., Mechanochemical coupling in actomyosin energy transduction studied by *in vitro* movement assay. *J Mol Biol*, 1990;**216**(1):49–68.
12. Lakowicz, J. R., *Principles of Fluorescence*. 2nd ed. New York: Kluwer Academic, 1999.
13. McKinney, S. A., et al., Structural dynamics of individual Holliday junctions. *Nat Struct Biol*, 2003;**10**(2):93–7.
14. Yildiz, A., et al., Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 2003;**300**(5628):2061–5.
15. Rasnik, I., S. A. McKinney, and T. Ha, Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods*, 2006;**3**(11):891–3.
16. Adachi, K., et al., Stepping rotation of F1-ATPase visualized through angle-resolved single-fluorophore imaging. *Proc Natl Acad Sci USA*, 2000;**97**(13):7243–7.
17. Sambongi, Y., et al., Mechanical rotation of the c subunit oligomer in ATP synthase (F0F1): direct observation. *Science*, 1999;**286**:1722–4.
18. Vale, R. D., The molecular motor toolbox for intracellular transport. *Cell*, 2003;**112**(4):467–80.
19. Asbury, C. L., Kinesin: world's tiniest biped. *Curr Opin Cell Biol*, 2005;**17**(1):89–97.
20. Spudich, J. A., The myosin swinging cross-bridge model. *Nat Rev Mol Cell Biol*, 2001;**2**(5):387–92.
21. Hua, W., J. Chung, and J. Gelles, Distinguishing inchworm and hand-over-hand processive kinesin movement by neck rotation measurements. *Science*, 2002;**295**(5556):844–8.
22. Mehta, A. D., et al., Myosin-V is a processive actin-based motor. *Nature*, 1999;**400**(6744):590–3.
23. Okten, Z., et al., Myosin VI walks hand-over-hand along actin. *Nat Struct Mol Biol*, 2004;**11**(9):884–7.
24. Yildiz, A., et al., Myosin VI steps via a hand-over-hand mechanism with its lever arm undergoing fluctuations when attached to actin. *J Biol Chem*, 2004;**279**(36):37223–6.
25. Kaseda, K., H. Higuchi, and K. Hirose, Alternate fast and slow stepping of a heterodimeric kinesin molecule. *Nat Cell Biol*, 2003;**5**(12):1079–82.
26. Asbury, C. L., A. N. Fehr, and S. M. Block, Kinesin moves by an asymmetric hand-over-hand mechanism. *Science*, 2003;**302**(5653):2130–4.
27. Yildiz, A., et al., Kinesin walks hand-over-hand. *Science*, 2004;**303**:676–8.
28. Reck-Peterson, S. L., et al., Single-molecule analysis of dynein processivity and stepping behavior. *Cell*, 2006;**126**(2):335–48.
29. Xiao, M., et al., Rapid DNA mapping by fluorescent single molecule detection. *Nucl Acids Res*, 2007;**35**(3):e16.

30. Ha, T., et al., Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci*, 1996;**93**(13):6264–8.
31. Gordon, M. P., T. Ha, and P. R. Selvin, Single-molecule high-resolution imaging with photobleaching. *Proc Natl Acad Sci USA*, 2004;**101**(17):6462–5.
32. Qu, X., et al., Nanometer-localized multiple single-molecule fluorescence microscopy. *Proc Natl Acad Sci USA*, 2004;**101**(31):11298–303.
33. Lacoste, T. D., et al., Ultrahigh-resolution multicolor colocalization of single fluorescent probes. *Proc Natl Acad Sci USA*, 2000;**97**(17):9461–6.
34. Churchman, L. S., et al., Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. *Proc Natl Acad Sci USA*, 2005;**102**(5):1419–23.
35. Warshaw, D. M., et al., Differential labeling of myosin V heads with quantum dots allows direct visualization of hand-over-hand processivity. *Biophys J*, 2005. **88**(5):L30–2.
36. Iino, R., I. Koyama, and A. Kusumi, Single molecule imaging of green fluorescent proteins in living cells: E-cadherin forms oligomers on the free cell surface. *Biophys J*, 2001;**80**(6):2667–77.
37. Harms, G. S., et al., Single-molecule imaging of l-type Ca(2+) channels in live cells. *Biophys J*, 2001;**81**(5):2639–46.
38. Murakoshi, H., et al., Single-molecule imaging analysis of Ras activation in living cells. *Proc Natl Acad Sci USA*, 2004;**101**(19):7317–22.
39. Douglass, A. D., and R. D. Vale, Single-molecule microscopy reveals plasma membrane microdomains created by protein-protein networks that exclude or trap signaling molecules in T cells. *Cell*, 2005;**121**(6):937–50.
40. Kohout, S. C., et al., Subunit organization and functional transitions in Ci-VSP. *Nat Struct Mol Biol*, 2008;**15**(1):106–8.
41. Tombola, F., M. H. Ulbrich, and E. Y. Isacoff, The voltage-gated proton channel Hv1 has two pores, each controlled by one voltage sensor. *Neuron*, 2008;**58**(4):546–56.
42. Sako, Y., S. Minoghchi, and T. Yanagida, Single-molecule imaging of EGFR signalling on the surface of living cells. *Nat Cell Biol*, 2000;**2**(3):168–72.
43. Dahan, M., et al., Diffusion dynamics of glycine receptors revealed by single-quantum dot tracking. *Science*, 2003;**302**(5644):442–5.
44. Lakadamyali, M., et al., Visualizing infection of individual influenza viruses. *Proc Natl Acad Sci USA*, 2003;**100**(16):9280–5.
45. Lakadamyali, M., M. J. Rust, and X. Zhuang, Ligands for clathrin-mediated endocytosis are differentially sorted into distinct populations of early endosomes. *Cell*, 2006;**124**(5):997–1009.
46. Bertrand, E., et al., Localization of ASH1 mRNA particles in living yeast. *Mol Cell*, 1998;**2**(4):437–45.
47. Golding, I., et al., Real-time kinetics of gene activity in individual bacteria. *Cell*, 2005;**123**(6):1025–36.
48. Yu, J., et al., Probing gene expression in live cells, one protein molecule at a time. *Science*, 2006; **311**(5767):1600–3.
49. Kural, C., et al., Kinesin and dynein move a peroxisome *in vivo*: a tug-of-war or coordinated movement? *Science*, 2005;**308**(5727):1469–72.
50. Kural, C., et al., Tracking melanosomes inside a cell to study molecular motors and their interaction. *Proc Natl Acad Sci USA*, 2007;**104**(13):5378–82.
51. Gennerich, A., and D. Schild, Sizing-up finite fluorescent particles with nanometer-scale precision by convolution and correlation image analysis. *Eur Biophys J*, 2005;**34**(3):181–99.
52. Nan, X., et al., Observation of individual microtubule motor steps in living cells with endocytosed quantum dots. *J Phys Chem B*, 2005;**109**(51):24220–4.
53. Watanabe, T. M., and H. Higuchi, Stepwise movements in vesicle transport of HER2 by motor proteins in living cells. *Biophys J*, 2007;**92**(11):4109–20.
54. Nan, X., P. A. Sims, and X. S. Xie, Organelle tracking in a living cell with microsecond time resolution and nanometer spatial precision. Chemphyschem, 2008;**9**(5):707–12.
55. Enderlein, J., E. Toprak, and P. R. Selvin, Polarization effect on position accuracy of fluorophore localization. *Opt Express*, 2006;**14**(18):8111–20.
56. Yasuda, R., et al., Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. *Nature*, 2001;**410**(6831):898–904.
57. Xie, X. S., and R. C. Dunn, Probing single molecule dynamics. *Science*, 1994;**265**(5170):361–4.
58. Ha, T., et al., Single molecule dynamics studied by polarization modulation. *Phys Rev Lett*, 1996;**77**: 3979–82.
59. Sase, I., et al., Axial rotation of sliding actin filaments revealed by single-fluorophore imaging. *Proc Natl Acad Sci USA*, 1997;**94**(11):5646–50.

60. Sosa, H., et al., ADP-induced rocking of the kinesin motor domain revealed by single-molecule fluorescence polarization microscopy. *Nat Struct Biol*, 2001;**8**(6):540–4.
61. Asenjo, A. B., N. Krohn, and H. Sosa, Configuration of the two kinesin motor domains during ATP hydrolysis. *Nat Struct Biol*, 2003;**10**(10):836–42.
62. Asenjo, A. B., Y. Weinberg, and H. Sosa, Nucleotide binding and hydrolysis induces a disorder–order transition in the kinesin neck-linker region. *Nat Struct Mol Biol*, 2006;**13**(7):648–54.
63. Forkey, J. N., et al., Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization. *Nature*, 2003;**422**(6930):399–404.
64. Syed, S., et al., Adaptability of myosin V studied by simultaneous detection of position and orientation. *EMBO J*, 2006;**25**(9):1795–803.
65. Bartko, A. P., K. Xu, and R. M. Dickson, Three-dimensional single molecule rotational diffusion in glassy state polymer films. *Phys Rev Lett*, 2002;**89**(2):026101/1–4.
66. Patra, D., I. Gregor, and J. Enderlein, Image analysis of defocused single-molecule images for three-dimensional molecule orientation studies. *J Phys Chem A*, 2004;**108**(33):6836–41.
67. Toprak, E., et al., Defocused orientation and position imaging (DOPI) of myosin V. *Proc Natl Acad Sci USA*, 2006;**103**(17):6495–9.
68. Shaner, N. C., et al., Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol*, 2004;**22**(12):1567–72.
69. Cinelli, R. A., et al., The enhanced green fluorescent protein as a tool for the analysis of protein dynamics and localization: local fluorescence study at the single-molecule level. *Photochem Photobiol*, 2000;**71**(6):771–6.
70. Howarth, M., et al., Monovalent, reduced-size quantum dots for imaging receptors on living cells. *Nat Methods*, 2008;**5**(5):397–9.
71. Smith, A. M., and S. Nie, Minimizing the hydrodynamic size of quantum dots with multifunctional multidentate polymer ligands. *J Am Chem Soc*, 2008;**130**(34):11278–9.
72. Dunn, A. R., and J. A. Spudich, Dynamics of the unbound head during myosin V processive translocation. *Nat Struct Mol Biol*, 2007;**14**(3):246–8.
73. Egner, A., S. Jakobs, and S. W. Hell, Fast 100-nm resolution three-dimensional microscope reveals structural plasticity of mitochondria in live yeast. *Proc Natl Acad Sci USA*, 2002;**99**(6):3370–5.
74. Klar, T. A., et al., Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc Natl Acad Sci USA*, 2000;**97**(15):8206–10.
75. Schermelleh, L., et al., Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science*, 2008;**320**(5881):1332–6.
76. Rust, M. J., M. Bates, and X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods*, 2006;**3**(10):793–5.
77. Betzig, E., et al., Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 2006;**313**(5793):1642–5.
78. Rosenberg, S. A., et al., Rotational motions of macro-molecules by single-molecule fluorescence microscopy. *Acc Chem Res*, 2005;**38**(7):583–93.

2

Single-Molecule Analysis of Biomembranes

Thomas Schmidt and Gerhard J. Schütz

Abstract Biomembranes are more than just a cell's envelope – as the interface to the surrounding of a cell they carry key signalling functions. Consequentially, membranes are highly complex organelles: they host about thousand different types of lipids and about half of the proteome, whose interaction has to be orchestrated appropriately for the various signalling purposes. In particular, knowledge on the nanoscopic organization of the plasma membrane appears critical for understanding the regulation of interactions between membrane proteins. The high localization precision of ∼20 nm combined with a high time resolution of ∼1 ms made single molecule tracking an excellent technology to obtain insights into membrane nanostructures, even in a live cell context. In this chapter, we will highlight concepts to achieve superresolution by single molecule imaging, summarize tools for data analysis, and review applications on artificial and live cell membranes.

2.1. Introduction

From the beginning of cell biology, biomembranes have been considered to be of major relevance for cellular function. Each cell is separated from its environment via a biomembrane. Membranes enable the division of the cell cytoplasm into chemically distinct subspaces, thereby governing the development of cell organelles. Moreover, they act as a matrix for integral, lipid-anchored, or peripheral membrane proteins and thereby considerably affect protein function.

The role of lipids as the major constituents of biomembranes was established in the 1970s, leading to the proposal of the fluid mosaic model for the plasma membrane [1]. According to this model, it is the lipids and not the proteins that constitute the matrix of the cell membrane. Since this matrix was shown to be in a fluid state, integral membrane proteins were expected to be free to diffuse laterally within the membrane, as long as no interaction with cytoskeletal elements hinders their Brownian motion. The relevance of lipid bilayers for cell biology initiated a vast number of studies on artificial systems, which aimed at the structural characterization of bilayers and monolayers under various environmental conditions [2].

T. Schmidt • Physics of Life Processes, Leiden Institute of Physics, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
G. J. Schütz • Biophysics Institute, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_2,

The restriction to the study of artificial systems allowed for full control over parameters such as lipid composition, lateral pressure, temperature, and ionic strength.

In particular, there has been increasing interest in obtaining detailed understanding of the structure and dynamics of the cellular plasma membrane itself [3], primarily based on recognition of its essential role in controlling cellular signaling processes. In recent years a picture has emerged that ascribes to the plasma membrane a high degree of organization at very short length scales of tens of nanometers [4–7]. Experiments performed on single biomolecules have added to this picture by providing access to spatial information below the diffraction limit of classic light microscopy.

A biomolecule moving in an artificial or cellular lipid membrane may experience multiple forces that influence its characteristic motion. In consequence, the path will deviate to a greater or lesser extent from a Brownian trajectory. Deviations from free diffusion may be caused by (1) the confinement of the tracer molecule in a meshwork of permeable barriers, resulting in hop diffusion [8]; (2) the transient binding of the tracer to an (immobile) membrane structure, yielding short periods of altered—generally reduced —mobility [9,10]; (3) periods of active transport via motor proteins [11]; (4) direct or indirect anchorage to the cytoskeleton, yielding immobilization or tethered motion of the tracer [12]; (5) fixed obstacles at high surface density [13]; and (6) the partitioning of the tracer to mobile or immobile membrane domains ("rafts") [14,15].

In the following we distinguish two strategies for following a single biomembrane constituent. Historically the first approach was termed *single-particle tracking (SPT)*, in that the molecule of interest was linked specifically to a larger particle that gave sufficient signal to be detectable as an individual point light source over multiple frames [16–25]. Single particles can be imaged with high signal-to-noise ratio, and therefore the location can be determined with high precision far beyond the limit of diffraction [23,26]. There is hardly any limit to the length of such a trajectory because no photophysical damage occurs to the traced particles. Termination may yet occur when using quantum dots for labeling. Due to inherent blinking [27], there will be dark periods in a trajectory; if the dark interval gets too long, correlation with the next appearance may be difficult.

The second approach was based on the attempt to further reduce the size of the label down to the ultimate limit of an *individual dye molecule*. These endeavors were based on the concern that larger particles might interfere with the motion and function of the tracer molecule. Pioneering studies in the mid 1990s demonstrated that single dye molecules indeed provide sufficient signal for imaging with subpixel localization precision and tracking over multiple frames [28–31]. Further technological advances, as well as improved understanding of the biological systems, have led to an increase in studies both on artificial membranes [32–39] and the live cell plasma membrane [40–61]. In contrast to particles, however, dye molecules are affected by prolonged excitation, which results in the photobleaching of the probe and the concomitant termination of the trajectory [62].

We focus in this chapter on reviewing analytical tools for the analysis of single-molecule tracking experiments in biomembranes and describe applications to artificial membranes and the live cell plasma membrane; the emphasis is on single-dye approaches. Initially, we discuss concepts of superresolution imaging based on single-molecule detection.

2.2. Superresolution

The major advantage of single-molecule approaches for biomembrane research is the enormous potential for characterization of subwavelength membrane structures. To

elucidate superresolution concepts for membrane research, we first briefly sketch the imaging constraints. An isotropic emitter (fluorescent bead or quantum dot) or scatterer (gold particle) will be imaged according to the point spread function of the imaging device, which is well approximated by an Airy disk of radius $\rho = 0.61 \times \lambda/NA$, with λ the imaging wavelength and NA the numerical aperture [63]. A single dye molecule with fixed emission dipole orientation will emit photons depending on its three-dimensional (3D) orientation in space, which in general leads to a significantly altered image [64]. In the case of a rapidly rotating molecule the contributions of the individual orientations add up, thereby yielding an image given by the point spread function. Since in fluid biomembranes the constituents show rapid rotation on a nanoseconds time scale [35], an approximation of the image by the point spread function centered on the actual position is justified.

The nonzero width of the Airy disk represents the basic limitation to the resolving power of the apparatus used. Imaging of an arbitrary object can be described by

$$f'(\vec{r}) = \int_{-\infty}^{-\infty} f(\vec{s})\,\delta(\vec{r}-\vec{s})\,d\vec{s} \Rightarrow \int_{-\infty}^{-\infty} f(\vec{s})\,psf(\vec{r}-\vec{s})\,d\vec{s} = g(\vec{r})$$

with $f'(\vec{r})$ representing the fluorescent object, which is described as a sum of point light sources, each of which is imaged according to the point spread function (PSF), yielding the total image $g(\vec{r})$; the arrow denotes the imaging processes. If we assume a Gaussian approximation for the PSF, a sample labeled with dyes at positions $\vec{s}_i$ would yield the image

$$g(\vec{r}) = \sum_i N_i \frac{1}{2\pi\rho^2} \exp\left(-\frac{(\vec{r}-\vec{s}_i)^2}{2\rho^2}\right)$$

where $N_{i\,\mathrm{is}}$ the number of photons emitted per dye. As a consequence, two point light sources with a distance smaller than ρ would yield highly overlapping images and thus could not be resolved as individual objects. This theorem, known as the "Abbe limit" or "Rayleigh criterion," is generally used to define resolution in light microscopy. Several modifications to standard imaging methods have been employed to obtain higher resolution via narrowing the point spread function by utilizing near-field effects [65,66], nonlinear excitation [67], by confocal excitation and detection schemes [68,69]. Recently, new concepts were introduced that use saturation effects when illuminating the sample with structured illumination [68,70,71] (see Ref. 72 for review).

Single–molecule imaging has enabled additional strategies for achieving superresolution. The general idea is to thin out the active probe such that signals become well separated and can be detected by classical diffraction-limited optics [73]. Then, a molecule can be localized to high precision of a few nanometers by determining the centroid of its image [26,64,74–76]. Different variants have been introduced to use this high localization precision for determination of molecular distances and for enhancement of imaging resolution in general (see Figure 2.1 for a sketch):

1. *Two-color imaging*. In the first approaches, different color channels were used to measure the distance between unresolvably close molecules of different spectrum [33,77,78]. Two-color imaging offers a way to sensitively detect colocalization between a ligand and its receptor in binding assays [79,80] and also in the live cell plasma membrane [10].
2. *Dynamic imaging*. Alternatively, one may exploit the time domain for diluting the biomolecule, for example, by allowing fluorescent ligands to bind and dissociate

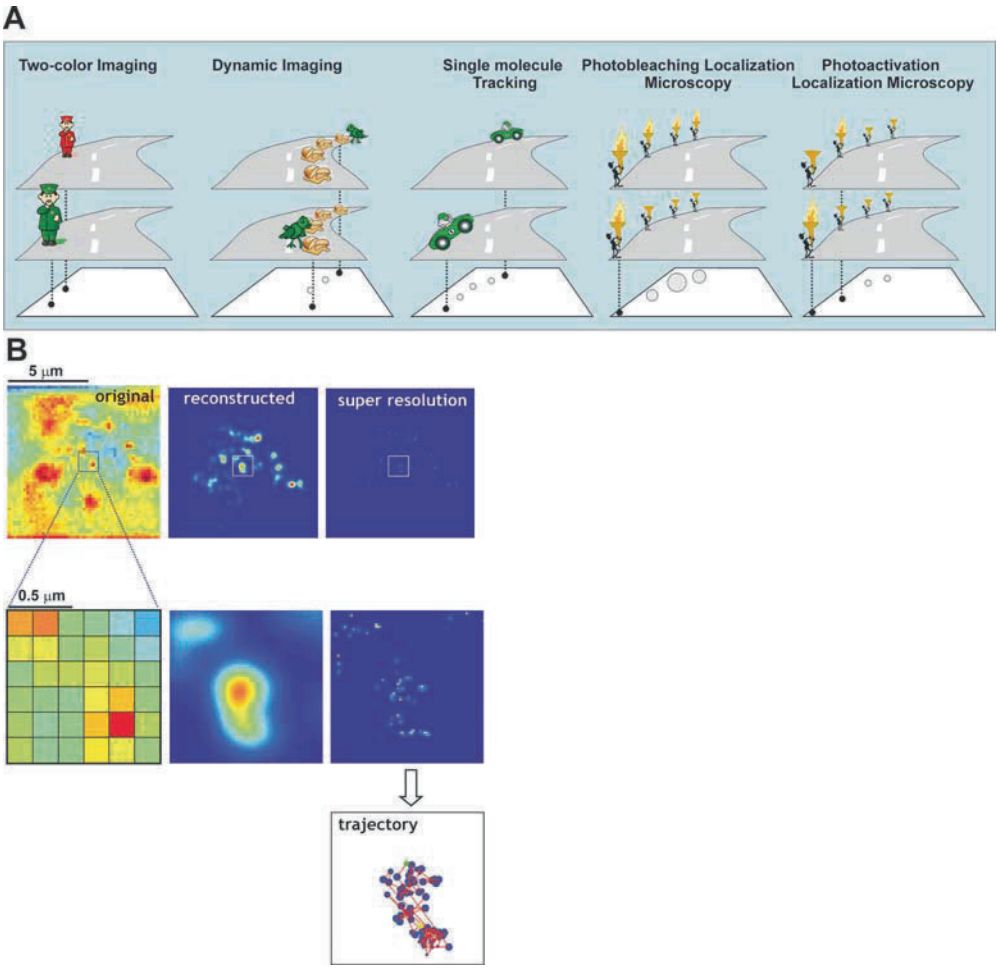


Figure 2.1. A. Concepts for achieving superresolution based on single-molecule imaging. The structure to be studied is represented as a road carrying different objects. The two top rows show two arbitrary observations of the system. The bottom row shows the results: The circles denote the detected single-molecule positions, and the size of the circles indicates the localization precision. See text for details on the individual approaches. **B.** The mean background-subtracted image taken from a stack of 500 images in an experiment described in Figure 2.2 exemplifies the construction of a superresolution image. Whereas the original mean image shows hardly any structure, the reconstructed image, as noise-free, indicates the existence of still-unresolved finer structure within the stack. The superresolution image resolves this structure given by the path of a moving molecule. The structure is better seen at 10× magnification (*lower image row*). In the right column, the single-molecule position data are plotted as Gaussians of widths given by the respective confidence limit of the positional fit. Such superresolution data can be subsequently used to construct molecular trajectories.

randomly from their receptors in the plasma membrane. If the applied concentration is low enough, single binding events can be detected and localized. An overlay of all determined positions provides a high-resolution map of the receptor organization in the plasma membrane [81].

3. *Single-molecule/particle tracking*. A similar strategy is based on the study of mobile molecules, which sample the accessible space within the biomembrane over time. Whenever structural features affect any measurable parameter of the molecule (e.g., its intensity, spectrum, the diffusion constant, or velocity), each trajectory can be

subdivided into distinct segments with a resolution only limited by the localization precision. This strategy has been used to identify, for example, step sizes of motor proteins [75,82] and confinements in artificial membranes [32] or in the cellular plasma membrane [40,46,56,57].

4. *Photobleaching localization microscopy.* To facilitate accurate determination of the positions of neighboring molecules with nanometer distance, Gordon and coworkers used photobleaching [83]: whenever a molecule within an assembly of subdiffraction dimension photobleaches, a concomitant shift in the signal distribution can be detected and used iteratively for the estimation of all dye positions. A similar approach has been used to determine the label positions on stretched DNA [84].
5. *Photoactivation or photoswitching localization microscopy*. Photobleaching localization microscopy is restricted to the analysis of weakly stained specimens because the localization analysis becomes inaccurate for a large number of dyes within a diffraction-limited spot. As an alternative, researchers have developed the inverse approach, which is based on the consecutive photoactivation, imaging, and photobleaching of a sample labeled with photoactivatable fluorophores (e.g., photoactivatable green fluorescent protein (paGFP) [85]) [86–88]. In this methodology, localization precision determines the resolution [89].

The listed approaches may all be applied for high-resolution biomembrane imaging; see the following sections for a detailed discussion of the individual examples. It should be noted, however, that all superresolution concepts—including the ensemble techniques described earlier —are comparably slow, in that they require addressing the pixels individually. More precisely, if two pixels are to be addressed simultaneously, they have to be separated at least by the width of the diffraction-limited point spread function. For two-dimensional (2D) microscopy, the minimum time required for recording an image at superresolution $T_{\min}$ is given by $T_{\min} = (\rho/\sigma_{xy})^2 t_{\text{frame}}$, with ρ the width of the point spread function, σ_{xy} the localization precision, and t_{frame} the time required for obtaining one frame; for example, recording an image at 10-times-improved resolution would require 100-fold longer recording times! If we further consider the imaging of dynamic structures moving with a diffusion coefficient D, the recording time should be faster than $T_{\max} \approx \text{resolution}^2/4D$, the time the molecule needs to move over one pixel. Superresolution imaging of dynamical processes would therefore actually require improved time resolution.

In contrast, tracking approaches do not necessarily lead to images; one can also use the information content of single trajectories to analyze structures that interfere with the motion of the tracer. In this case, the timing requirements are less strict: It may not even be necessary to image faster than the time the molecule needs to traverse the structure of interest. This advantage made single-particle/molecule tracking a preferred methodology in biomembrane research.

2.3. Detection and Tracking

The possibility of detecting individual fluorophores emerged in 1989 [90] with the detection of individual pentacene molecules at liquid-He temperatures in a confocal setting. Since then, the developments in optical dielectric filters and charge-coupled device (CCD) cameras has led to sensitivities that makes it possible to detect individual fluorophores in wide-field microscopy and at video rates in the living cell. The selective throughput of optical filters now approaches 90% at the same time blocking unwanted wavelengths by $>10^5$, and

the detection efficiency of back-illuminated CCD cameras is up to 95% throughout the whole visible spectral range. The reliable detection of molecules is hence just limited by the noise in acquisition governed by detector noise (which can be as low as 3 electrons per pixel root mean square (RMS) on a N_2-cooled CCD) and the sample background. Optimization in preparation protocols reduces the latter to basically zero for artificial systems (e.g., polymer films, biomimetic systems) and for cells observed in total internal reflection (TIR) mode, whereas for cells observed in regular wide-field applications the background is still substantial. Typically 30 counts per pixel per millisecond illumination at excitation intensity $\sim$1 kW/cm^2 is seen for live-cell wide-field applications. At those conditions about 100 counts are detected from an individual fluorophore [91]. A typical image is shown in Figure 2.2 ("original").

The sheer amount of data generated by such images requires a stable and unattended automatic analysis. In a first step any static or slowly varying background has to be subtracted. Techniques like high-order [4–8] polynomial fitting, low-spatial-frequency filtering, mean image calculation, and pixel-by-pixel low-temporal frequency filtering have been used to generate a reliable background image (Figure 2.2, "background"). The background-subtracted images are subsequently optimally filtered by cross-correlation with the PSF of the microscope approximated by a Gaussian (Figure 2.2, "target"). Thresholding of those optimally filtered images in which the threshold is dynamically adjusted to the image noise makes it possible reliably to identify signals that are above noise by a factor of >3 (Figure 2.2, "thresholded"). The final analysis then proceeds by fitting each of the identified signals to the PSF and extracting the information on position, signal strength, spatial width of the signal, and signal of the background, as well as an estimate on the accuracy in each of the parameters (Figure 2.2, "fitting"). Depending on the type of data, also 3D information can be retrieved. Within the Rayleigh length [$z_0 = \pi\lambda/(4\mathrm{NA}^2)$] the PSF widens with z position as $\omega(z) = \omega_0\sqrt{1 + z^2/z_0^2}$, which permits extraction of the z position [41]. For small deviations from the plane of focus (small z) it is advantageous to introduce a small astigmatism into the detection light path to increase the accuracy in determination of the z position [92]. The positional accuracy is further increased by taking fast image stacks, which are subsequently analyzed simultaneously. In this way the 3D position of an individual object is determined to within 30 nm in xyz, respectively, and image volumes of $10\times10\times10$ μm^3 are analyzed at frame rates of 10 sec^{-1} (Laurent Holtzer, unpublished results; [93]).

Image analysis as just described is the basis for further data treatment. It becomes possible to restrict further analysis to signals of, for example, a certain signal strength or a desired positional accuracy (Figure 2.2, "reconstructed"). In particular the latter is key for the novel nanometer-resolution microscopy methods described in the previous subsection. Correlations between the parameters can be used to obtain information on, for example, the aggregation state of the molecule or local differences in pH seen as a local change in signal strength.

In addition, temporal information from image movies together with the positional information is used to obtain trajectories of individual molecules. The generation of trajectories is straightforward for sparsely occurring and immobile objects [25,94]. However, in general we face the problem of correlating M mobile molecules in image n with L mobile molecules in image $n + 1$ (Figure 2.3). This is a classical NP-complete problem ("traveling salesmen") for which we follow an optimization procedure from operations research developed by Vogel. First a complete translational matrix is built up that includes the probability that molecule j in image i at position $r_{j,i}$ moves to molecule k in image $I + 1$ at position $r_{k,i+1}$ by diffusion as characterized by a diffusion constant D,

$$p(j,k) = \exp\left(-\frac{(\vec{r}_{j,i} - \vec{r}_{k,i+1})^2}{4Dt}\right)$$

In addition, molecules are allowed to disappear by diffusion out of the observed area or by photobleaching, $p(j,k > L) = p_{\text{bleach}}$, and molecules are allowed to move into the observed area or get reactivated, $p(j > M,k) = p_{\text{activation}}$. Subsequently trajectories are constructed that optimize for the total probability of all connections between two images,

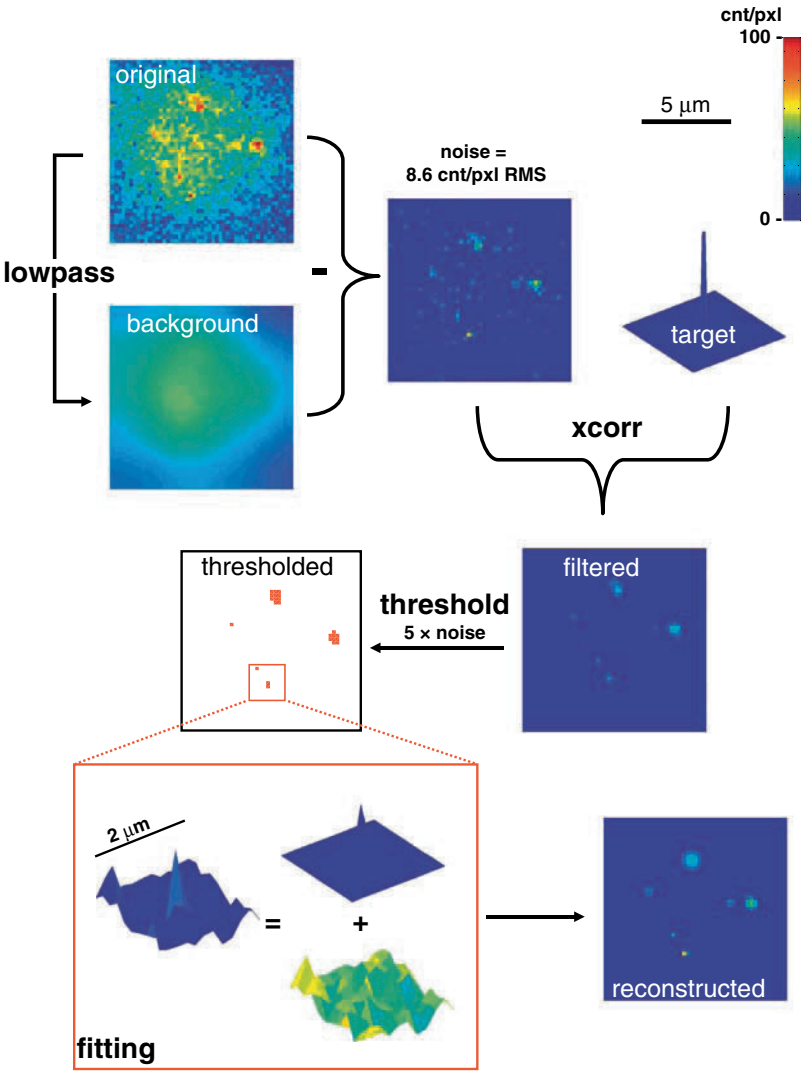


Figure 2.2. Steps in image analysis for automated single-molecule detection. *Dictyostelium* cells were transfected with a yellow fluorescent protein (YFP)–labeled cyclic adenosine monophosphate (cAMP) receptor located in the plasma membrane of the cells. The cells were illuminated at an intensity of 2 kW/cm^2 for 5 ms to create the images on a N_2-cooled charge-coupled device (CCD) camera. The images have a size of 10×10 µm^2. Low-pass filtering of the original image was used to create a smooth background image. The background-subtracted original image (noise: 8.6 counts/pixel root mean square) was subsequently cross-correlated (xcorr) by the Gaussian target image, leading to an optimally filtered image. Signals that exceed a threshold criterion are subsequently fitted to the Gaussian target image. Results of the fitting were used to obtain a noise-free reconstruction image.

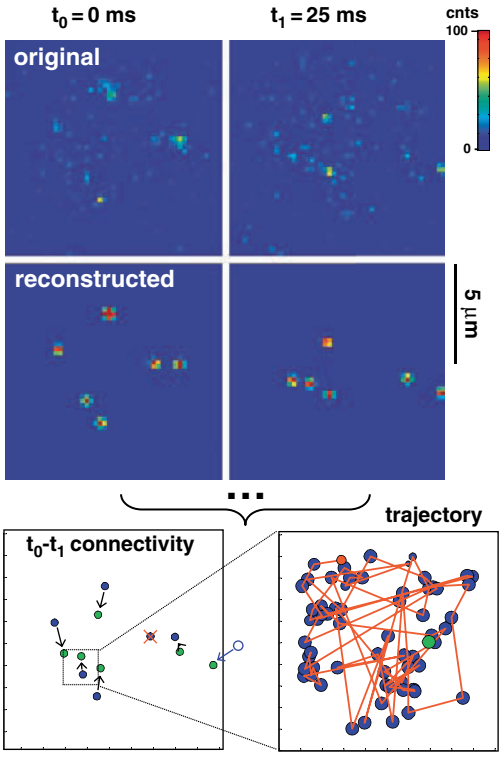


Figure 2.3. Two subsequent images (25-ms delay) from the experiment described in Figure 2.2 exemplify the construction of trajectories. A Vogel algorithm was used to obtain the connectivity map in which the various processes are visualized: (1) diffusion (*black arrows*), (2) bleaching (*red cross*), and (3) diffusion into the observation volume (*blue arrow*). The position of the molecules at t_0 are shown in blue, those at t_1 are shown in green. On the bottom right the trajectory of a single receptor is shown.

$\log(P) = \sum_{j,k} \log(p(j,k))$. This algorithm enhances the number of faithfully reconstructed trajectories even in the case of a sizable amount of molecules per image.

Since one step in trajectory reconstruction contains the assumption of diffusional motion, it is obvious that trajectories can be used to learn about molecular mobility on the nanometer scale. A detailed description will be given in the next subsection. It should be mentioned, however, that mobility information can be directly extracted from the position data. Here the position data are used to construct spatial image correlations that directly yield the desired molecular mobilities at nanometer accuracy (particle image correlation spectroscopy [PICS]; [95]). The advantage of correlation analysis is that it is very robust, and the accuracy of the mobility parameter extracted is clearly defined by statistics. The disadvantage clearly is that it is impossible to trace back to individual events or individual molecules as sometimes desired, as described later.

2.4. Learning from Trajectories

Although there is currently an intense discussion on the movements of membrane lipids over molecular distances [96–98], Brownian motion is considered as the appropriate model

when studying movements over length scales of tens of nanometers. We summarize here a few essentials of diffusion models; a more detailed overview can be found in Ref. 99.

A Brownian walker starting at time $t = 0$ at position $\vec{x}_0$ will be found at a later time t within the interval $\left[\vec{x},\vec{x} + d\vec{x}\right]$ with a probability

$$p\left(\vec{x},t\right) dx = \frac{1}{\sqrt{\left(4\pi Dt\right)^d}} \exp\left(-\frac{\left(\vec{x} - \vec{x}_0\right)^2}{4Dt}\right) d\vec{x}$$

where D denotes the diffusion coefficient and d the number of dimensions. Because diffusion has no preference for a particular direction, the average position of the tracer at any instant of time will be its starting position, $\langle \vec{x}\left(t\right)\rangle = \int \vec{x} p\left(\vec{x},t\right) d\vec{x} = \vec{x}_0$. Concurrently, it will be localized at a characteristic distance away from the starting point $\left\langle x^2\left(t\right)\right\rangle = \int \left(\vec{x} - \vec{x}_0\right)^2 p\left(\vec{x},t\right) d\vec{x} = 2dDt$. The distance $\left\langle x^2\left(t\right)\right\rangle$ is frequently termed mean square displacement (msd); it increases linearly with time, with the slope specifying the diffusion coefficient. It is further worthwhile to inspect the distribution $p\left(\vec{x},t\right) d\vec{x}$ more closely. For convenience, let us set $\vec{x}_0 = \vec{0}$. In the case of two-dimensional diffusion, the cumulative density function

$$\mathrm{cdf}(r^2,t) = \int_0^{r^2} p(\rho^2,t) d\rho^2$$

is given by

$$\mathrm{cdf}(r^2,t) = 1 - \exp\left(-\frac{r^2}{4Dt}\right)$$

with r denoting the 2D distance; note that this simple monoexponential expression is only valid for 2D diffusion.

Experimental data, however, contain localization errors, which have to be accounted for in the analysis. Localization errors can be well characterized using immobilized particles, and typically follow a Gaussian distribution

$$p_{err}\left(\vec{x}_{err}\right) d\vec{x}_{err} = \frac{1}{\sqrt{2\pi\sigma_{xy}^2}} \exp\left(-\frac{x_{err}^2}{2\sigma_{xy}^2}\right) d\vec{x}_{err}$$

Since the measured displacement is given by $\vec{x}_{\exp} = \vec{x} + 2\vec{x}_{err}$, the experimentally accessible probability distribution is a convolution of $p\left(\vec{x},t\right)$ and $p_{err}\left(\vec{x}_{err}\right)$, yielding

$$p\left(\vec{x},t\right) dx = \frac{1}{\sqrt{\left(4\pi Dt + 4\sigma_{xy}^2\right)^d}} \exp\left(-\frac{x^2}{4Dt + 4\sigma_{xy}^2}\right) d\vec{x}$$

For 2D diffusion, the cumulative density function changes to

$$\mathrm{cdf}(r^2,t) = 1 - \exp\left(-\frac{r^2}{4Dt + 4\sigma_{xy}^2}\right)$$

and the mean square displacement to $\left\langle x^2\left(t\right)\right\rangle = 4Dt + 4\sigma_{xy}^2$.

If the traced molecules follow free Brownian motion, tracking data may be used to study the diffusion coefficient under various conditions. Already 100 years ago, Einstein derived the famous fluctuation-dissipation theorem $D = k_{\mathrm{B}}T/\gamma$, with k_BT the Boltzmann energy and γ the Stokes friction coefficient; for a sphere of radius R moving in a fluid of viscosity η,

the mobility is characterized by $D = k_B T/6\pi\eta R$. Diffusion in a membrane, however, turned out to be more difficult to address. The seminal work by Saffman and Delbruck in 1975 found a logarithmic dependence of the diffusion coefficient on the radius of a cylindrical membrane protein [100]. Hughes et al. provided an extension for larger drag coefficients and concomitantly larger objects [101,102]; an approximation of the Hughes formula can be found in Ref. 103. In contrast, recent experimental data show strong deviations from the hydrodynamic models and indicate a Stokes-Einstein-like $1/R$-dependence of the diffusion coefficient [104].

Most researchers, however, have attempted to use single-molecule trajectories for identifying deviations from free Brownian motion. If the length of the observed trajectories exceeds hundreds of observations—for example, in single-particle tracking experiments—a detailed investigation at the level of a single trajectory is possible [105,106]. However, the lengths of single-molecule trajectories are limited by photobleaching to a couple of tens of observations [62], rendering such detailed analysis impracticable. We restrict our discussion here to approaches that average over multiple single-molecule trajectories. A more detailed discussion of diffusion analysis can be found in Ref. 107.

For example, multiple mobile fractions can be discriminated by analyzing the step-size distribution [32]. A mixture of two different fractions α and $(1-\alpha)$ with mobility D_1 and D_2, respectively, can be identified as different exponential components in the cumulative density function

$$\mathrm{cdf} = 1 - \alpha \ \exp\left(-\frac{r^2}{4D_1 t}\right) - (1-\alpha) \ \exp\left(-\frac{r^2}{4D_2 t}\right)$$

The strategy has been successfully applied to identify subfractions of distinct mobility in model systems [32,38] and living cells [47].

Frequently, the diffusion of membrane proteins or lipids has been found to deviate from free Brownian motion and to be better described by anomalous subdiffusion, where the msd dependence on the time is characterized by a sublinear increase, $\mathrm{msd} \propto t^{\alpha}$, with $\alpha < 1$ the anomalous diffusion exponent [49,108–110]; in other words, the apparent diffusion coefficient appears to decrease with increasing time according to $D_{\mathrm{app}} = \mathrm{msd}/(4t) \propto t^{\alpha-1}$. Anomalous subdiffusion leads to an improved screening of the local environment and may therefore represent an important mechanism for enhancing the binding probability of the biomolecule to a nearby target [111]. Multiple studies have addressed potential sources of anomalous diffusion on cell membranes via Monte Carlo simulations [13,112–116]. In particular, fixed obstacles at concentrations close to the percolation threshold lead to a pronounced anomalous subdiffusion [13]. However, anomalous subdiffusion may also be misinferred by disregarding or incorrectly accounting for the finite localization precision [117]. In the following, we discuss the dependence of the mean-square displacement on time for different matrix geometries.

The confined diffusion of plasma membrane proteins or lipids within domains [8,18,40,42, 46,53,118] can be regarded as a specific case of subdiffusion. Analytical treatments have been provided for certain shapes of the confinement zones, thereby enabling quantitative estimates of the confinement size and the characteristic mobilities. For example, the diffusion of a tracer with mobility D_{micro} inside an impermeable circular or square corral has been described analytically [18,116]. For short time-lags, the function can be approximated by free diffusion, $\mathrm{msd}(t \to 0) \approx 4D_{\mathrm{micro}}t$; in the limit of long time-lags a constant offset is reached at the radius of gyration, which is $L^2/3$ for a square of length L and is R^2 for a circle of radius R [116].

We recently generalized the analytical treatment to the case of a meshwork of periodic permeable corrals, so that the tracer can escape into the adjacent compartment, performing a type of motion referred to as hop diffusion [42]. In this case, a transition between rapid microscopic diffusion within the domains and slow macroscopic diffusion from domain to domain can be observed. Figure 2.4 shows the time dependence of both msd (Figure 2.4A) and D_{app} (Figure 2.4B), the latter in a double-logarithmic plot. In that representation,

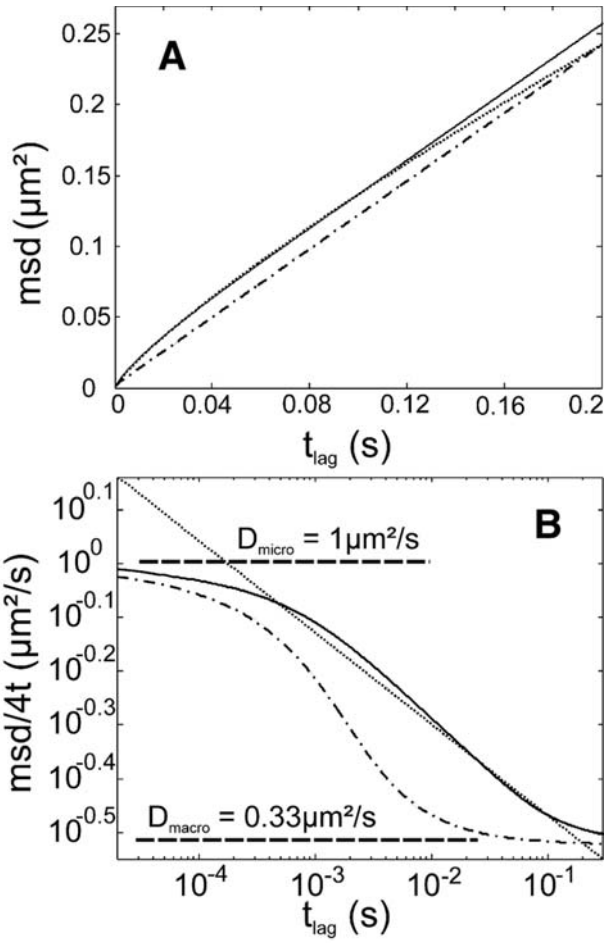


Figure 2.4. Effect of different geometries on the time dependence of the mean square displacement (msd) [107]. Hop diffusion in a meshwork of squares yields an apparent anomalous subdiffusion, in particular when the measurement is performed within a small time window. Hop diffusion was simulated with a mesh-size of $L = 0.1$ μm, a microscopic mobility of 1 μm^2/sec, and $\hat{\tau} = 3.33$ (*dot-dashed line*). **A.** The msd as a function of t_{lag}. **B.** The apparent diffusion constant $D_{\text{app}} = \text{msd}/(4t_{\text{lag}})$ versus t_{lag} on a log scale. In panel B the transition between the microscopic mobility on the left and the macroscopic mobility on the right is clearly visible. The transition region yields a rather straight line within a time window of about one order of magnitude. The effect smears out in the case of a distribution of confinement sizes, as shown in a simulation of equal contributions from $L = 0.1$, 0.2, and 0.5 μm (*solid line*). The curve can be well approximated by anomalous subdiffusion with msd $\propto t_{\text{lag}}^{0.83}$ over a time window of two orders of magnitude (*dotted line*). (Reprinted from Wieser and Schütz [107].)

anomalous subdiffusion would yield a linear decrease of $\log(D_{\text{app}})$ with $\log(t)$. Hop diffusion shows a different behavior, with the two mobility regimes being clearly discriminated as two plateaus for D_{micro} and D_{macro}, separated by a rather linear transition region. However, to identify both plateaus a time window of about four orders of magnitude is required, which is typically not available from experimental data. Within experimentally reasonable time windows of about one order of magnitude, the curves can be also well approximated by standard anomalous subdiffusion, making it difficult to distinguish those models. The linear transition region gets even larger if a natural variation in the domain sizes occurs. Figure 2.4 shows a scenario assuming equal contributions from domains with size 100, 200, and 500 nm. In this case, a time window of two orders of magnitude can be well fitted by msd $\propto t^{\alpha}$; researchers would not doubt to ascribe such results to anomalous subdiffusion.

Many more 2D diffusion models can be constructed that yield mutually similar or identical results. For example, a well-known model for organelle transport assumes the diffusion of the tracer in an impermeable mobile domain [119]. This model may well be applicable for describing protein mobility in cell membranes, in particular in view of recent studies indicating the existence of protein islands in the cellular plasma membrane [120]. However, when analyzing msd versus t, a moving-domain model yields the same result as a hop diffusion model, rendering the two cases indistinguishable. It should be emphasized that such ambiguities require cautious application of diffusion analysis based on msd, in particular when different diffusion models are to be distinguished.

Measurements of the lateral diffusion typically assume that the biomolecule moves in a flat plane. Yet, undulations or fluctuations of the membrane on a length scale close to the localization precision absolutely can be expected. In this case, the projection of the movement onto the focal plane will be observed. Ruffled surfaces or thermal membrane fluctuations will essentially reduce the diffusion constant [121–123]; more dramatic effects may be expected when the biomolecule enters highly curved 3D structures. We recently calculated and measured the one-dimensional mobility of membrane proteins diffusing along the circumference of tunneling nanotubules (TNTs) [44], cylindrical structures with a radius $R \approx 130$ nm that interconnect different cells for membrane and organelle exchange [124]. Since the molecules are tied to the cylindrical surface of the TNT, the msd for transverse motion saturates at R^2 for long time-lags. For short time-lags, the mobility can be approximated by msd $\approx D_{\perp} t_{\text{lag}}$, concomitant with a reduction in the apparent mobility by a factor of two.

Finally, the motion of the tracer during its illumination will change the recorded displacements to reduced values [107]. This effect is present for a freely diffusing molecule [125] and becomes dramatic for confined molecules [42,110,126].

2.5. Application 1—Synthetic Lipid Bilayers

The application of single-molecule microscopy to synthetic bilayers [29] was key for the development of single-molecule technology for biosciences and cell biology [127]. Similarly, synthetic bilayers are an excellent arena in which to understand the underlying properties of cellular membranes and as a substrate for bio-nanoscience applications: For both, the microscopic structure of mixed bilayers systems has been of prime scientific interest. When lipids of different phase transition temperature are mixed and the temperature lowered, demixing and macroscopic phase separation occurs [128]. Domain size and domain dynamics of such systems can be addressed by single-molecule methods. Labeling of a trace amount of lipids with a fluorescence marker allowed us directly to visualize the diffusional motion of lipids

in a supported membrane [32], in a free-standing membrane spanning a micron-sized hole [34], and in the membrane of giant unilamellar vesicles (GUVs) (S. Semrau and T. Schmidt, unpublished, 2007). The significant increase in mobility from $D = 4\ \mu m^2/sec$ in supported membranes to $D = 10–20\ \mu m^2/sec$ in free-standing membranes and GUVs clearly shows the influence of the membrane–substrate interaction on lipid mobility (Figure 2.6A). Depending on the preparation technique, this interaction can be quite inhomogeneous, leading to domains or diffusion-barrier formation on the hundred-nanometer length scale as resolved by single-molecule tracking experiments [32]. Whereas those domains, due to imperfections of bilayers formation even in unimolecular systems, are undisputed, the appearance of nanometric domains in more complicated mixed systems like the universal phosphocholine-sphingolipid-cholesterol 1:1:1 "raft" mixture close to the demixing point is still under investigation (Figure 2.5). This is because the lateral mobility of lipids in the liquid-ordered and liquid-disordered phases, respectively, is very similar and hence difficult to distinguish. Multiparameter single-molecule microscopy, however, might be able to solve the experimental problem. Dual-channel imaging with the fluorescence anisotropy as readout has been developed to address fast rotational motion [35] and fluorescence resonant energy transfer [129] in bilayers systems. As the rotational motion of lipids in the liquid-ordered state is significantly lower than in the liquid-disordered state, trajectory analysis as described in the last subsection combined with simultaneous analysis of the rotational mobility might give an indication of whether nanoscopic phase separation does occur. At least the macroscopic phase separation

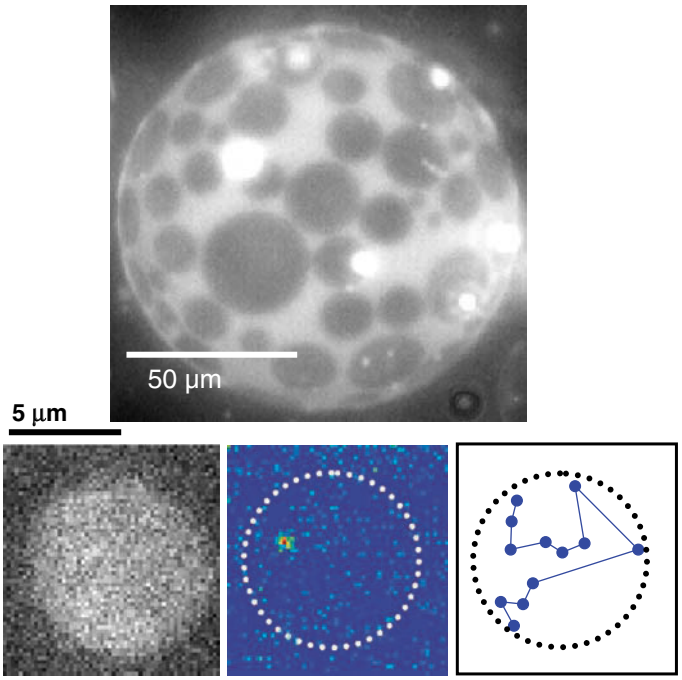


Figure 2.5. Individual lipids in a phase-separated giant unilamellar vesicle **(top)** were followed over time. Dual-color imaging was used to visualize the fluid domain in the green using a DiI stain **(bottom left)** and individual Cy5-labeled phospholipids **(bottom middle)** simultaneously. The lipid performed Brownian motion confined to the fluid-disordered domain. (S. Semrau, unpublished, 2008.)

far from the demixing point seen in GUVs did allow us to estimate the expected size of those nanodomains close to the demixing point. Analysis of the shape of macroscopic phase-separated 1:1:1 GUVs yields a line tension of 4 pN, which allowed us to estimate upper and lower limits for the size of phase-separated domains in the range of 10–30 nm [95]. This finding based on the mechanical properties of bilayers goes along with the lipid-shell models currently being discussed for cellular membranes [130].

A novel domain that comes into reach by single-molecule studies on bilayers systems is the analysis of mobility in complex and crowded environments [111,131]. The artificial environment of bilayers allows us precisely to adjust the amount of molecular players, including some that are attached to the membrane, and by that occupying space make it unavailable to other molecules. Such molecular crowding is believed to play an important role in cellular processes on membranes, in the cytosol, and in the nucleus, and is technically used for example, to enhance protein crystallization. Since part of the total volume is occupied by crowding agents, dynamic processes like signaling or random searching can be largely enhanced. Single-molecule microscopy does allow us now to directly follow individual players in, for example, a signaling pathway, and by such experiments we will be able to observe the influence of the crowding agent on the effectiveness and reliability of diffusion-controlled interactions. Similarly, groups are using membrane systems like GUVs for the construction of artificial biological systems [132], systems that at least in part mimic processes seen in nature.

As a conclusion to this subsection, we note that single-molecule microscopy has opened new opportunities for studies of artificial bilayer systems. Building on a vast amount of literature on the physical chemistry of monolayer and bilayers systems, the possibility to observe and follow an individual molecule within the bilayer or a vesicle as an artificial cell will spur novel and exciting research in a field localized among physics, material sciences, and cell biology [4].

2.6. Application 2—Live Cell Plasma Membrane

The successful utilization of single-molecule microscopy for the study of model membranes has yielded the promising perspective of a rather instantaneous application to living cells. However, the frequent occurrence of many species of endogenous fluorescent molecules inside cells has made such an application difficult. Cellular autofluorescence has been characterized in terms of spectral properties [133], lifetime [134], and spatial distribution [135]. In the visible regime, flavins [133] and lipofuscin [136] are currently regarded as the major source of endogenous fluorescence. Flavins are mainly located in mitochondria, whereas lipofuscins predominantly reside in lysosomes. In fluorescence images, both organelles appear as diffraction-limited spots randomly distributed in the cytoplasm of the cell. The high variability of the fluorescence intensity of such spots, even within one cell, makes unambiguous distinction between fluorophores and autofluorescence a challenging task.

In general, it turned out that the brightness of autofluorescent structures decreases with increasing wavelength, rendering red excitation beyond 600 nm the pragmatic choice [40,41]. However, frequently the biological problem defines both the cell type and the excitation color to use. For such cases, methods for the decomposition of optical signals originating from two spectrally distinct components may be used [137].

Having selected the proper dye, there are several ways for attaching it to the biomolecule of interest:

1. The purified or synthesized biomolecule is directly labeled covalently with a fluorescent marker and subsequently added to the living cell. This strategy has been applied for the study of small molecules like lipids [40] and for larger stable structures such as viruses [138]. In addition, exogenously applied lipid-anchored proteins were found to incorporate into the plasma membrane and acquire the correct signaling capacity [139].
2. The protein of interest is genetically fused to a fluorescent protein [140], a specific tag, or a protein that can be covalently labeled with an exogenously added dye (Snap-tag) [141]. In particular, since the spectroscopic properties of the available fluorescent protein mutants are inferior to those of many organic fluorophores [91], covalent labeling strategies have found widespread approval of the researchers.
3. Membrane proteins can be specifically labeled by fluorescent ligands such as Fab fragments [42, 43] or toxins [142].

Whereas specific monovalent labeling is rather straightforward for single-dye approaches, it represents a big challenge for SPT. Commonly, particles are coated with a mix of specific antibody and a blocking protein, typically bovine serum albumin (BSA). By reducing the amount of antibody, researchers attempt to reduce the likelihood of particles containing more than one accessible reactive group. As expected, the multivalency of the particle reduces the mobility of the tracer [19]. Indeed, results obtained by SPT can differ significantly from the undisturbed motion of the molecule of interest. In one study, residual cross-linking of the diffusing probe was found to reduce the diffusion coefficients significantly when compared to data obtained from tracking the same probe labeled via a fluorescent antibody [52]. A different study reported the alteration of the diffusion behavior from free diffusion to anomalous subdiffusion on labeling with quantum dots [143]. With the development of small passivated quantum dots, however, monovalency of a particle may indeed come into reach [144].

Most single-molecule studies on the plasma membrane have aimed at the identification of confinements to the free diffusion of the tracer. In particular, when molecules undergo direct or indirect interactions with the membrane skeleton, the diffusional motion becomes transiently restricted to membrane domains. Such confinements are highly relevant for theoretical modeling because they affect the interaction probability between arbitrary molecules. These confinement zones were postulated long before they were actually seen [145]. The first evidence for the existence of structural domains within the plasma membrane of cells came from biochemical extraction of the membrane together with subsequent sucrose density gradient centrifugation, which made it possible qualitatively to distinguish a distinct part of the plasma membrane that is not soluble in mild detergent [146]. Careful investigation of such detergent-resistant membranes (DRMs) further yielded their protein content, which was found to be consistently different from the remaining fractions of the membrane [147,148]. Of interest, DRMS were found to be enriched in proteins involved in signaling [149].

The physical origin and *in vivo* counterpart of DRMs are under debate. The detergent insolubility is most likely mediated by self-interaction between glycosphingolipids [15]. In addition, DRMs contain a distinct set of lipids [146], such as sphingolipids [150], fully saturated fatty acids [151], and cholesterol [146]. Detergent-resistant liposomes were found to be in liquid-ordered phase [15,152]. This ordered environment highly enhances the partitioning of proteins that are linked to saturated acyl chains, such as glycosylphosphatidylinositol (GPI)-anchored proteins or proteins acylated with myristate or palmitate [153,154]. Taken together, those studies were the experimental basis for the formulation of the "raft theory,"

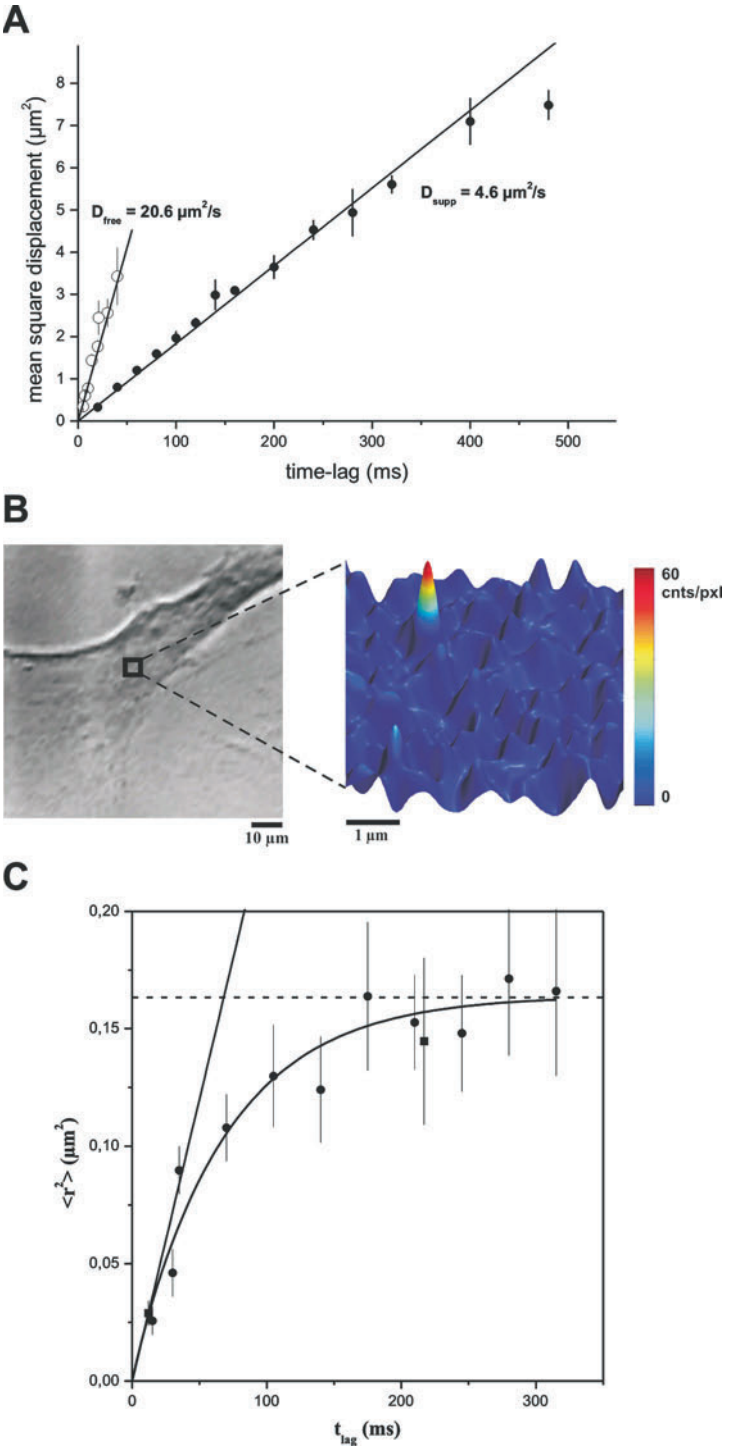


Figure 2.6. (continued)

in which the existence of stable platforms within the plasma membrane, membranes of organelles, and membranes of transport vesicles has been postulated [15].

Beside its initial importance for the discovery of plasma membrane domains, detergent extraction methods are still widely used to study the targeting of membrane proteins, either to rafts or to the remaining part of the plasma membrane. However, there are limitations to this method. First, artifacts due to relocalization of proteins during detergent extraction have been reported [155,156]. Second, a note of caution was recently raised against the structural interpretation of biochemical data by pointing out that rafts may be altered or even formed on detergent treatment [157]. Third, to avoid degradation processes, solubilization of cells is commonly performed at low temperatures, which most likely has an effect on the phase state of the domain and might alter its physical properties [158,159]. Note that analysis of the solubility in detergent reveals no information on the actual size, shape, and dynamical properties of domains.

SPT has been developed as a technique to study the compartmentalization of cell surfaces [8,17,18,94,160,161]. Restrictions to the lateral mobility have been observed for proteins such as Band 3 [162], E-cadherin [12], neural cell adhesion molecule (NCAM) [163], the transferrin receptor [164], and even for phospholipids themselves [8]. In most of those studies, the structural origin of the constraint was found to be the membrane-skeleton meshwork. Recent studies identified details in the molecular mechanism for transient immobilization of clustered GPI-anchored proteins [9,10,165].

In addition, lipid rafts have been reinterpreted in view of single-particle tracking data that revealed hop diffusion of essentially all investigated membrane probes [5]. The results of those studies provoked the current view of lipid rafts as highly dynamic objects, which are expected to mediate the coupling of the cortical actin meshwork to the exoplasmic membrane leaflet [166]. We and others therefore attempted to confirm or correct this model, based on the less invasive single-molecule imaging approach.

By pushing instrumentation to the theoretical limits, we were able to measure the mobility of single GPI-anchored proteins in the live cell plasma membrane at a resolution of $\sim$20 nm in space and <1 ms in time. For the raft protein CD59 we found no indication for hop diffusion and, as consequence, no indication for the presence of short-lived lipid rafts [42]. We further addressed this aspect by analyzing the mobility of the same GPI protein diffusing on the surface of cellular nanotubules— recently discovered thin structures that

Figure 2.6. A. Comparison of the diffusion in a free-standing ($\bigcirc$) and supported membrane ($\bullet$) [34]. For both, the mean square displacement increases linearly with time lag, yielding diffusion constants of $D_{\text{free}} = 20.6 \pm 0.9\ \mu\text{m}^2/\text{secand}\ D_{\text{supp}} = 4.6 \pm 0.1\ \mu\text{m}^2/\text{sec}$, respectively. (Reprinted by permission from Sonnleitner et al. [34], copyright 1999 by the Biophysical Society.) **B.** Single-lipid imaging in the plasma membrane of a human smooth muscle cell [40]. On the left, the cell is shown in a transmission light image at 40× magnification. The square at the center indicates the area selected for fluorescence imaging (*right*). The cell was pretreated with Cy5-labeled dioleylphosphatidylethanolamine (DOPE-Cy5) so that a low concentration of the fluorescent lipid was incorporated into the plasma membrane. The clearly resolved peak corresponds to a single DOPE-Cy5 molecule. (Reprinted from Schütz et al. [40].) **C.** Mobility of single Cy5-labeled dimyristoylphosphatidylethanolamine (DMPE-Cy5) molecules in the plasma membrane of living human smooth muscle cells [40]. The mean square displacement $\left\langle r^2 \right\rangle$ as a function of the time-lag shows saturation for $t_{\text{lag}} > 100$ ms. From a fit to the model of restricted diffusion, the diffusion constant within a membrane microdomain, $D = 0.6 \pm 0.04\ \mu\text{m}^2/\text{sec}$ (*solid line*), and the domain size of 700 ± 20 nm (*dashed line*) were determined. (Reprinted from Schütz et al. [40].)

connect different cells. The high resolution in space and time allowed us to follow the diffusion process not only in the longitudinal direction, but also along the circumference. We found perfectly isotropic mobility, in line with our previous observation that the anisotropic membrane skeleton does not impose a meshwork of periodic boundaries to the diffusion of exoplasmic plasma membrane constituents [44].

Although the general mechanism of hop diffusion remains controversial, there is no doubt that numerous membrane constituents become trapped in specific cell surface regions. Confinements to immobile domains were indeed reported for a fully saturated lipid in a human smooth muscle cell line [40] (Figure 2.6B and C), the membrane anchor of Ras- and Src-family proteins [45–47], a G protein–coupled receptor [57], and hemagglutinin from the influenza virus [55]. Whether these structures correspond in general to lipid rafts, represent a subclass of lipid rafts, or can be considered as different organization units remains to be solved.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF project Y250-B10) and the GEN-AU project of the Austrian Federal Ministry for Science and Research (GJS). This work is part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie (FOM), which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) (TS).

References

1. Singer, S. J., and G. L. Nicolson. 1972. The fluid mosaic model of the structure of cell membranes. Science 175:720–731.
2. Lipowsky, R., and E. Sackmann, editors. 1995. Structure and dynamics of membranes. Amsterdam: Elsevier.
3. Vereb, G., J. Szollosi, J. Matko, P. Nagy, T. Farkas, L. Vigh, L. Matyus, T. A. Waldmann, and S. Damjanovich. 2003. Dynamic, yet structured: the cell membrane three decades after the Singer-Nicolson model. Proc Natl Acad Sci USA 100:8053–8058.
4. Jacobson, K., O. G. Mouritsen, and R. G. Anderson. 2007. Lipid rafts: at a crossroad between cell biology and physics. Nat Cell Biol 9:7–14.
5. Kusumi, A., C. Nakada, K. Ritchie, K. Murase, K. Suzuki, H. Murakoshi, R. S. Kasai, J. Kondo, and T. Fujiwara. 2005. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. Annu Rev Biophys Biomol Struct 34:351–378.
6. Lenne, P. F., L. Wawrezinieck, F. Conchonaud, O. Wurtz, A. Boned, X. J. Guo, H. Rigneault, H. T. He, and D. Marguet. 2006. Dynamic molecular confinement in the plasma membrane by microdomains and the cytoskeleton meshwork. EMBO J 25:3245–3256.
7. Anderson, R. G., and K. Jacobson. 2002. A role for lipid shells in targeting proteins to caveolae, rafts, and other lipid domains. Science 296:1821–1825.
8. Fujiwara, T., K. Ritchie, H. Murakoshi, K. Jacobson, and A. Kusumi. 2002. Phospholipids undergo hop diffusion in compartmentalized cell membrane. J Cell Biol 157:1071–1081.
9. Chen, Y., W. R. Thelin, B. Yang, S. L. Milgram, and K. Jacobson. 2006. Transient anchorage of cross-linked glycosyl-phosphatidylinositol-anchored proteins depends on cholesterol, Src family kinases, caveolin, and phosphoinositides. J Cell Biol 175:169–178.
10. Suzuki, K. G., T. K. Fujiwara, F. Sanematsu, R. Iino, M. Edidin, and A. Kusumi. 2007. GPI-anchored receptor clusters transiently recruit Lyn and G{alpha} for temporary cluster immobilization and Lyn activation: single-molecule tracking study 1. J Cell Biol 177: 717–730.
11. Demond, A. L., K. D. Mossman, T. Starr, M. L. Dustin, and J. T. Groves. 2008. T cell receptor microcluster transport through molecular mazes reveals mechanism of translocation. Biophys J 94:3286–3292.

12. Sako, Y., A. Nagafuchi, S. Tsukita, M. Takeichi, and A. Kusumi. 1998. Cytoplasmic regulation of the movement of E-cadherin on the free cell surface as studied by optical tweezers and single particle tracking: corralling and tethering by the membrane skeleton. J Cell Biol 140:1227–1240.
13. Saxton, M. J. 1994. Anomalous diffusion due to obstacles: a Monte Carlo study. Biophys J 66:394–401.
14. Marguet, D., P. F. Lenne, H. Rigneault, and H. T. He. 2006. Dynamics in the plasma membrane: how to combine fluidity and order. EMBO J 25:3446–3457.
15. Simons, K., and E. Ikonen. 1997. Functional rafts in cell membranes. Nature 387:569–572.
16. Barak, L. S., and W. W. Webb. 1981. Fluorescent low density lipoprotein for observation of dynamics of individual receptor complexes on cultured human fibroblasts. J Cell Biol 90:595–604.
17. Kusumi, A., Y. Sako, and M. Yamamoto. 1993. Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. Biophys J 65:2021–2040.
18. Saxton, M. J., and K. Jacobson. 1997. Single-particle tracking: applications to membrane dynamics. Annu Rev Biophys Biomol Struct 26:373–399.
19. Lee, G. M., A. Ishihara, and K. A. Jacobson. 1991. Direct observation of Brownian motion of lipids in a membrane. Proc Natl Acad Sci USA 88:6274–6278.
20. Wilson, K. M., I. E. Morrison, P. R. Smith, N. Fernandez, and R. J. Cherry. 1996. Single particle tracking of cell-surface HLA-DR molecules using R-phycoerythrin labeled monoclonal antibodies and fluorescence digital imaging. J Cell Sci 109 (Pt 8):2101–2109.
21. Kao, H. P., and A. S. Verkman. 1994. Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. Biophys J 67:1291–1300.
22. Felsenfeld, D. P., D. Choquet, and M. P. Sheetz. 1996. Ligand binding regulates the directed movement of beta1 integrins on fibroblasts. Nature 383:438–440.
23. Anderson, C. M., G. N. Georgiou, I. E. Morrison, G. V. Stevenson, and R. J. Cherry. 1992. Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera. Low-density lipoprotein and influenza virus receptor mobility at 4 degrees C. J Cell Sci 101(Pt 2): 415–425.
24. Fein, M., J. Unkeless, F. Y. Chuang, M. Sassaroli, R. da Costa, H. Vaananen, and J. Eisinger. 1993. Lateral mobility of lipid analogues and GPI-anchored proteins in supported bilayers determined by fluorescent bead tracking. J Membr Biol 135:83–92.
25. Geerts, H., M. De Brabander, R. Nuydens, S. Geuens, M. Moeremans, J. De Mey, and P. Hollenbeck. 1987. Nanovid tracking: a new automatic method for the study of mobility in living cells based on colloidal gold and video microscopy. Biophys J 52: 775–782.
26. Thompson, R. E., D. R. Larson, and W. W. Webb. 2002. Precise nanometer localization analysis for individual fluorescent probes. Biophys J 82:2775–2783.
27. Kuno, M., D. P. Fromm, H. F. Hamann, A. Gallagher, and D. J. Nesbitt. 1999. Nonexponential "blinking" kinetics of single CdSe quantum dots: a universal power law behavior. J Chem Phys 112:3117–3120.
28. Schmidt, T., G. J. Schütz, W. Baumgartner, H. J. Gruber, and H. Schindler. 1995. Characterization of photophysics and mobility of single molecules in a fluid lipid membrane. J Phys Chem 99: 17662–17668.
29. Schmidt, T., G. J. Schütz, W. Baumgartner, H. J. Gruber, and H. Schindler. 1996. Imaging of single molecule diffusion. Proc Natl Acad Sci USA 93:2926–2929.
30. Funatsu, T., Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida. 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous-solution. Nature 374:555–559.
31. Sase, I., H. Miyata, J. E. Corrie, J. S. Craik, and K. Kinosita, Jr. 1995. Real time imaging of single fluorophores on moving actin with an epifluorescence microscope. Biophys J 69: 323–328.
32. Schütz, G. J., H. Schindler, and T. Schmidt. 1997. Single-molecule microscopy on model membranes reveals anomalous diffusion. Biophys J 73:1073–1080.
33. Schütz, G. J., W. Trabesinger, and T. Schmidt. 1998. Direct observation of ligand colocalization on individual receptor molecules. Biophys J 74:2223–2226.
34. Sonnleitner, A., G. J. Schutz, and T. Schmidt. 1999. Free Brownian motion of individual lipid molecules in biomembranes. Biophys J 77:2638–2642.
35. Harms, G. S., M. Sonnleitner, G. J. Schütz, H. J. Gruber, and T. Schmidt. 1999. Single-molecule anisotropy imaging. Biophys J 77:2864–2870.
36. Ke, P. C., and C. A. Naumann. 2001. Hindered diffusion in polymer-tethered phosopholipid monolayers at the air–water interface: a single molecule fluorescence imaging study. Langmuir 17:5076–5081.
37. Ke, P. C., and C. A. Naumann. 2001. Single molecule fluorescence imaging of phospholipid monolayers at the air–water interface. Langmuir 17:3727–3733.

38. Deverall, M. A., E. Gindl, E. K. Sinner, H. Besir, J. Ruehe, M. J. Saxton, and C. A. Naumann. 2005. Membrane lateral mobility obstructed by polymer-tethered lipids studied at the single molecule level. Biophys J 88: 1875–1886.
39. Kiessling, V., J. M. Crane, and L. K. Tamm. 2006. Transbilayer effects of raft-like lipid domains in asymmetric planar bilayers measured by single molecule tracking. Biophys J 91:3313–3326.
40. Schütz, G. J., G. Kada, V. P. Pastushenko, and H. Schindler. 2000. Properties of lipid microdomains in a muscle cell membrane visualized by single molecule microscopy. EMBO J 19:892–901.
41. Schütz, G. J., V. P. Pastushenko, H. J. Gruber, H.-G. Knaus, B. Pragl, and H. Schindler. 2000. 3D Imaging of individual ion channels in live cells at 40 nm resolution. Single Mol. 1:25–31.
42. Wieser, S., M. Moertelmaier, E. Fuertbauer, H. Stockinger, and G. J. Schutz. 2007. (Un)Confined diffusion of CD59 in the plasma membrane determined by high-resolution single molecule microscopy. Biophys J 92: 3719–3728.
43. Drbal, K., M. Moertelmaier, C. Holzhauser, A. Muhammad, E. Fuertbauer, S. Howorka, M. Hinterberger, H. Stockinger, and G. J. Schutz. 2007. Single-molecule microscopy reveals heterogeneous dynamics of lipid raft components upon TCR engagement. Int Immunol 19:675–684.
44. Wieser, S., G. J. Schutz, M. E. Cooper, and H. Stockinger. 2007. Single molecule diffusion analysis on cellular nanotubules: implications on plasma membrane structure below the diffraction limit. Appl Phys Lett 91:233901.
45. Lommerse, P. H., K. Vastenhoud, N. J. Pirinen, A. I. Magee, H. P. Spaink, and T. Schmidt. 2006. Single-molecule diffusion reveals similar mobility for the Lck, H-ras, and K-ras membrane anchors. Biophys J 91:1090–1097.
46. Lommerse, P. H., B. E. Snaar-Jagalska, H. P. Spaink, and T. Schmidt. 2005. Single-molecule diffusion measurements of H-Ras at the plasma membrane of live cells reveal microdomain localization upon activation. J Cell Sci 118:1799–1809.
47. Lommerse, P. H., G. A. Blab, L. Cognet, G. S. Harms, B. E. Snaar-Jagalska, H. P. Spaink, and T. Schmidt. 2004. Single-molecule imaging of the H-Ras membrane-anchor reveals domains in the cytoplasmic leaflet of the cell membrane. Biophys J 86:609–616.
48. Harms, G. S., L. Cognet, P. H. Lommerse, G. A. Blab, H. Kahr, R. Gamsjager, H. P. Spaink, N. M. Soldatov, C. Romanin, and T. Schmidt. 2001. Single-molecule imaging of l-type Ca(2+) channels in live cells. Biophys J 81:2639–2646.
49. Vrljic, M., S. Y. Nishimura, S. Brasselet, W. E. Moerner, and H. M. McConnell. 2002. Translational diffusion of individual class II MHC membrane proteins in cells. Biophys J 83:2681–2692.
50. Vrljic, M., S. Y. Nishimura, W. E. Moerner, and H. M. McConnell. 2005. Cholesterol depletion suppresses the translational diffusion of class II major histocompatibility complex proteins in the plasma membrane. Biophys J 88:334–347.
51. Nishimura, S. Y., M. Vrljic, L. O. Klein, H. M. McConnell, and W. E. Moerner. 2006. Cholesterol depletion induces solid-like regions in the plasma membrane. Biophys J 90: 927–938.
52. Umemura, Y. M., M. Vrljic, S. Y. Nishimura, T. K. Fujiwara, K. G. Suzuki, and A. Kusumi. 2008. Both MHC class II and its GPI-anchored form undergo hop diffusion as observed by single-molecule tracking. Biophys J 95:435–450.
53. Murase, K., T. Fujiwara, Y. Umemura, K. Suzuki, R. Iino, H. Yamashita, M. Saito, H. Murakoshi, K. Ritchie, and A. Kusumi. 2004. Ultrafine membrane compartments for molecular diffusion as revealed by single molecule techniques. Biophys J 86:4075–4093.
54. Douglass, A. D., and R. D. Vale. 2005. Single-molecule microscopy reveals plasma membrane microdomains created by protein-protein networks that exclude or trap signaling molecules in T cells. Cell 121: 937–950.
55. Hess, S. T., T. J. Gould, M. V. Gudheti, S. A. Maas, K. D. Mills, and J. Zimmerberg. 2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. Proc Natl Acad Sci USA 104:17370–17375.
56. Manley, S., J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz. 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods 5:155–157.
57. Jacquier, V., M. Prummer, J. M. Segura, H. Pick, and H. Vogel. 2006. Visualizing odorant receptor trafficking in living cells down to the single-molecule level. Proc Natl Acad Sci USA 103:14325–14330.
58. James, J. R., S. S. White, R. W. Clarke, A. M. Johansen, P. D. Dunne, D. L. Sleep, W. J. Fitzgerald, S. J. Davis, and D. Klenerman. 2007. Single-molecule level analysis of the subunit composition of the T cell receptor on live T cells. Proc Natl Acad Sci USA 104:17662–17667.

59. Morimatsu, M., H. Takagi, K. G. Ota, R. Iwamoto, T. Yanagida, and Y. Sako. 2007. Multiple-state reactions between the epidermal growth factor receptor and Grb2 as observed by using single-molecule analysis. Proc Natl Acad Sci USA 104:18013–18018.
60. Sako, Y., S. Minoghchi, and T. Yanagida. 2000. Single-molecule imaging of EGFR signalling on the surface of living cells. Nat Cell Biol 2:168–172.
61. Ueda, M., Y. Sako, T. Tanaka, P. Devreotes, and T. Yanagida. 2001. Single-molecule analysis of chemotactic signaling in *Dictyostelium* cells. Science 294:864–867.
62. Füreder-Kitzmüller, E., J. Hesse, A. Ebner, H. J. Gruber, and G. J. Schütz. 2005. Non-exponential bleaching of single bioconjugated Cy5 molecules. Chem Phys Lett 404:13–18.
63. Hecht, E. 1987. Optics. Reading, MA: Addison-Wesley.
64. Enderlein, J., E. Toprak, and P. R. Selvin. 2006. Polarization effect on position accuracy of fluorophore localization. Opt Express 14:8111–8120.
65. Pohl, D. W., W. Denk, and M. Lanz. 1984. Optical stethoscopy: image recording with resolution λ/20. Appl Phys Lett 44:651–653.
66. Betzig, E., and J. K. Trautman. 1992. Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit. Science 257:189–195.
67. Denk, W., J. H. Strickler, and W. W. Webb. 1990. Two-photon laser scanning fluorescence microscopy. Science 248:73–76.
68. Klar, T. A., S. Jakobs, M. Dyba, A. Egner, and S. W. Hell. 2000. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. Proc Natl Acad Sci USA 97:8206–8210.
69. Pawley, J. B., editor. 1995. Handbook of biological confocal microscopy. New York: Plenum Press.
70. Gustafsson, M. G. 2005. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proc Natl Acad Sci USA 102:13081–13086.
71. Heintzmann, R., T. M. Jovin, and C. Cremer. 2002. Saturated patterned excitation microscopy—a concept for optical resolution improvement. J Opt Soc Am A – Opt Image Sci Vis 19:1599–1609.
72. Hell, S. W. 2007. Far-field optical nanoscopy. Science 316:1153–1158.
73. Betzig, E. 1995. Proposed method for molecular optical imaging. Opt Lett 20:237–239.
74. Bobroff, N. 1986. Position measurement with a resolution and noise-limited instrument. Rev Sci Instrum 57:1152–1157.
75. Yildiz, A., J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin. 2003. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science 300:2061–2065.
76. Ober, R. J., S. Ram, and E. S. Ward. 2004. Localization accuracy in single-molecule microscopy. Biophys J 86:1185–1200.
77. Ha, T., T. Enderle, D. S. Chemla, and S. Weiss. 1996. Dual-molecule spectroscopy: molecular rulers for the study of biological macromolecules. IEEE J Sel Top Quant Electr 2: 1115–1128.
78. van Oijen, A. M., J. Kohler, J. Schmidt, M. Muller, and G. J. Brakenhoff. 1999. Far-field fluorescence microscopy beyond the diffraction limit. J Opt Soc Am A 16:909–915.
79. Trabesinger, W., G. J. Schütz, H. J. Gruber, H. Schindler, and T. Schmidt. 1999. Detection of individual oligonucleotide pairing by single-molecule microscopy. Anal Chem 71:279–283.
80. Trabesinger, W., B. Hecht, U. P. Wild, G. J. Schütz, H. Schindler, and T. Schmidt. 2001. Statistical analysis of single-molecule colocalization assays. Anal Chem 73:1100–1105.
81. Baumgartner, W., G. J. Schütz, J. Wiegand, N. Golenhofen, and D. Drenckhahn. 2003. Cadherin function probed by laser tweezer and single molecule fluorescence in vascular endothelial cells. J Cell Sci 116: 1001–1011.
82. Yildiz, A., M. Tomishige, R. D. Vale, and P. R. Selvin. 2004. Kinesin walks hand-over-hand. Science 303: 676–678.
83. Gordon, M. P., T. Ha, and P. R. Selvin. 2004. Single-molecule high-resolution imaging with photobleaching. Proc Natl Acad Sci USA 101:6462–6465.
84. Qu, X., D. Wu, L. Mets, and N. F. Scherer. 2004. Nanometer-localized multiple single-molecule fluorescence microscopy. Proc Natl Acad Sci USA 101:11298–11303.
85. Patterson, G. H., and J. Lippincott-Schwartz. 2002. A photoactivatable GFP for selective photolabeling of proteins and cells. Science 297:1873–1877.
86. Betzig, E., G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. 2006. Imaging intracellular fluorescent proteins at nanometer resolution. Science 313:1642–1645.
87. Hess, S. T., T. P. Girirajan, and M. D. Mason. 2006. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophys J 91:4258–4272.

88. Rust, M., M. Bates, and X. Zhuang. 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods 3:793–795.
89. Shroff, H., C. G. Galbraith, J. A. Galbraith, and E. Betzig. 2008. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nat Methods 5:417–423.
90. Orrit, M., and J. Bernard. 1990. Single pentacene molecules detected by fluorescence excitation in a *p*-terphenyl crystal. Phys Rev Lett 65:2716–2719.
91. Harms, G. S., L. Cognet, P. H. Lommerse, G. A. Blab, and T. Schmidt. 2001. Autofluorescent proteins in single-molecule research: applications to live cell imaging microscopy. Biophys J 80:2396–2408.
92. Holtzer, L., T. Meckel, and T. Schmidt. 2007. Nanometric three-dimensional tracking of individual quantum dots in cells. Appl Phys Lett 90:053902.
93. Prabhat, P., Z. Gan, J. Chao, S. Ram, C. Vaccaro, S. Gibbons, R. J. Ober, and E. S. Ward. 2007. Elucidation of intracellular recycling pathways leading to exocytosis of the Fc receptor, FcRn, by using multifocal plane microscopy. Proc Natl Acad Sci USA 104:5889–5894.
94. Ghosh, R. N., and W. W. Webb. 1994. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. Biophys J 66:1301–1318.
95. Semrau, S., and T. Schmidt. 2006. Particle image correlation spectroscopy (PICS) Retrieving nanometer-scale correlations from high-density single-molecule position data. Biophys J 92:613–621.
96. Falck, E., T. Rog, M. Karttunen, and I. Vattulainen. 2008. Lateral diffusion in lipid membranes through collective flows. J Am Chem Soc 130:44–45.
97. Almeida, P. F., W. L. Vaz, and T. E. Thompson. 2005. Lipid diffusion, free area, and molecular dynamics simulations. Biophys J 88:4434–4438.
98. Falck, E., M. Patra, M. Karttunen, M. T. Hyvonen, and I. Vattulainen. 2005. Response to comment by Almeida et al.: free area theories for lipid bilayers—predictive or not? Biophys J 89:745–752.
99. Berg, H. C. 1983. Random walks in biology. Princeton, New Jersey: Princeton University Press.
100. Saffman, P. G., and M. Delbruck. 1975. Brownian motion in biological membranes. Proc Natl Acad Sci USA 72:3111–3113.
101. Hughes, B. D., B. A. Pailthorpe, and L. R. White. 1981. The translational and rotational drag on a cylinder moving in a membrane. J. Fluid. Mech. 110:349–372.
102. Hughes, B. D., B. A. Pailthorpe, L. R. White, and W. H. Sawyer. 1982. Extraction of membrane microviscosity from translational and rotational diffusion coefficients. Biophys J 37:673–676.
103. Petrov, E. P., and P. Schwille. 2008. Translational diffusion in lipid membranes beyond the Saffman-Delbruck approximation. Biophys J 94:L41–43.
104. Gambin, Y., R. Lopez-Esparza, M. Reffay, E. Sierecki, N. S. Gov, M. Genest, R. S. Hodges, and W. Urbach. 2006. Lateral mobility of proteins in liquid membranes revisited. Proc Natl Acad Sci USA 103: 2098–2102.
105. Saxton, M. J. 1995. Single-particle tracking: effects of corrals. Biophys J 69:389–398.
106. Simson, R., E. D. Sheets, and K. Jacobson. 1995. Detection of temporary lateral confinement of membrane proteins using single-particle tracking analysis. Biophys J 69:989–993.
107. Wieser, S., and G. J. Schütz 2008. Tracking single molecules in the live cell plasma membrane—Do's and Don't's. Methods 46:131–140.
108. Feder, T. J., I. Brust-Mascher, J. P. Slattery, B. Baird, and W. W. Webb. 1996. Constrained diffusion or immobile fraction on cell surfaces: a new interpretation. Biophys J 70:2767–2773.
109. Smith, P. R., I. E. Morrison, K. M. Wilson, N. Fernandez, and R. J. Cherry. 1999. Anomalous diffusion of major histocompatibility complex class I molecules on HeLa cells determined by single particle tracking. Biophys J 76:3331–3344.
110. Ritchie, K., X. Y. Shan, J. Kondo, K. Iwasawa, T. Fujiwara, and A. Kusumi. 2005. Detection of non-Brownian diffusion in the cell membrane in single molecule tracking. Biophys J 88:2266–2277.
111. Guigas, G., and M. Weiss. 2007. Sampling the cell with anomalous diffusion—the discovery of slowness. Biophys J 94:90–94.
112. Nicolau, D. V., Jr., J. F. Hancock, and K. Burrage. 2007. Sources of anomalous diffusion on cell membranes: a Monte Carlo study. Biophys J 92:1975–1987.
113. Saxton, M. J. 1996. Anomalous diffusion due to binding: a Monte Carlo study. Biophys J 70:1250–1262.
114. Saxton, M. J. 2007. A biological interpretation of transient anomalous subdiffusion. II. Reaction kinetics. Biophys J 94:760–771.
115. Saxton, M. J. 2007. A biological interpretation of transient anomalous subdiffusion. I. Qualitative model. Biophys J 92:1178–1191.
116. Saxton, M. J. 1993. Lateral diffusion in an archipelago. Single-particle diffusion. Biophys J 64:1766–1780.

117. Martin, D. S., M. B. Forstner, and J. A. Kas. 2002. Apparent subdiffusion inherent to single particle tracking. Biophys J 83:2109–2117.
118. Sako, Y., and A. Kusumi. 1995. Barriers for lateral diffusion of transferrin receptor in the plasma membrane as characterized by receptor dragging by laser tweezers: fence versus tether. J Cell Biol 129: 1559–1574.
119. Daumas, F., N. Destainville, C. Millot, A. Lopez, D. Dean, and L. Salome. 2003. Confined diffusion without fences of a g-protein–coupled receptor as revealed by single particle tracking. Biophys J 84:356–366.
120. Lillemeier, B. F., J. R. Pfeiffer, Z. Surviladze, B. S. Wilson, and M. M. Davis. 2006. Plasma membrane-associated proteins are clustered into islands attached to the cytoskeleton. Proc Natl Acad Sci USA 103: 18992–18997.
121. King, M. R. 2004. Apparent 2-D diffusivity in a ruffled cell membrane. J Theor Biol 227:323–326.
122. Reister, E., and U. Seifert. 2005. Lateral diffusion of a protein on a fluctuating membrane. Europhys Lett 71:859–865.
123. Aizenbud, B. M., and N. D. Gershon. 1982. Diffusion of molecules on biological membranes of nonplanar form. A theoretical study. Biophys J 38:287–293.
124. Rustom, A., R. Saffrich, I. Markovic, P. Walther, and H. H. Gerdes. 2004. Nanotubular highways for intercellular organelle transport. Science 303:1007–1010.
125. Goulian, M., and S. M. Simon. 2000. Tracking single proteins within cells. Biophys J 79:2188–2198.
126. Destainville, N., and L. Salome. 2006. Quantification and correction of systematic errors due to detector time-averaging in single-molecule tracking experiments. Biophys J 90:L17–19.
127. Weiss, S. 1999. Fluorescence spectroscopy of single biomolecules. Science 283:1676–1683.
128. Baumgart, T., S. T. Hess, and W. W. Webb. 2003. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. Nature 425:821–824.
129. Cognet, L., G. S. Harms, G. A. Blab, P. H. M. Lommerse, and T. Schmidt. 2000. Simultaneous dual-color and dual-polarization imaging of single molecules. Appl Phys Lett 77:4052–4054.
130. Jacobson, K., E. D. Sheets, and R. Simson. 1995. Revisiting the fluid mosaic model of membranes. Science 268:1441–1442.
131. Hac, A. E., H. M. Seeger, M. Fidorra, and T. Heimburg. 2005. Diffusion in two-component lipid membranes—a fluorescence correlation spectroscopy and Monte Carlo simulation study. Biophys J 88: 317–333.
132. Loose, M., E. Fischer-Friedrich, J. Ries, K. Kruse, and P. Schwille. 2008. Spatial regulators for bacterial cell division self-organize into surface waves *in vitro*. Science 320:789–792.
133. Benson, R. C., R. A. Meyer, M. E. Zaruba, and G. M. McKhann. 1979. Cellular autofluorescence—is it due to flavins? J Histochem Cytochem 27:44–48.
134. Konig, K., P. T. So, W. W. Mantulin, B. J. Tromberg, and E. Gratton. 1996. Two-photon excited lifetime imaging of autofluorescence in cells during UVA and NIR photostress. J Microsc 183:197–204.
135. Andersson, H., T. Baechi, M. Hoechl, and C. Richter. 1998. Autofluorescence of living cells. J Microsc 191(Pt 1):1–7.
136. Schnell, S. A., W. A. Staines, and M. W. Wessendorf. 1999. Reduction of lipofuscin-like autofluorescence in fluorescently labeled tissue. J Histochem Cytochem 47:719–730.
137. Moertelmaier, M. A., E. J. Kögler, J. Hesse, M. Sonnleitner, L. A. Huber, and G. J. Schütz. 2002. Single molecule microscopy in living cells: subtraction of autofluorescence based on two color recording. Single Mol. 3:225–231.
138. Seisenberger, G., M. U. Ried, T. Endress, H. Buning, M. Hallek, and C. Brauchle. 2001. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. Science 294:1929–1932.
139. van den Berg, C. W., T. Cinek, M. B. Hallett, V. Horejsi, and B. P. Morgan. 1995. Exogenous glycosyl phosphatidylinositol-anchored CD59 associates with kinases in membrane clusters on U937 cells and becomes Ca(2+)-signaling competent. J Cell Biol 131:669–677.
140. Shaner, N. C., P. A. Steinbach, and R. Y. Tsien. 2005. A guide to choosing fluorescent proteins. Nat Methods 2:905–909.
141. Gronemeyer, T., G. Godin, and K. Johnsson. 2005. Adding value to fusion proteins through covalent labelling. Curr Opin Biotechnol 16:453–458.
142. Freudenthaler, G., M. Axmann, H. Schindler, B. Pragl, H. G. Knaus, and G. J. Schütz. 2002. Ultrasensitive pharmacological characterisation of the voltage-gated potassium channel K(V)1.3 studied by single-molecule fluorescence microscopy. Histochem Cell Biol 117:197–202.
143. Nechyporuk-Zloy, V., P. Dieterich, H. Oberleithner, C. Stock, and A. Schwab. 2008. Dynamics of single potassium channel proteins in the plasma membrane of migrating cells. Am J Physiol Cell Physiol 294: C1096–1102.

144. Howarth, M., W. Liu, S. Puthenveetil, Y. Zheng, L. F. Marshall, M. M. Schmidt, K. D. Wittrup, M. G. Bawendi, and A. Y. Ting. 2008. Monovalent, reduced-size quantum dots for imaging receptors on living cells. Nat Methods 5:397–399.
145. Morrisett, J. D., H. J. Pownall, R. T. Plumlee, L. C. Smith, and Z. E. Zehner. 1975. Multiple thermotropic phase transitions in *Escherichia coli* membranes and membrane lipids. A comparison of results obtained by nitroxyl stearate paramagnetic resonance, pyrene excimer fluorescence, and enzyme activity measurements. J Biol Chem 250:6969–6976.
146. Brown, D. A., and J. K. Rose. 1992. Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. Cell 68:533–544.
147. Melkonian, K. A., T. Chu, L. B. Tortorella, and D. A. Brown. 1995. Characterization of proteins in detergent-resistant membrane complexes from Madin-Darby canine kidney epithelial cells. Biochemistry 34: 16161–16170.
148. Lisanti, M. P., P. E. Scherer, J. Vidugiriene, Z. Tang, A. Hermanowski-Vosatka, Y. H. Tu, R. F. Cook, and M. Sargiacomo. 1994. Characterization of caveolin-rich membrane domains isolated from an endothelial-rich source: implications for human disease. J Cell Biol 126:111–126.
149. Horejsi, V., K. Drbal, M. Cebecauer, J. Cerny, T. Brdicka, P. Angelisova, and H. Stockinger. 1999. GPI-microdomains: a role in signalling via immunoreceptors. Immunol Today 20:356–361.
150. Brown, R. E. 1998. Sphingolipid organization in biomembranes: what physical studies of model membranes reveal. J Cell Sci 111(Pt 1):1–9.
151. Schroeder, R., E. London, and D. Brown. 1994. Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. Proc Natl Acad Sci USA 91:12130–12134.
152. Brown, D. A., and E. London. 1998. Structure and origin of ordered lipid domains in biological membranes. J Membr Biol 164:103–114.
153. Brown, D. A., and E. London. 1998. Functions of lipid rafts in biological membranes. Annu Rev Cell Dev Biol 14:111–136.
154. Resh, M. D. 1999. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. Biochim Biophys Acta 1451:1–16.
155. Peirce, M., and H. Metzger. 2000. Detergent-resistant microdomains offer no refuge for proteins phosphorylated by the IgE receptor. J Biol Chem 275:34976–34982.
156. Kurzchalia, T. V., E. Hartmann, and P. Dupree. 1995. Guilt by insolubility—does a protein's detergent insolubility reflect a caveolar location? Trends Cell Biol 5:187–189.
157. Heerklotz, H. 2002. Triton promotes domain formation in lipid raft mixtures. Biophys J 83:2693–2701.
158. London, E., and D. A. Brown. 2000. Insolubility of lipids in triton X-100: physical origin and relationship to sphingolipid/cholesterol membrane domains (rafts). Biochim Biophys Acta 1508:182–195.
159. Li, X. M., J. M. Smaby, M. M. Momsen, H. L. Brockman, and R. E. Brown. 2000. Sphingomyelin interfacial behavior: the impact of changing acyl chain composition. Biophys J 78:1921–1931.
160. Edidin, M., S. C. Kuo, and M. P. Sheetz. 1991. Lateral movements of membrane glycoproteins restricted by dynamic cytoplasmic barriers. Science 254:1379–1382.
161. Lee, G. M., F. Zhang, A. Ishihara, C. L. McNeil, and K. A. Jacobson. 1993. Unconfined lateral diffusion and an estimate of pericellular matrix viscosity revealed by measuring the mobility of gold-tagged lipids. J Cell Biol 120:25–35.
162. Tomishige, M., Y. Sako, and A. Kusumi. 1998. Regulation mechanism of the lateral diffusion of band 3 in erythrocyte membranes by the membrane skeleton. J Cell Biol 142:989–1000.
163. Simson, R., B. Yang, S. E. Moore, P. Doherty, F. S. Walsh, and K. A. Jacobson. 1998. Structural mosaicism on the submicron scale in the plasma membrane. Biophys J 74:297–308.
164. Sako, Y., and A. Kusumi. 1994. Compartmentalized structure of the plasma membrane for receptor movements as revealed by a nanometer-level motion analysis. J Cell Biol 125:1251–1264.
165. Suzuki, K. G., T. K. Fujiwara, M. Edidin, and A. Kusumi. 2007. Dynamic recruitment of phospholipase C{gamma} at transiently immobilized GPI-anchored receptor clusters induces IP3-Ca2+ signaling: single-molecule tracking study 2. J Cell Biol 177:731–742.
166. Pike, L. J. 2006. Rafts defined. J Lipid Res 47:1597–1598.

3

Single-Molecule Imaging in Live Cells

Jie Xiao

Abstract This chapter provides a comprehensive overview of how single-molecule imaging is achieved in live cells. The main focus is on fluorescent proteins, which are the most widely used fluorescent labels for live-cell imaging. The chromophore structures and the associated photochemical and photophysical properties of fluorescent proteins are discussed in detail, with a particular focus on how they influence single-molecule imaging in live cells. A few fluorescent proteins in the yellow-to-red spectral range, including newly discovered photoinducible ones, are selected for more detailed discussions due to their superior properties in single-molecule imaging. Special considerations for live-cell imaging and general instrumentations for single-molecule detection are also described. Finally, a few representative applications using single-molecule imaging in live cells are provided to illustrate how important biological knowledge can be obtained using this powerful technique.

3.1. Introduction

Nearly 20 years ago, when the first single molecule was detected in solid at extreme temperatures [1,2], it was hard to foresee imaging a single molecule in the noisy background of a live cell. Yet now, due to the efforts of many research groups and technical advances in modern microscopes, optics, and detectors, single-molecule imaging[1] in live cells is no longer impossible.

Why is it important to image single molecule inside live cells? The reasons are multifold. First, many biologically important molecules such as transcription factors exist at low copy numbers [11,12]. In yeast, nearly 40% of genes are expressed at such low levels that their gene products are not detectable using standard fluorescence microscopy or quantitative Western blotting [11]. A more sensitive detection method—preferably at the single-molecule level—would reveal what is otherwise imperceptible. Second, due to the finite-number effect,

J. Xiao • Johns Hopkins University School of Medicine, 708 WBSB, Baltimore, MD 21205, USA

[1] Here and throughout, single-molecule imaging refers to single-fluorophore imaging. Alternatively, a single molecule can be imaged by labeling a single molecule using multiple fluorophores or detected by amplifying the fluorescent signal using a fluorogenic substrate [3–10]. These approaches are discussed briefly at the end of the chapter.

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_3,

the interactions of molecules at low copy numbers are often stochastic [13]. This stochasticity has been proven to contribute to molecular noise in various cellular processes such as gene expression [14–16], signal transduction [17–20], and cell fate decisions [21–25]. However, this stochasticity is often masked in ensemble measurements. Single-molecule experiments avoid ensemble averaging and allow the stochasticity of molecular interactions to be observed, providing unprecedented details about the origins and roles of molecular noise. Third, all cells (including bacterial cells) exhibit considerable heterogeneity in their cellular environments. Depending on where and when a molecule is present in the cell, its behavior may change accordingly. Single-molecule imaging allows the heterogeneity of molecular behaviors to be probed in time and space, directly correlating them with cellular states. Finally, *in vitro* single-molecule studies have contributed significantly to the understanding of how individual molecules and macromachineries work. Taking these experiments into live cells allows the direct observation of the actions of these molecules and molecular machines in a native cellular context, establishing significant physiological relevance.

It is necessary to note that this chapter focuses on the formation of an image of a single molecule in live cells on an array detector such as a charge-coupled device (CCD) rather than the collection of a signal originating from a single molecule on a point detector such as a photo avalanche diode (PAD). In this sense, single-molecule detection in live cells by fluorescence correlation spectroscopy (FCS) [26–28] and by enzymatic amplification of a fluorescent signal [9, 10] are not included, and readers are referred to the original articles and Chapter 8 for further discussion. The main focus of this chapter is to elucidate how single-molecule imaging in live cells is achieved—which fluorescent label is best suited for this purpose, what special considerations are needed, and so on—and how it can be employed to address significant biological problems. Readers are assumed to have basic knowledge of fluorescence microscopy and, wherever as needed, are referred to available literature for further discussion.

3.2. Fluorescent Labels

To probe the dynamics of biomolecules in live cells using fluorescence microscopy, the molecules have to be labeled specifically using a fluorescent probe. There are three types of labels: small organic dyes, fluorescent semiconductor quantum dots (QDs), and fluorescent proteins (FPs) (Table 3.1).

Small organic dyes need to be conjugated to purified biomolecules through *in vitro* chemical reactions and reintroduced into cells by microinjection [29]. Cell

Table 3.1. A comparison of the three commonly used fluorescent labeling schemes

	Size (nm)	Labeling	Delivery	Brightness	Photostability
Organic dyes	<1	*In vitro* chemical reaction, peptide targeting	Perfusion, microinjection	High	High
Quantum dots	10–30	*In vitro* chemical reaction, streptavidin/antibody conjugation	Endocytosis, microinjection	Extremely high	Extremely high
Fluorescent proteins	3	Genetic fusion	Endogenous	Low to high	Low to high

membrane–permeable small organic dyes can also be incubated with cells expressing the protein of interest genetically fused with a short peptide sequence that binds the dye specifically [30]. Cells are then washed to remove excess dye molecules before imaging. One advantage of using small organic dyes is that because of their small size (<1 nm), they rarely perturb the functionality of labeled biomolecules. Some organic dyes are also brighter and more photostable than FPs, and their emission spectra are far red–shifted compared to that of currently available FPs, facilitating single-molecule imaging in live cells. Indeed, organic dyes such as Cy3 [31], Cy5 [32], BodipyTR [33], and Atto647 [34] have been employed in live-cell imaging at the single-molecule level to probe receptor and membrane dynamics on the cell surface. New ones, such as the photoactivatable azido-DCDHF [35,36], are also emerging. However, this type of labeling often encounters a large fluorescence background due to the incomplete removal of excess dye molecules in the labeling reaction and nonspecific binding of dye molecules to other cellular components. For these reasons, small organic dyes are not widely used in single-molecule imaging in live cells.

Quantum dots are inorganic nanocrystals that emit strong fluorescence between 400 and 1,350 nm, depending on their sizes and chemical compositions [37]. QDs have broad and overlapping absorbance bands but narrow emission bands; therefore, one single-wavelength excitation, for example, at 488 nm, of QDs of different sizes will lead to distinct and spectrally separable emission in the visible range. QDs are often coated with streptavidin or antibodies to allow their conjugation with specific cellular target or macromolecules. The fluorescence intensities of QDs are 10- to 100-fold stronger than those of FPs and small organic dyes, and they are extremely photostable, permitting long-time tracking of tagged molecules. However, the large size of QDs (10–30 nm) precludes them from passing through cell membranes. Delivery to the inside of eukaryotic cells is often by microinjection or peptide- mediated endocytosis [38]; there is no study reporting the delivery of QDs to bacterial cells due to the presence of the cell wall.

Because small organic dyes and QDs have to be delivered into the cell through invasive methods, they are most commonly used to label proteins or lipids on the outer surface of the cell membrane. In contrast, the genetically encoded FPs are the most popular fluorescent labels for specifically labeling of molecules inside live cells. Although the size of FPs (3 nm) is larger than that of small organic dyes, and the fluorescence intensity and photostability of FPs are less impressive than those of QDs, they can be genetically engineered through molecular biological means and do not require the introduction of any exogenous cofactors to become fluorescent. The labeling of a protein by a FP is extremely specific, and the expression level of a FP fusion protein can be tuned to allow single-FP molecule detection. Because of these reasons, this chapter focuses on using FPs as the fluorescent label for single-molecule imaging in live cells. The structural, biochemical, and especially photochemical and photophysical properties (whenever available) of a collection of FPs are examined in detail, and their suitability for single-molecule imaging is discussed. The emphasis is on the underlying biophysics of FPs, knowledge of which is the prerequisite for the appropriate selection of a FP for a particular experimental design, optimization of imaging conditions, correct interpretation of single-molecule experimental data, and successful extraction of biological information from single-molecule images. A brief introduction to the discovery and structure of the *Aequorea victoria* green fluorescent protein (GFP) is first given, followed by general discussions of important spectral, photophysical, and biochemical properties of FPs, and, finally of a collection of FPs that have potential for or have been demonstrated in single-molecule, live-cell imaging.

3.3. Green Fluorescent Protein

3.3.1. Discovery of GFP

The green fluorescence of the light organs of the jellyfish *Aequorea victoria* was first described by Davenport and Nicol in 1955 [39], but the discovery of GFP, which is responsible for the green fluorescence, did not come until the early 1960s, when Osamu Shimomura and his coworkers purified a Ca^{2+}-dependent bioluminescent protein, aequorin, from luminous jellyfish at the Friday Harbor Laboratories, University of Washington [40]. They found that another protein, which was not luminescent but exhibited "a very bright, greenish fluorescence in the ultraviolet of a Mineralite," coexisted with aequorin during the early stages of the purification. They called this protein "green protein," and it was renamed "green fluorescent protein" a few years later by Hastings and Morin [41,42]. Hastings and Morin also suggested that the natural green glow of jellyfish came from the green fluorescence of GFP that was the result of the energy transferred from the aequorin protein: When aequorin binds to Ca^{2+}, it emits blue light (peaks at 469 nm) through the oxidation of its prosthetic group coelenterazine. The emitted photons are then absorbed by GFP, leading to its emission of green light (peaks at 508 nm). This was further confirmed by characterizing the photochemical properties of purified GFP [43]. The structure of the GFP fluorophore was successfully deduced by Shimomura to be a 4-(*p*-hydroxybenzylidene)-5-imidazolinone.

In the following period of more than 10 years, however, the study of GFP seemed to be at halt, possibly partially due to difficulties involved in collecting huge amount of jellyfish for purification; 100,000 jellyfish yield only less than 1/2 g of pure GFP after months of purification [43], and overhunting gradually depleted the jellyfish population in and around Friday Harbor [44]. Finally, in 1992 a critical breakthrough came as Prasher and his coworkers cloned the gene coding for GFP [45]. The expression and subsequent imaging of recombinant GFP in both *Escherichia coli* and *Caenorhabditis elegans* soon followed [46–48]. These studies revealed the most remarkable property of GFP, which is that it does not require any exogenous cofactor or enzyme except molecular oxygen to become fluorescent, and all the information for the formation of the fluorophore is genetically encoded. Thus it can be expressed and become fluorescent in foreign organisms other than jellyfish. This significantly boosted the use of GFP as a genetically encoded marker for various live-cell applications. Further development of GFP and finding of its homologs in the following years resulted in variants with fluorescence emission profiles spanning almost the entire visible light spectrum, including blue (BFP), cyan (CFP), green (GFP), yellow (YFP), orange (OFP), and red regions (RFP),[2] essentially revolutionized modern cell biology.

3.3.2. Structure of GFP and the Fluorophore

The crystal structures of jellyfish GFP (referred to as *Aequorea victoria* GFP [avGFP] from now on) and an enhanced GFP mutant [49] (S65T, [EGFP]) were solved independently by two groups in 1996 [50,51]. Both GFPs adopt a cylindrical fold of 11-stranded β-barrels wrapped around a single central helix, with a diameter about 30 Å and a length of about 40 Å (Figure 3.1A). The fluorophore is located on the central helix, almost at the geometric center of the cylinder. The top and bottom of the barrel are capped by small, somewhat distorted

[2] In the field of fluorescent proteins, the wavelength of each color is defined slightly differently from what is defined in physics. Following the convention of the field, in this chapter the wavelength of each color is as follows: violet: ~400 nm; blue: ~480 nm; green: ~510 nm; yellow: ~550 nm; orange: ~580 nm; red: ~620 nm.

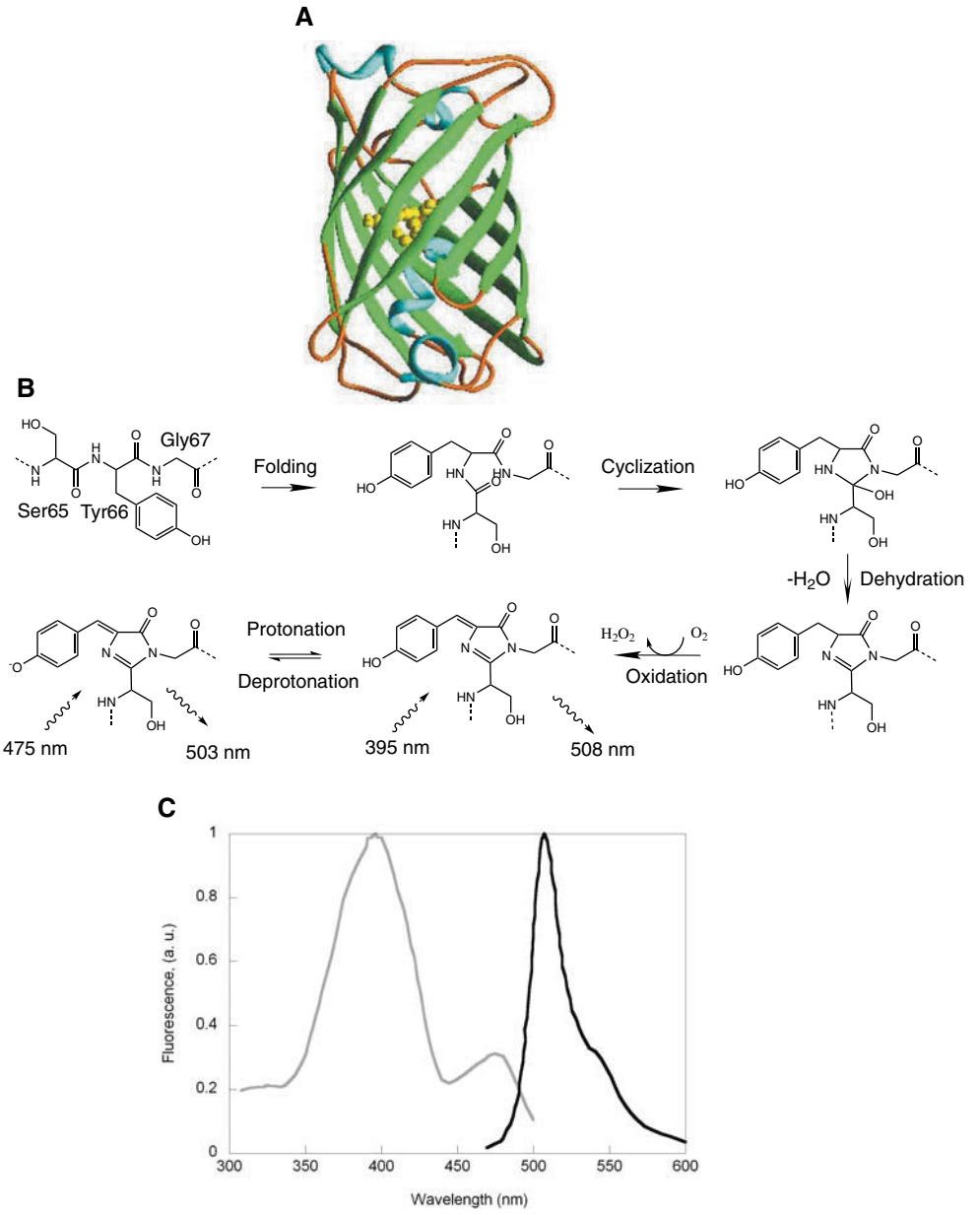


Figure 3.1. Crystal structure of green-fluorescent protein (GFP) from *Aequorea victoria* (**A**) [50], the fluorophore maturation process (**B**), and its excitation (*gray*) and emission (*black*) spectra (**C**).

sections of α-helix. The hydrogen bonds formed by the tightly fitted β-sheet strands and the caps at the top and bottom of the barrel serve well for the extraordinary stability of GFP, and at the same time shield the fluorophore from quenching by colliding water or oxygen molecules. All available FPs, natural or engineered, share the same cylindrical fold, although there may be extensive differences in the primary sequences.

The fluorophore of avGFP is formed by the posttranslational, covalent modifications of three amino acids, Ser65, Tyr66, and Gly67, which only occurs after the protein folds into its native state [47,52–56]. The fluorophore forms an extensive hydrogen-bonding network with surrounding residues, and the state of bonding changes the spectral and photophysical properties of the fluorophore substantially. The following describes the accepted mechanism of fluorophore formation, although some mechanistic details are under debate [55–57] (Figure 3.1B). The first step is the cyclization of the tripeptide main chain by the nucleophilic attack of the amide nitrogen of Gly67 at the carbonyl carbon of Ser65, generating a heterocyclic intermediate. The second step is the dehydration of the five-member heterocyclic ring, which leads to the formation of an imidazolinone ring. The oxidation of this intermediate by molecular oxygen at the position of Tyr66 C_α–C_β then finally produces a full π conjugation that extends from the phenolic group of Tyr66 to the imidazolinone ring in *cis* configuration—the mature GFP fluorophore. Among these steps, the final oxidation by molecular oxygen is rate limiting, with a time constant of about 2 hr for avGFP [53].

The absorption spectrum of avGFP has two peaks—a major one at 395 nm and a minor one at 475 nm (Figure 3.1C). Excitation at 395 and 475 nm gives similar but not identical emission spectra peaking at 508 or 503 nm, respectively [47]. Based on pH-dependent studies, it was shown that the 395-nm absorption peak is due to the protonated phenol (Tyr66) in the fluorophore, whereas the 475-nm peak is due to the deprotonated form [47,58]. Of interest, the former undergoes excited-state proton transfer (ESPT), as the neutral fluorophore becomes more acidic in the excited state [59], to become the anionic fluorophore, giving a similar but not identical emission peak at 508 nm [60,61].

In all naturally occurring GFP-like proteins (>100), four amino acids—two in the fluorophore tripeptide, Tyr66[3] and Gly67, and two that directly interact with the fluorophore through hydrogen bonding, Arg96 and Glu222 (avGFP numbering)—are absolutely conserved [64]. This suggests that the basic mechanisms of fluorophore formation in different GFP-like proteins are similar. Indeed, so far all mature fluorophores of available FPs share one common feature, which is the oxidized tyrosine C_α–C_β bond. However, the subsequent modifications of the GFP fluorophore, which change the extension of the π systems, either by natural evolution or human-made engineering, produce various FPs with colors extending from blue to far red.

3.4. Properties of Fluorescent Proteins

This section will discuss important spectral, photophysical, and biochemical properties of FPs. The goal is to equip the reader with a clear understanding of these important parameters so that when it comes to choosing which FP to use in a particular experiment, the researcher will know which property of different FPs to compare.

3.4.1. Brightness

The most essential parameter to ensure successful single-molecule imaging is the brightness of the fluorophore. At the ensemble level, brightness (B_e) is defined as the product

[3] It is unclear why Tyr66 is universally conserved in all GFP-like proteins in nature. In engineered FPs it was found that other two aromatic groups, Trp and Phe, or even His, can substitute Tyr66 and produce blue-shifted fluorescence as in CFPs and BFPs [47,62,63].

of molar absorption coefficient (ε, M^{-1} cm^{-1}) and quantum yield (Φ) at a particular pair of excitation and emission wavelengths:

$$B_e = \Phi_{\lambda ex/\lambda em} \cdot \varepsilon_{\lambda ex}$$

At the single-molecule level, the brightness is defined as

$$B_s = \Phi_{\lambda ex/\lambda em} \cdot \sigma_{\lambda ex} = \Phi_{\lambda ex/\lambda em} \cdot \frac{2.303\varepsilon_{\lambda ex}}{N_A}$$

where $\sigma_{\lambda ex}$ is the absorption cross section of a single molecule in cm^2, which can be calculated from ε using the Avogadro number, N_A.

In a typical experiment, the number of photons emitted by a single fluorophore during a particular exposure time can be calculated from its photon emission rate, which is related to its brightness, in that

$$I_{\lambda em} = I_{\lambda ex} \cdot B_s$$

where $I_{\lambda em}$ is the photon emission rate in photons per second per molecule and $I_{\lambda ex}$ is the incident light intensity in photons per second per square centimeter. For example, given an incident light flux of 514 nm at 200 W/cm^2, a typical room-temperature fluorophore absorption cross section at 2 $Å^2$, and a quantum yield of 0.6, the photon emission rate can be calculated to be $\sim 6 \times 10^4$ photons/sec. If an exposure time of 100 ms is used and the overall microscope detection efficiency is 5%, then ~300 photons will be detected during the 100 ms. This amount of photons is more than enough to enable single-molecule imaging; many detectors are capable of single-photon detection. However, the available photons become a major limitation when a reaction time course is monitored and the total photons are spread over a number of separate frames. In addition, as discussed later, a high brightness is necessary but not sufficient to achieve single-molecule detection in live cells because the photobleaching property of the fluorophore and the autofluorescence background of live cells complicate detection. In practice, an experience-based empirical threshold of brightness can be set at about a B_e of 3×10^4 (B_s of 1.0) to ensure that a single fluorophore will emit enough photons to allow its detection at the single-molecule level. Depending on the emission wavelength, this threshold for selecting fluorophores will change. For example, in the blue/green region a fluorophore that is brighter than the empirical threshold is usually needed because cells' autofluorescence background is higher at these wavelengths, whereas in the red region the threshold could be lower because the autofluorescence background is lower.

3.4.2. Fluorescence Lifetime

Fluorescence lifetime is another important parameter of a fluorophore. It measures on average how fast a fluorophore emits a photon, thus determining the saturating excitation intensity required to extract all possible photons in the shortest time. However, in live-cell imaging this scenario is rarely if ever encountered. Besides this, most FPs suitable for live-cell imaging have similar fluorescence lifetimes in the range of a few nanoseconds, rendering gating fluorescence lifetime difficult. Therefore, this parameter usually has little influence when choosing an appropriate FP.

3.4.3. Photobleaching Quantum Yield

Photobleaching is the process of irreversible destruction of a fluorophore in its excited state due to its interactions with molecular oxygen or other surrounding molecules. The fluorophore may be chemically modified and return to the ground state as a new molecule that no longer absorbs light at the excitation wavelength, For example, it was reported that photobleached enhanced yellow-fluorescent protein (EYFP) molecules irreversibly lose a mass of 44 daltons, possibly due to the decarboxylation of the E222 residue during intense laser illumination [61,65]. The exact photobleaching mechanisms of FPs in live cells have not been elucidated, and different FPs may have different mechanisms due to different structural flexibility and local fluorophore environments. For example, photobleaching of mOrange and TagRFP-T are reported to be sensitive to the presence of oxygen [66], while EGFP is insensitive to oxygen, singlet oxygen, or general radicals [67–69]. It has also been reported that when a FP molecule bleaches, it emits one singlet oxygen, which could further damage adjacent FP molecules [70,71].

Quantitatively, the sensitivity of a fluorophore to photobleaching is described by its photobleaching quantum yield Φ_b, defined as the probability of photobleaching per photon absorbed, that is, the reciprocal of the average number of excitation/emission cycles a fluorophore can undergo before it photobleaches. When multiplied by the quantum yield, the product represents the average number of photons a fluorophore can emit. For organic dyes used in single-molecule studies, Φ_b is usually in the range around 10^{-6}, while that for FPs is around 10^{-5} [68,72–74].

Photobleaching quantum yield is also a key parameter in determining whether the fluorophore is suitable for single-molecule imaging because it limits how many photons can be collected before a fluorophore is irreversibly destroyed. For example, if a fluorophore is dim but highly photostable (low Φ_b), one may compromise time resolution to accumulate enough photons using long integration time. However, if a fluorophore is photoliable (high Φ_b), it may only emit a few photons before it photobleaches, rendering its detection difficult even if it has a high brightness.

Φ_b is dependent on the particular molecular structure of a fluorophore and its local environment but not on excitation intensity or how excitation is delivered (continuous or pulsed). Therefore, Φ_b can be used to compare the photostability of fluorophores across different spectral classes and in different experimental setups. By definition, Φ_b is best measured at the single-molecule level, where the number of photons each fluorophore emits before it photobleaches is counted and the average Φ_b calculated. Note here that although the total number of photons a fluorophore can emit before it photobleaches is independent of excitation intensity, the photon emission rate of the fluorophore is. The stronger the excitation intensity, the higher is the photon emission rate, and thus the shorter is the time the fluorophore will fluoresce. If achieving longer time traces is the goal, low excitation intensity should be used at the sacrifice of lower signal levels per acquisition time.

Currently published photostability measurement of FPs were usually done at the ensemble level by monitoring the time at which the photon emission rate of high-concentration fluorophores has dropped to half of its initial rate under arc lamp illumination [75]. This method does not allow the calculation of the absolute Φ_b values but still provides a qualitative comparison among different FPs that are measured under the same experimental conditions. However, the time constants in this type of ensemble measurements are often much longer than what would be obtained in single-molecule measurements [76]. This is likely due to the fact that in the ensemble there are molecules existing in dark states that do not

fluoresce immediately at the onset of excitation [68], whereas in single-molecule experiments the true photobleaching time can be obtained by recording only the time when a fluorophore is fluorescent. In addition, the ensemble photobleaching time curves often exhibit multiphasic exponential decay, which cannot be simply described by one time constant. Because of these considerations, a systematic characterization of the photobleaching quantum yields of available fluorophores at the single-molecule level is much needed.

In most single-molecule experiments (especially tracking experiments), photostability of a fluorophore is preferred to obtain long trajectories. However, photobleaching can also be used to one's advantage. For example, in gene expression experiments in which accumulation of fluorescent background by newly expressed FP molecules is a problem, a photolabile FP such as the YFP variant Venus is preferred so that the previously generated molecules can be quickly photobleached to allow the next round of detection.

3.4.4. Fluorescence Blinking

Fluorescence blinking is the switching of a fluorophore between a fluorescent and a nonfluorescent state spontaneously on a time scale usually ranging from milliseconds to seconds. At the ensemble level, blinking of individual molecules is not observed, and this behavior merely reduces the average fluorescence intensity because the on- and off-switching of individual molecules is stochastic and averaged out in large ensembles. At the single-molecule level, however, it is important to take the blinking behaviors of individual molecules into account to avoid misinterpretation of data.

Nearly all fluorophores exhibit some kinds of dynamic blinking behaviors on different time scales. In live-cell imaging, the integration time is usually in the range of a few to a few hundreds of milliseconds. Therefore, only blinking at that time scale is considered in the following text because faster or slower blinking behaviors will not be reflected in the measurements.

The on-off blinking behavior was first reported on EGFP [69] and YFP [77]. Figure 3.2 shows one typical time trace of a single YFP (Venus) molecule on continuous illumination at 514 nm. It is evident that the fluorescence intensity fluctuates between bright (on) and dark (off) states on the time scale of seconds. A three-level system has been proposed to explain the blinking behavior. The on-state is the result of photon emission during the transitions from the singlet excited state to the ground state, as described in a standard fluorescence two-level system. During this cycle, however, there is a small probability that a molecule will go into a long-lived dark state that cannot emit a photon. This contributes to the off-period. Only when the molecule is returned to the singlet ground state capable of absorbing new photons will fluorescence resume. The transition probability from a bright to dark state can be calculated by taking the reciprocal of the average number of excitation/emission cycles during the on-time, which has been measured to be at about 5×10^6 for enhanced GFP (EGFP) [74]. The average off-time for both EGFP and EYFP was measured to be between a few and a few tens of seconds [65,68,74,77].

The exact nature of the molecule at the dark state, however, has not been completely elucidated. It was first proposed that the protonated fluorophore is responsible for the off-state and that the conversion between protonation and deprotonation of the fluorophore results in the blinking behavior—the anionic form of the fluorophore fluoresces, while the protonated form does not [77]. Later this process turned out to be more complicated than expected. McAnaney et al. demonstrated that at sufficiently low power densities (<1 kW/cm^2), the ground-state protonation and deprotonation is indeed responsible for the blinking behavior

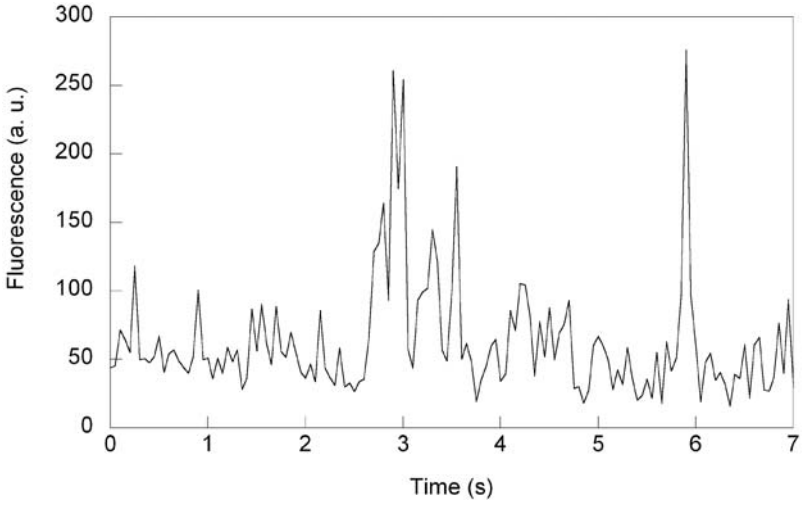


Figure 3.2. Fluctuations (blinking) in the fluorescence intensity of a single yellow-fluorescent protein (YFP) Venus molecule under continuous illumination of 514 nm at 2 kW/cm^2 with an integration time of 50 ms.

[65]. However, Peterman et al. showed that there is a clear excitation power dependence—between 0.5 and 5 kW/cm^2 the fluorescence on-time becomes shorter when the power gets higher. In addition, at high excitation intensity (5 kW/cm^2), the average fluorescence on-time is independent of pH between 6 and 10 [68]. These experiments suggest that at high power densities, the photon-driven excited-state protonation reaction of the fluorophore is faster than the ground-state process [65,74,78]. The fluorescence off-time, however, was found to be independent of excitation power [68], indicating that YFP molecules in the off-state spontaneously relax to the on-state [65,79]. Moreover, photon-induced isomerization of the fluorophore conformation [80] and rearrangement of hydrogen-bonding network around the fluorophore [81] have been suggested to contribute to the blinking behavior.

3.4.5. Maturation Time

The fluorophore maturation process in FPs determines how fast a fluorescent response can be detected following the production of a FP molecule. The oxidation step of the fluorophore formation process as described earlier is usually the slowest and is thus considered the rate-limiting step for the maturation process. To speed up the maturation process, increased access of the fluorophore to molecular oxygen is important [82]. However, it is possible that this increased accessibility may also lead to increased environmental sensitivity and proneness to photobleaching. For some cellular processes that do not require the detection of a FP molecule immediately after its production, slow maturation is of less concern. However, a cell expressing a fast-maturing FP will contain more copies of fully matured FP than a cell expressing a slow-maturing FP at the steady state, affecting data throughput of an experiment. If monitoring the fast dynamics of cellular processes such as gene expression is the goal, using a fast-maturing FP as the gene expression reporter is critical. Because gene regulation is mainly controlled by differential binding modes of transcription factors, changes of gene expression state usually occur on the time scale of a few minutes or even shorter. To follow this process faithfully with a high time resolution, a FP reporter molecule needs to become

fluorescent as soon as it is synthesized. Unfortunately, all current FPs mature slowly, on the time scale of a few tens of minutes to a few hours, except one YFP variant, Venus, which has a maturation time constant measured at about 2 min *in vitro* and 7 min *in vivo* [83,84]. This property makes Venus an attractive gene expression reporter, allowing the direct counting of protein molecules generated one at a time [84]. If another fast-maturing FP emitting in the red region were developed, the temporal expression correlation of two related genes could be monitored simultaneously at the single-molecule level, greatly facilitating the investigation of gene regulation mechanisms in large networks.

3.4.6. Construction and Expression of Fusion Proteins

It is critical that being tagged by a FP does not perturb the function of the tagged protein. Before a single-molecule experiment is conducted, one should verify whether the fusion protein is functional. This can be done in different ways. The most stringent test is to check whether the expression of the fusion protein in a cell strain that has the gene for the untagged protein deleted will rescue the deletion phenotype, provided that the deletion of the wild-type gene has an identifiable phenotype. If no obvious phenotype of the deletion strain can be identified, one should verify whether other genetic and biochemical properties of the fusion protein, such as expression level, cellular localization pattern, enzymatic activity, or gene regulation activity, reproduce those of the untagged protein. Based on empirical observations, in general GFPs cause least perturbations to the tagged proteins, while some engineered RFPs and especially some photoinducible FPs such as rsFastLime and EosFP (discussed later) are difficult to work with. In the latter cases, when the fusion protein is not functional, the linker sequence needs to be carefully scrutinized, and changing the terminus to which the FP is fused may improve the performance of the fusion protein. It was found that by fusing the first and last seven amino acid residues of avGFP to the N and C termini of some RFPs, the tolerance of the tagged proteins to the RFP fusions was enhanced [85]. In some cases, lowering the expression temperatures helped to improve the folding of the fusion proteins.

To enable single-FP-molecule detection in live cells, it is necessary to repress the expression level of the fusion FP protein to avoid large fluorescence background. In bacterial cells, this can be done by using a weak or repressed promoter such as the *lac* promoter and incorporating the gene encoding the fusion protein into chromosome or a single-copy plasmid such as the F plasmid. In eukaryotic cells, a regulated promoter such as the cytomegalovirus (CMV) promoter repressed by Tet repressors can be used. The original chromosomal gene of the tagged protein can be left intact. In the case of photoinducible proteins (see later discussion), however, the expression level does not have to be strongly repressed because high concentrations of FP can be expressed with only a few molecules becoming fluorescent upon photoinduction.

3.4.7. General Guidelines

In general, a FP suitable for single-molecule imaging in live cells should possess one or more of the following properties: high brightness, high photostability, emission in the visible range, and fast maturation time. Note that although the monomeric state of the FPs was not discussed earlier, it is mandatory for protein tagging. This is because dimeric or tetrameric FPs tend to cause self-association and aggregation of tagged proteins, preventing proper interpretation of experimental data [75]. All the FPs to be discussed are monomeric. One should also keep it in mind that there is no single FP that can meet the demanding criteria of all

experiments. Therefore, one needs to decide what the most important property for a particular experiment is and make the necessary compromise. In the following section a collection of FPs that have the potential or have already been demonstrated in single-molecule imaging in live cells is discussed in detail. Their properties are compiled in Tables 3.2 and 3.3.

3.5. Derivatives of avGFP and Other GFP-Like Proteins

Extensive mutagenesis efforts have been exerted on the original avGFP to improve its properties. This is because the avGFP has complicated photochemistry [86,87], does not fold well above room temperature [58], and matures slowly [47]. Through the combination of random mutagenesis and rational design, different versions of FPs have been generated, with their absorption and emission wavelengths blue- or red-shifted as much as 40 nm and with improvements in their fluorescence brightness, photostability, environmental sensitivity, folding, and maturation kinetics (Figure 3.3 and Table 3.2) (for extensive reviews, see Refs. 75 and 86]).

For example, the first EGFP (Figure 3.3) has an S65T mutation, which diminishes the 395-nm absorption peak of the avGFP but significantly enhances the 475-nm peak, greatly facilitating excitation in the blue region instead of the violet region, with the latter being usually detrimental to cells [47,49,88]. This is possibly because the threonine at position 65 is bulkier than the original serine and cannot adopt the same conformation as Ser65 to donate a hydrogen bond to Glu222. This change leads to changes in the hydrogen-bonding status of a few polar residues around the fluorophore, and the phenolic hydroxyl group of Tyr67 is effectively anionized, contributing to the greatly enhanced absorption at 475 nm [51,89]. Another example is the YFP class. The red-shifted excitation and emission (515 and 525 nm, respectively) of YFPs is mainly due to the T203Y mutation, where the tyrosine stacks its aromatic ring next to the phenolate anion of the fluorophore, forming an extended $\pi-\pi$ interaction and lowering the excited-state energy level (Figure 3.3) [63,90].

Among many improved versions of avGFP, the YFP class is preferable over GFPs for single-molecule imaging in live cells due to its superior spectral properties [76], the reduced cellular autofluorescence at YFP emission peaks, and the available high-quality imaging optics and detectors. Blue or cyan FPs are usually not considered for single-molecule imaging in live cells because of the presence of prominent cellular autofluorescence in their spectral region, increased phototoxicity to cells caused by ultraviolet (UV)/violet excitation, and low detection efficiency of UV optics and detectors. For the GFP class, although improved GFPs such as EGFP has been successfully applied to single-molecule imaging in live cells to reveal cell surface protein distribution patterns [91] and actin filament dynamics [92], and the newly developed superfolder GFP (sfGFP) [93] possesses the necessary brightness for its single-molecule detection, this class does not have a clear advantage over the YFP class or the RFP class in single-molecule live-cell imaging. Therefore, blue and green FPs are not included here.

3.5.1. Derivatives of avGFP

EYFP

Enhanced YFP (EYFP) contains four mutations (S65G, V68L, S72A, and T203Y) compared to avGFP. As discussed earlier, the T203Y mutation is rationally designed to shift

Table 3.2. Properties of selected fluorescent proteins that have potential or have been demonstrated in single-molecule imaging in live cells

	Ex (nm)	Em (nm)	ε (M^{-1} cm^{-1})	σ ($Å^2$)[1]	Φ	B_s[2]	Φ_{bl}[3a] ($\times 10^{-5}$)	Photostability[3b] (sec)	$\tau_{maturation}$ (min)[4]	pKa[5]	SMD[6] *in vitro*	SMD[7] *in vivo*
EGFP[8]	488	507	53,000 [214]	2	0.60 [214]	1.2	6.9 [76]	174	94 [82]	6.0	[76,215,68,69]	[91,92]
sfGFP	485	510	83,300 [93]	3.2	0.65 [93]	2.1	ND	157	ND	5.5	NP	NP
EYFP[9]	514	527	83,400 [63]	3.2	0.61 [63]	1.9	6 [76]	60	6.6 [83]	6.9 [83]	[65,68,76–78]	[33,94–96,202]
Citrine	516	529	77,000 [100]	2.9	0.76 [100]	2.2	2.6 [73]	49	ND	5.7 [100]	[73]	[216]
Venus	515	528	92,200[10] [83]	3.5	0.57 [83]	2.0	ND	15	2.0 [83]	6.0 [83]	NP	[84,196]
Ypet	517	530	104,000 [103]	4.0	0.77 [103]	3.1	ND	49	ND	5.6 [103]	NP	NP
mKO	548	559	51,600 [117]	2	0.60 [117]	1.2	ND	122	ND	5.0	NP	NP
mOrange2	549	565	58,000 [66]	2.2	0.60 [66]	1.3	ND	228 [66]	390 [66]	6.5	NP	NP
tdTomato	554	581	138,000 [85]	5.3	0.69 [85]	3.6	ND	98	87 [85]	4.7	NP	[170]
TagRFP-T	555	584	81,000 [66]	3.1	0.41 [66]	1.2	ND	337	144	4.6	NP	NP
mCherry	587	610	72,000 [85]	2.8	0.22 [85]	0.6	ND	96	22 [85]	<4.5	[218]	[116]
mKate	588	635	45,000 [122]	1.7	0.33 [122]	0.6	ND	166	108 [122]	6.0 [122]	NP	NP

[1] The cross section σ of single molecules is calculated from the ensemble molar extinction coefficient ε and expressed in $Å^2$.
[2] Brightness at the single-molecule level expressed as the product of cross section σ and quantum yield Φ, in units of $Å^2$.
[3a] Photobleaching quantum efficiency measured at the single-molecule level. For most fluorescent proteins (FPs) this parameter is not determined.
[3b] Photostability was measured at the ensemble level by the time (in seconds) it takes for the photon emitting of each FP to drop to half of its initial value at 1,000 photons/sec under continuous illumination of an arc lamp. Values of photostability are taken from Ref. 75 if not specifically indicated.
[4] Time constant of the oxidation step during the maturation of the chromophore. In some published work the maturation half-life $\tau_{1/2}$ was used. In this table published $\tau_{1/2}$ is converted to $\tau_{maturation}$ using $\tau_{maturation} = \tau_{1/2}/\ln(2)$.
[5] The values of pKa are taken from Ref. 75 if not specifically indicated.
[6] Representative published work involving single-molecule detection *in vitro*.
[7] Representative published work involving single-molecule detection *in vivo*.
[8] The commonly used enhanced green-fluorescent protein (EGFP) from Clontech is the GFPmut1 variant containing F64L and S65T mutations [88]. The listed ε (23,000 M^{-1} cm^{-1}) and Φ (0.7) from Clontech at http://www.clontech.com/upload/images/downloads/fp_overview_and_apps_CDv2.zip are lower than those listed in the table and have not been verified.
[9] Enhanced yellow-fluorescent protein (EYFP) contains S65G, V68L, S72A, and T203Y mutations and was called 10C in the original publication [63].
[10] In two other studies [103,217] the molar extinction coefficient and quantum efficiency of Venus were measured at ~105,000 and ~0.7, respectively. The differences may be due to the different percentages of nonfluorescent protein in the protein preparation.
Em, emission; ex, excitation; mKO, monomeric orange-fluorescent protein Kusabira-Orange; ND, not determined; NP, not published; sfGFP, superfolder green-fluorescent protein; YPet, yellow-fluorescent protein for energy transfer.

Table 3.3. Properties of photoinducible fluorescent proteins

	Pre-photoinduction				Photoinduction wavelength/Φ[1]	Reversibility	Post-photoinduction				pKa	Ref.	SMD *in vitro*	SMD *in vivo*
	Ex (nm)	Em (nm)	σ (Å^2)	Φ			Ex (nm)	Em (nm)	σ (Å^2)	Φ				
Dronpa	NA	NA	NA	NA	405 nm/0.37	Yes (488 nm/3.2×10^{-4})	503	518	3.6	0.85	5	[127]	[138,145,158]	NP
rsFastLime	NA	NA	NA	NA	405 nm/0.37 or none	Yes (488 nm/0.17)[2]	496	518	1.5	0.77	ND	[140]	[152]	NP
rsCherry	NA	NA	NA	NA	550 nm/ND	Yes (450 nm/ND)	572	610	ND	ND	ND	[153]	[153]	NP
rsCherryRev	NA	NA	NA	NA	450 nm/ND	Yes (550 nm/ND)	572	608	ND	ND	ND	[153]	[153]	[153]
Dendra2	490	507	1.7	0.50	405 or 488 nm/ND	No	553	573	1.3	0.55	6.6–6.9	Evrogen [131]	NP	[159]
mEosFP (tdEosFP)	505	516	2.6	0.64	405/ND	No	569	581	1.4	0.62	5.8	[130]	[158,130,150]	[123]

[1] Φ is the photoinduction quantum efficiency. In the case of rsFastlime, it is assumed to be the same as that of Dronpa because the measured switching-on rates for Dronpa and rsFastlime are similar to each other [140]. rsFastlime can also spontaneously relax to the fluorescent state without the need for an activation laser line.

[2] The switching-off quantum efficiency of rsFastlime is calculated based on the switching-off rate ratio between rsFastlime and Dronpa [140].

[3] The quantum efficiency and absorption cross section of rsCherry and rsCherryRev have not been measured, but fluorescence correlation spectroscopy study shows that at the single-molecule level they have similar brightness to mCherry [153].

Em, emission; Ex, excitation; NA, not applicable; ND, not determined; NP, not published.

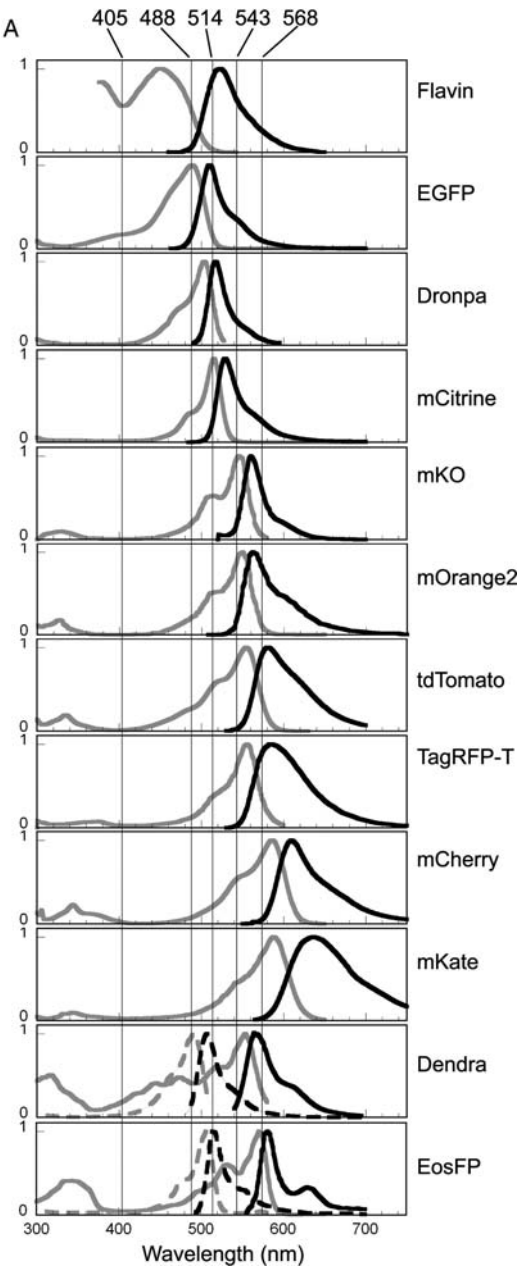


Figure 3.3. Comparison of fluorescence excitation (*gray*), emission (*black*) spectra (**A**), and the corresponding fluorophore structure of fluorescent proteins (FPs) as discussed in the text (**B**). The excitation and emission spectra of flavin adenine dinucleotide (FAD) are also shown as a comparison. For the photoconvertible FPs Dendra and EosFP, the spectra of green forms are shown in broken lines and red forms in solid lines. For each fluorophore the conjugated π system is color-coded for its emission wavelength. For the yellow-fluorescent protein (YFP) class, only mCitrine's spectra and fluorophore structure are shown because they are representative for other YFPs such as enhanced yellow-fluorescent protein (EYFP), Venus, and YFP for energy transfer (YPet). The spectra of rsFastlime, rsCherry, and rsCherry-Rev are not shown because they are similar to their corresponding parents Dronpa and mCherry with slight wavelength shift, respectively (see Table 3.3 for details). The crystal structures of fluorophores in some FPs are not

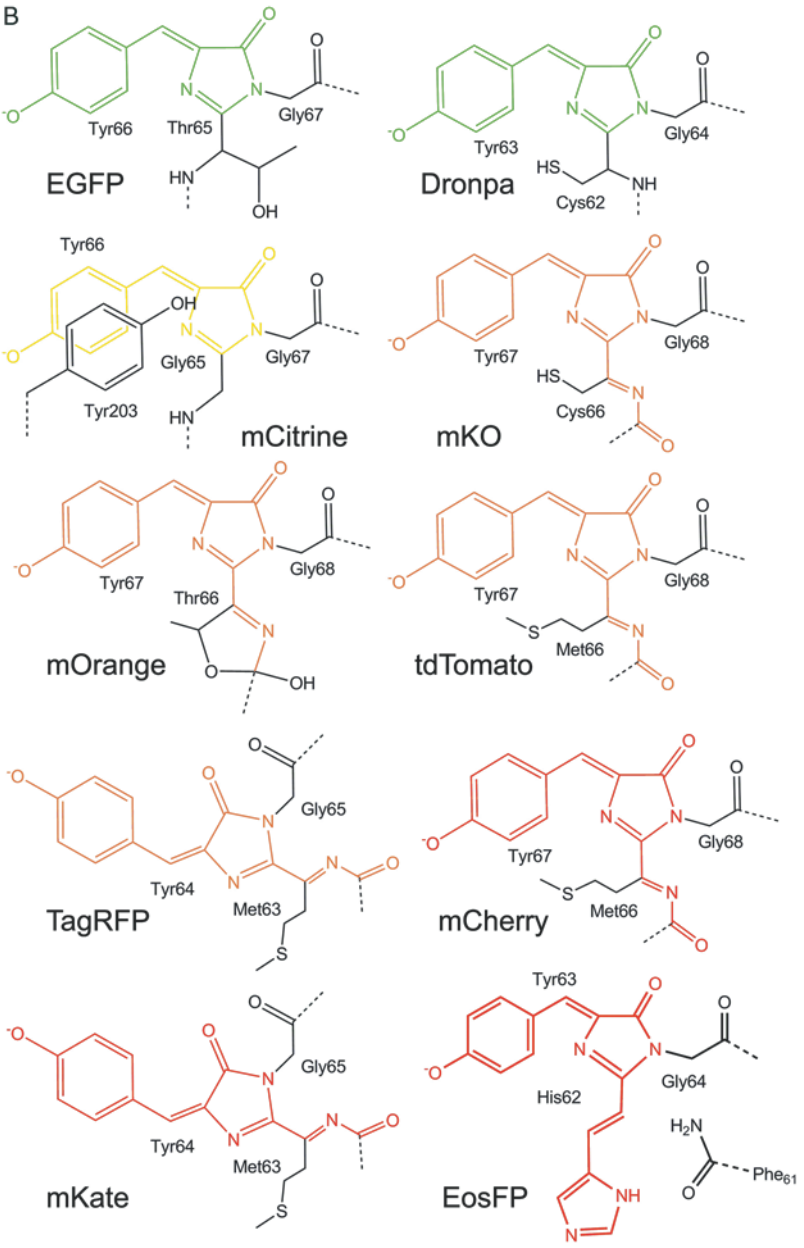


Figure 3.3. (continued) available, and the shown structures are modeled based on the closest available structures when possible: monomeric orange-fluorescent protein Kusabira-Orange (mKO) and tdTomato from *Discosoma striata* red-fluorescent protein (dsRed) [112,113], mOrange2 from mOrange [118], and TagRFP-T and mKate from eqFP611 [121]. For Dendra the mechanism for its photoconversion is unclear, and therefore its chromophore structure is not shown. Five commonly used laser lines at 405 (solid state), 488 (argon ion), 514 (argon ion), 543 (helium neon), and 568 nm (krypton) are superimposed on the spectra to indicate which laser line is best suited for the excitation of which FP. EGFP, enhanced green-fluorescent protein.

emission into the yellow-green region of the spectrum. The commercially available version of EYFP from Clontech has also been codon-optimized to allow more efficient translation and higher expression levels in eukaryotic cells. EYFP is bright, with a high absorption coefficient of 83,400 M^{-1} cm^{-1} and a quantum yield of 0.61 [63]. It has been widely used to track the movement and distribution patterns of single proteins molecules on the membrane of live cells [76,94–96]. However, its high pKa (6.9) and sensitivity to halides cause its fluorescence to fluctuate in a cellular environment, where usually pH is 6.8–7.3 and Cl^- concentration is 10–60 mM [97,98]. In addition, EYFP is a weak dimer, with a K_d of 0.11 mM [99], although this should not be of significant concern for single-molecule studies. This is because most of the time, to achieve single-molecule detection in live cells, FPs are usually expressed at concentrations far below those required for significant dimer formation. A true monomeric YFP (mYFP) was generated by introducing the mutation A206K, in which the positive charge of the lysine side chain disrupts the hydrophobic interaction at the dimeric interface. The spectral properties of mYFP are essentially the same as those of EYFP.

Citrine

By introducing one additional mutation, Q69M, into EYFP, a new YFP named Citrine was generated, Citrine has a lower pKa at 5.7, is less sensitive to halide, and is about twice as resistant to photobleaching as EYFP [73,100]. The low pKa and insensitivity to halide could be explained by the Q69M mutation. The side chain of the Met was shown in the crystal structure to occupy a halide-binding cavity near the fluorophore, shielding the fluorophore from Cl^- or protons. The resistance to photobleaching may also be a direct result of the elimination of the cavity near the fluorophore, so that the fluorophore is better packed and shielded from the environment. Like its parent EYFP, Citrine is also a weak dimer, but its monomeric version was easily generated by introducing the same mutation A206K, and the resulting mCitrine has the same spectroscopic properties [73,75,100,101].

Venus

A particularly notable and popular YFP mutant is the fast-maturing Venus generated by the Miyawaki group [83]. Venus has comparable brightness, low pKa, and insensitivity to halide to those of mCitrine, but its most distinguishing feature is its fast fluorophore maturation process, the time constant of which is measured to be about 2 min *in vitro* [83] and 7 min inside live *E. coli* cells [84]. The accelerated fluorophore maturation results from a combined effect of the five point mutations (F46L, F64L, M153T, V163A, S175G), which likely introduce structural flexibility, remove steric and energetic constraints in the folding of the polypeptide chain, and enhances oxygen access to the fluorophore [83,102]. Given that many other FPs mature on the time scale of a few tens of minutes or hours, the accelerated maturation time of Venus is particularly attractive for monitoring cellular processes with fast dynamics such as gene expression. In a recent work by the Xie group, Venus was used as a single-molecule gene expression reporter to probe the real-time dynamics of a repressed *lac* promoter in live *E. coli* cells [84]. Had a FP with slower maturation kinetics been used in this work, the detected FP molecules generated from one single mRNA would have a much wider distribution in time, rendering it difficult to discern when an mRNA molecule was generated. One drawback of Venus is its low photostability—it photobleaches at least twice as fast as mCitrine. Therefore, Venus is not the best fluorophore to use in experiments such as

single-molecule tracking where photostability is crucial. Venus is also a weak dimer, but, just as in the cases of mYFP and mCitrine, an A206K mutation generates monomeric mVenus [101].

Yellow-Fluorescent Protein for Energy Transfer

Yellow-fluorescent protein for energy transfer (YPet) was originally generated during a screening for improved fluorescence resonance energy transfer (FRET) between CFP and YFP [103], but it was later found that the improved FRET performance of YPet is likely due to its tendency to dimerize with its coevolved partner CyPet [104,105]. Nevertheless, YPet is superior for single-molecule imaging because it is the brightest YFP available, with a very high absorption coefficient at 104,000 M^{-1} cm^{-1} and a quantum yield of 0.77. It was based on another bright YFP variant, Topaz [63], which was the brightest YFP before the appearance of YPet. Multiple mutations were introduced into Topaz through random mutagenesis, and the roles of many of them in improving its brightness are unclear. The maturation kinetics of YPet is similar to that of EYFP, but its photostability has not been addressed. There is no report on the use of YPet in single-molecule studies, but given its brightness and low pKa (5.6), its use in single-molecule imaging in live cells should be imminent.

3.5.2. Orange- and Red-Fluorescent Proteins

Fluorescent proteins that excite and emit at longer wavelengths than GFP or YFP are highly desirable for single-molecule imaging in live cells for three reasons. First, excitation at long wavelengths avoids exciting the flavin protein, which is the major source for high cellular autofluorescence background in the 500-to 600-nm region. Second, orange or red excitation in general causes less photodamage to cells because blue or green excitation carries more energy and can be absorbed by DNAs and proteins. Therefore, longer time traces can be obtained without exposing cells to severe photodamage. Third, RFPs can be spectrally distinguished from GFPs or YFPs, making it possible for two-color, single-molecule imaging in live cells.

Earlier efforts to engineer jellyfish avGFP to produce variants in the orange or red region had been unsuccessful,[4] although blue light–induced conversion from green to red fluorescence under limited oxygen supply had been reported [107,108]. This situation did not change until 1999, when the first RFP, dsRed, was discovered in a nonbioluminescent reef coral species, *Discosoma striata*, commonly called the striped mushroom [109]. This breakthrough started the second round of extensive mutagenesis efforts to improve the original dsRed because it matures slowly ($t_{1/2} = 27$ hr) through a green intermediate [110], has dim fluorescence [109], and is an obligate tetramer [110]. The results were successful, and a series of monomeric RFPs called mFruits were generated by the Tsien group [85]. Together with RFPs generated from other research groups and companies, there are now FPs ranging from orange to far infrared with much improved biochemical and photochemical properties (Figure 3.3) (also see comprehensive reviews in Refs. 75 and 101).

The formation of these RFP fluorophores (Figure 3.4) shares similar steps with GFP maturation but includes additional chemistry, which explains why the efforts to engineer

[4] There is one recent study reporting the generation of the first red fluorescence-emitting derivative (excitation and emission maxima at 555 and 585 nm, respectively) of avGFP [106]. Purified proteins or cells expressing this mutant, however, showed both green (strong) and red (dim) emission. It is likely that only a small population of the mutant matures to the red chromophore. Nonetheless, this study indicates that the full mutagenesis potential of GFP has yet to be reached.

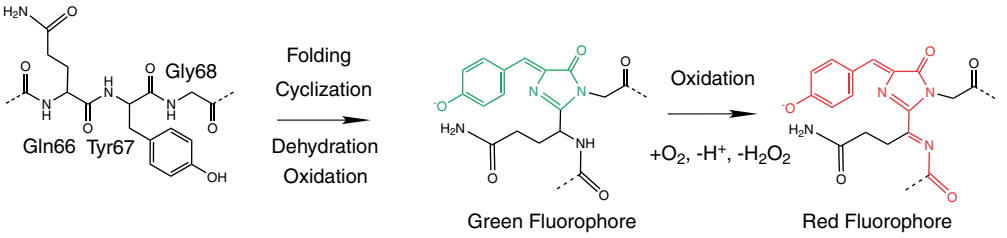


Figure 3.4. Maturation of the fluorophore in *Discosoma striata* red-fluorescent protein (dsRed) [112,113]. The first step is to form a green-emitting fluorophore as in green-fluorescent proteins through cyclization, dehydration, and oxidation, and the second step involves oxidation of the bond between the α-carbon and the nitrogen of Gln66 to form an acylimine, which extends the π-conjugation system.

avGFP to produce variants in the red region have been unsuccessful [111–113]. For dsRed, the initial steps also involve a tripeptide (Gln66, Tyr67, and Gly68) to go through similar cyclization, dehydration, and oxidation to form a coplanar *p*-hydroxybenzylidene-imidazolinone π system like that in GFP. However, additional oxidation at C_α—N bond of Gln66 takes place to form an acylimine group, leading to an extended π-bonding resonance system. This extended conjugation system results in greater delocalization of electrons upon excitation, giving rise to the longer excitation and emission wavelengths (558 and 583 nm, respectively). In addition to this extra oxidation step, the Phe65 and Gln66 of dsRed are connected by an unexpected and unique *cis* peptide bond, the isomerization of which from the usual *trans* bond formed initially upon protein folding is believed to be a key step that slows down the maturation of the fluorophore [112,114]. In other RFPs that derived based on dsRed, such as mOrange, mCherry, and mStrawberry, the fluorophores may undergo further chemical transformations [115]. In the following, a few OFPs and RFPs with promising potential for single-molecule imaging in live cells are discussed.

The brightness of OFPs and RFPs is usually less than that of the YFPs; however, as the cells' autofluorescence background is much less at the longer wavelengths, the slightly decreased brightness does not pose a significant problem for single-molecule imaging. In fact, mCherry, the dimmest among the few selected RFPs to be discussed, has been demonstrated to be detectable as single molecules in live bacteria cells [116]. Since many of the OFPs and RFPs to be described were just discovered very recently, some of their properties, such as fluorophore structures, maturation kinetics, photobleaching, and fluorescence-blinking behaviors, have not been fully characterized. Therefore, the following selection of FPs is based on the best knowledge available and serves as a starting point for further detailed investigations, especially at the single-molecule level.

Monomeric Orange-Fluorescent Protein Kusabira-Orange

The orange-fluorescent protein Kusabira-Orange (KO) was originally cloned from stony coral and selected as a FRET partner with CFP [117]. Because KO exists in the heterogeneous oligomeric state including dimers and tetramers, extra rounds of mutagenesis were carried out to introduce more than 20 mutations to generate the monomeric version mKO. mKO absorbs strongly at 548 nm, with a small shoulder at 515 nm, and emits maximally at 561 nm. Its overall brightness is comparable to that of EGFP, and its photostability appears to be better than that of many FPs, making it an attractive fluorescent label especially in single-molecule

tracking experiments [66,75]. mKO behaves well when fused to most proteins tested, although it was reported to cause problems in the localization of α-tubulin. Adding longer/flexible linkers or the N- or C-terminus of GFP to mKO may lessen this problem [66].

mOrange and mOrange2

mOrange (excites and emits maximally at 548 and 562 nm, respectively) was created during one substantial effort of reengineering dsRed and its derivatives to generate a series of orange- and red-emitting FPs called mFruits. Each FP in this series is named after a similarly colored fruit [85]. mOrange is the brightest monomeric FP in this series, but it photobleaches quickly. In addition, its pKa of 6.5 may pose a disadvantage compared to other mFruits, which usually have pKa values around 5 or lower. The absorption spectrum of mOrange has a minor peak at 470 nm, and it has a 510-nm emission peak under 470-nm excitation. This likely represents a population of fluorophores whose maturation stalls at a GFP-like fluorophore stage [118]. This can be a serious drawback in using mOrange in two-color imaging experiments. Work aimed at improving the photostability of FPs led to a new version, mOrange2, with four mutations that may hinder a critical oxidation reaction and loss of fluorescence of the fluorophore [66]. mOrange2 is less bright than mOrange but still compares favorably to other mFruits. The absorbance spectrum of mOrange2 has the similar additional peak around 490 nm, but excitation at this peak does not produce fluorescence. The most distinctive property of mOrange2 is its photostability, which is significantly (>25-fold) improved over that of mOrange and is among the best of all available FPs. However, its sensitivity to pH is not changed, and mOrange2 appears to mature twice as slowly as mOrange (4.5 instead of 2.5 hr).

tdTomato

The tandem dimer tdTomato is the brightest of all available FPs; it was created by fusing two copies of the dimeric dTomato gene, and thus possesses two bright fluorophores. Its excitation and emission wavelengths (554 and 581 nm, respectively) are red-shifted compared to those of mKO and mOrange, and it is also very photostable. The major drawback of its use in tagging proteins of interest is its relatively large size, which has been problematic when it was fused to α-tubulin and connexin, where tight packing of these fusion proteins is needed [66]. However, this should be case dependent, in that its fusion with myosin light chain appears to be successful [119]. Therefore, tdTomato fusion proteins need to be more carefully scrutinized for their functionality before any meaningful conclusion is drawn.

TagRFP and TagRFP-T

TagRFP-T is the optimized, monomeric version of the early tetrameric eqFP578 cloned from sea anemone [120]. TagRFP possesses a high absorption coefficient (100,000 M^{-1} cm^{-1}) and an intermediate quantum yield (0.48), with excitation and emission maxima at 555 and 584 nm, respectively. Of interest, the fluorophore of TagRFP, which is formed by the Met-Tyr-Gly triad, likely adopts a coplanar but *trans* conformation as in its homolog eqFP611 cloned from the same species [121]. This is different from the fluorophore conformations of the majority of other FPs, which are usually in the *cis* form. In addition, as will be discussed later for photoinducible FPs (PI-FPs), a mechanism of *trans*-to-*cis*

isomerization of the fluorophore is suggested to account for the photoinduction of the PI-FPs from the nonfluorescent to a fluorescent form. It is not clear why the *trans*-fluorophore of eFP611 is still highly fluorescent and how its conformation contributes to its photophysical properties.

Although TagRFP already possesses reasonable photostability, the Tsien group further optimized TagRFP by selecting mutants that exhibit prolonged fluorescence after intense photobleaching [66]. One mutation, S158T, was found to enhance the TagRFP's photostability ninefold. This new version, named TagRFP-T, is the most photostable FP across all spectral classes available, although its brightness is 20% lower than that of TagRFP. In addition, the absorption spectrum of TagRFP-T does not have other minor peaks that absorb but do not fluoresce, indicating that the fluorophore only exists in one fluorescent state.

mKate

The RFP mKate is also derived from eqFP587, the same progenitor as that of TagRFP [122]. In the red and far-red region where most RFPs are dim, mKate performs reasonably well, with an extinction coefficient of 45,000 M^{-1} cm^{-1}, a quantum yield of 0.33 (excitation and emission maxima at 588 and 635 nm, respectively), and a brightness that is comparable to that of mCherry. mKate also demonstrates high photostability, which is in the same range as that of the two newly developed OFP variants mOrange2 and TagFP-T. However, its photobleaching curve measured in ensemble exhibits complex behaviors, with an initial rise of fluorescence intensity followed by gradual decay, which was suggested to resemble that of photoinducible FPs. It is likely that some mKate molecules that originally exist in a dark state can be irreversibly driven into the bright state by intense laser irradiation at its absorption maximum, and continuous irradiation eventually leads to complete photobleaching. Further investigations, especially at the single-molecule level, are needed to characterize this photobleaching behavior.

mCherry

Another RFP, mCherry, has been one of the most promising RFPs in the mFruits series, in that it has exceptional photostability, its excitation and emission wavelengths are red-shifted further to around 600 nm, and it matures relatively rapidly [85]. Although its brightness is among the dimmest of the FPs selected in this chapter, its single-molecule detection has been realized in live bacterial cells [116]. Compared to mCherry, the newly developed mKate emits in the same far-red region, has similar brightness, and has better photostability. However, because the complex photobleaching behavior of mKate is not well characterized, mCherry remains a good fluorescent label to use in single-molecule imaging in live cells.

3.5.3. Photoinducible Fluorescent Proteins (PI-FPs)

Recent advances in FP engineering have resulted in another exciting class: the photoinducible FPs (PI-FPs).[5] This new class of FPs (Table 3.3) possesses a unique property in

[5] Here the term "photoinducible FPs" is introduced to characterize the common feature of this class of fluorescent proteins: They all undergo light-induced processes to change their fluorescence emission properties. In the literature, the terms photoactivation, photoswitching, and photoconversion are used, sometimes interchangeably. In the context of this chapter *photoactivation* refers specifically to the process of fluorescence turn-on from a dark state

that they undergo pronounced light-induced spectral changes in response to irradiation at a particular wavelength. This remarkable property enables their fluorescence to be controlled at will and is especially suitable for selectively marking macromolecules of interest in space and time and tracking their movement and redistribution inside a cell. For example, a protein can be labeled with a PI-FP, expressed at high density (10^5 molecules/μm^2) inside a cell. This high expression level still contributes minimally to the background, in that most of the PI-FP molecules are kept at the fluorescence off-state. Upon the application of photoinduction at a certain level, a few PI-FP molecules are selectively induced to become fluorescent, enabling single-molecule detection, localization, and tracking. These molecules can then be turned off by photoinduction or bleaching to allow the next round of recording. Thus, thousands of cycles can be carried out in the same cell to generate a large amount of data, facilitating statistical data analysis [123,124].

Available PI-FPs can be divided into two categories. The first category includes PI-FPs whose excitation/emission spectra are constant but whose fluorescence on- or off-state is changed by photoinduction. This category includes the photoactivatable RFP PA-mRFP1-1 [125], the photoswitchable RFP asFP595 [126], and the GFP Dronpa [127]. The second category includes PI-FPs whose excitation/emission color is changed by photoinduction. Typical examples are the irreversible, green-to-red photoconvertible FPs Kaede [128], KiKGR [129], EosFP [130], and Dendra [131] and the cyan-to-green PS-CFP [132]. Comprehensive reviews on the PI-FPs are available [101,133–135], and readers are directed there for the complete coverage of available PI-FPs. The focus here is on the PI-FPs that have the potential for or have already been demonstrated in single-molecule imaging in live cells. These PI-FPs include the photoswitchable Dronpa together with its derivative, the fast-switching rsFastlime; the newly developed rsCherry and rsCherryRev; the green-to-red photoconvertible monomeric mEosFP; and the improved version of Dendra, Dendra2. All of these PI-FPs are monomeric and possess sufficient brightness to be imaged at the single-molecule level. Their spectral properties are summarized in Table 3.3.

Many members in the GFP, YFP, and RFP classes also exhibit similar photochromatic behavior, although the photoinduction quantum yields are usually much lower than that of PI-FPs [66,68,77,79, 80,136,137]. Some molecules that appeared to be "photobleached" are actually molecules that stayed in long-lived dark states. These molecules can be photoinduced to fluoresce again with illumination with violet or blue light. The mechanisms of their switching in general may be similar to that of the PI-FPs, as discussed later.

Dronpa and Its Derivative rsFastLime

Dronpa and its derivatives are the best-studied PI-FPs [127,138]. Their fluorescence can be detected with blue excitation (488 nm), but prolonged blue excitation will drive Dronpa into a dark state (switch-off). Dronpa can escape from the dark state by irradiation of violet light (405 nm) to become fluorescent and detected by 488-nm excitation again (switch-on). This switching-on and -off cycle can be repeated more than 100 times for a single Dronpa molecule imbedded in polyvinyl alcohol film before it finally photobleaches [138].

The photoinducible switching between the on- and off-states of Dronpa is likely due to the light-induced fluorophore protonation state and/or conformational changes [138–144].

irreversibly (off $\rightarrow$ on); *photoswitching* to the process of reversible fluorescence turn-on and turn-off cycles (off $\leftrightarrow$ on); and *photoconversion* to the process of light-induced change of excitation/emission spectra between two color regions (on_1 on_2).

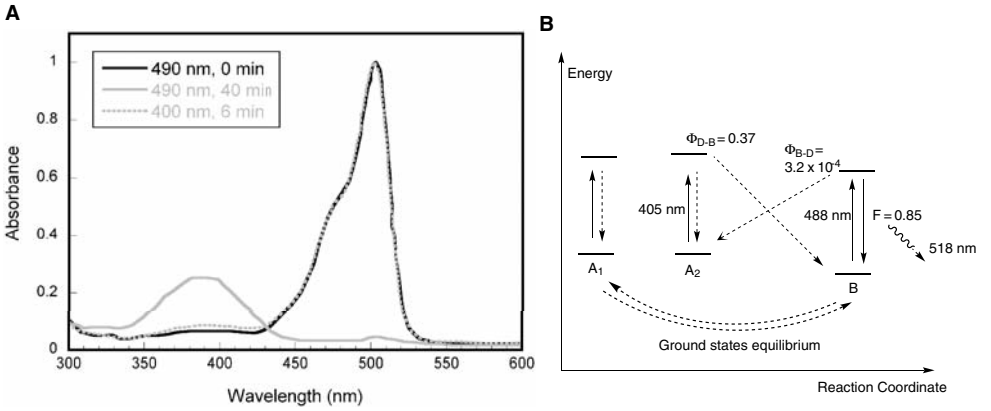


Figure 3.5. A. Absorption spectra of Dronpa before 490-nm illumination (*black*), followed by 490-nm illumination for 40 min (*gray*) and then after 400-nm illumination for 6 min (*dotted gray*) The three curves exhibit an isosbestic point at 428 nm. Irradiation at 490 nm for a prolonged time completely diminishes the absorption peak around 500 nm and increases the population that absorbs at 390 nm. Brief irradiation at 405 nm completely restores the 500-nm peak and diminishes the 390-nm peak. (Adapted from Ando et al. [127].) **B.** Proposed photoswitching states of Dronpa. Form B is the deprotonated form, which can be excited at 488 nm to emit fluorescence; A_1 is the protonated form, which exists in equilibrium with form B through ground-state deprotonation and protonation and does not fluoresce; A_2 is the protonated form, which only forms through the excited form B and cannot relax spontaneously to B but can be converted to B through irradiation at 405 nm.

Dronpa's absorption spectrum (Figure 3.5A) has one major peak at 503 nm and a minor one at 390 nm, corresponding to the deprotonated (B) and neutral (A_1) forms of the fluorophore, respectively (Figure 3.5B). Only the deprotonated form B is fluorescent upon excitation at 488 nm (emission peak at 518 nm, with quantum yield Φ of 0.85). When Dronpa is illuminated at 488 nm for a prolonged time, the deprotonated form B undergoes an efficient ESPT from the excited state to a protonated ground state (A_2), which is not fluorescent. This constitutes the switching-off of the photocycle of Dronpa, with a switching quantum yield $\Phi_{B\cdot D}$ of about 3×10^{-4}. The protonated form A_2 relaxes very slowly (days) to the deprotonated form B, is different from the originally neutral form A_1, which exists in equilibrium with the deprotonated form B, and, most remarkably, can be rapidly (within milliseconds) converted to the deprotonated form B by 405-nm irradiation (Figure 3.5B). The conversion of A_2 to B constitutes the switching-on of the photocycle because after 405-nm irradiation, bright fluorescence is regenerated with 488-nm excitation. The quantum yield for the switching-on $\Phi_{D\cdot B}$ is very high (~ 0.37). The on-time, that is, the time for a single Dronpa molecule to stay at the deprotonated form B during 488-nm illumination, is linearly dependent on the excitation power in the range tested (0.2–3 kW/cm^2)—the higher the power, the shorter is the on-time. Similarly, Dronpa's switching-on rate is also power- dependent, with stronger 405-nm irradiation giving higher switching-on rates [145].

Structural studies using X-ray crystallography and nuclear magnetic resonance (NMR) confirmed that the fluorophore, formed by the Cys62, Tyr63, and Gly64 tripeptide, is deprotonated in the bright state and protonated in the dark state [140,141,144]. In addition, these studies also suggested that conformation changes in the fluorophore, such as a *cis–trans* transformation as in asFP595 [140], the overall β-barrel structure flexibility [144], and/or the

extensive hydrogen-binding network of the fluorophore with surrounding polar groups [141] contribute to light-induced photoswitching.

Dronpa has been successfully applied to probe the shuttling of signaling molecules between different cellular locations and track cell fates using whole-cell labeling at the ensemble level [127,146–148]. It also has been detected at the single-molecule level *in vitro* [138,142,149] and in fixed cells [150]. However, its imaging at the single-molecule level in live cells has not been reported. Based on Dronpa's high quantum efficiency (0.85) and high absorption coefficient (95,000 M^{-1} cm^{-1}), it is the brightest FP among the selected PI-FPs and there is no reason why its single-molecule imaging in live cells cannot be done. The only drawback, however, could be the reported residual fluorescence when Dronpa is in the off-state (10% of its on-state fluorescence measured at the ensemble level [151]). Therefore, it may be difficult to image single Dronpa molecules in a densely labeled cellular environment because the residual fluorescence of Dronpa molecules in the off-state will contribute to a high background. It is not clear whether the residual fluorescence of Dronpa in the off-state is an intrinsic property of the fluorophore or whether 488-nm irradiation can also switch on Dronpa molecules in the off-state (although inefficiently), creating a dynamic equilibrium between Dronpa molecules in the off- and on-states.

Further engineering of Dronpa produced three mutants that accelerate the switching between the bright and dark states, possibly by reducing the steric hindrance of fluorophore conformational changes [140]. One of these, the V157G mutant rsFastLime, has a good combination of brightness and fast switching kinetics (faster switching-off and spontaneous switching-on rates) (Table 3.3). The other two mutants, M159T and M159S, although switching much faster, are both too dim to be useful in single-molecule imaging in live cells. The fast switching-off property of rsFastLime has been successfully used in photoactivated localization microscopy (PALM) in fixed cells to shorten the imaging time from more than 10 hr to a few minutes [152]. In single-molecule tracking experiments, however, the fast switching-off property of rsFastLime does not necessarily pose an advantage over Dronpa because it causes short fluorescence on-times, limiting the duration of tracking trajectories. However, the spontaneous switching-on property of rsFastLime simplifies instrumentation because only one laser line (488 nm) rather than two alternating lines (488 and 405 nm) satisfies both the switching-on and -off requirements, although the addition of a 405-nm line significantly enhances the switching-on probability.

rsCherry and rsCherryRev

The newest additions to the reversibly photoswitching FP family are the red-emitting rsCherry and rsCherryRev, which were semirationally derived from mCherry [153]. By incorporating into mCherry mutations that likely facilitate *cis–trans* transition and modify the immediate hydrogen-bonding environment of the fluorophore, rsCherry and rsCherryRev were produced with opposite photoswitching properties. In the on-state, both FPs absorb and emit maximally around 570 and 610 nm, which are slightly blue-shifted compared to mCherry. Prolonged illumination using a yellow light (550 nm) turns off rsCherryRev, just as Dronpa is reversibly inactivated on prolonged illumination at its detection light (488 nm). For rsCherry, however, the same yellow light illumination switches on its fluorescence from its dark state. Similarly, blue light (450 nm) has the opposite effect on the two proteins—it turns on rsCherryRev but turns off rsCherry. The brightness of these two proteins is reported to be similar to that of mCherry based on FCS measurements, rendering their single-molecule

detection achievable in live cells. Indeed, rsCherryRev has been applied successfully to live mammalian cells, in which single-molecule, superresolution imaging with a 70-nm spatial resolution was achieved [153].

EosFP

EosFP was cloned from the stony coral *Lobophyllia hemprichii* in 2004 [130]. It was named after Eos, the goddess of dawn in Greek mythology, "who brings light to all the mortals of this earth" (Hesiod, *Theogony*), as it changes its color from green to red irreversibly on violet irradiation. EosFP was first isolated as a tightly associated tetramer but was eventually engineered to be a monomer (mEosFP) by breaking the interfaces of subunits in a similar way to that of DsRed [154]. mEosFP has similar spectral properties and brightness to those of EosFP but was reported to fold poorly in mammalian cells at 37°C [130]. This presents an obstacle for live-cell imaging at the ensemble level but may not have significant adverse effect for live-cell imaging at the single-molecule level because only folded and fluorescent molecules will be selected and imaged. Another way to circumvent this is to use a tandem dimer tdEosFP in which two subunits are connected through a flexible linker sequence, which expresses well at 37°C [155].

The crystal structures of EosFP in the green and red states have been elucidated (Figure 3.6) [156]. The fluorophore of EosFP is formed by the His62-Tyr63-Gly64 tripeptide. In its green-emitting state, the fluorophore is the anionic hydroxybenzylideneimidazolinone, the same as that of a typical GFP. This anionic fluorophore absorbs and emits maximally at 505 and 516 nm, respectively, with a fairly high absorption coefficient ε of 72,000 M^{-1} cm^{-1} and a quantum yield Φ of 0.7 (the ε and Φ of mEosFP are similar to those of EosFP). On irradiation around 390 nm, the fluorophore undergoes a peptide backbone breakage between the N_α and C_α atoms of His62 via a β-elimination mechanism [157], resulting in two peptide fragments (8 and 20 kD, respectively) that still bind to each other to form the β-barrel structure but separate on denaturing gel electrophoresis [130]. The reaction is likely mediated by Glu212 and Phe61, in which Glu212 may act as a base to extract one proton from the C_β atom of His62, whereas the imidazole ring of His62, which is transiently protonated during photoinduction, simultaneously donates a proton to the Phe61 carboxyl, making the peptide bond between Phe61 and His62 an ideal carboximidic leaving group in the elimination reaction. After the leaving of Phe61, the π-bonding system extends to the imidazole ring of His61 via a *trans*-alkenylene and forms the red-emitting fluorophore with shifted absorption and emission maxima at 569 and 581 nm, respectively.

It is noteworthy that the initial substrate for photoconversion is the protonated instead of deprotonated fluorophore, as evidenced by the drastically increased yield of photoconversion at low pH. This property may be used to one's advantage when imaging EosFP in an acidic environment such as the early endosomes.

The photoconverted mEosFP is about half bright after conversion, but because it emits in the red region where cells' autofluorescence is also reduced, single-molecule detection of EosFP in live cell is still possible. Indeed, the tandem dimer tdEosFP has been used successfully in single-molecule tracking in live cells [123]. Because of its photoconvertibility and the high contrast ratio between the converted and unconverted forms, thousands of single-molecule tracking trajectories can be obtained from a single cell by alternating the converting and detecting laser beams [123]. In another study, tdEosFP was paired with Dronpa in a clever imaging sequence to demonstrate dual-color PALM imaging [158]. Although this work was

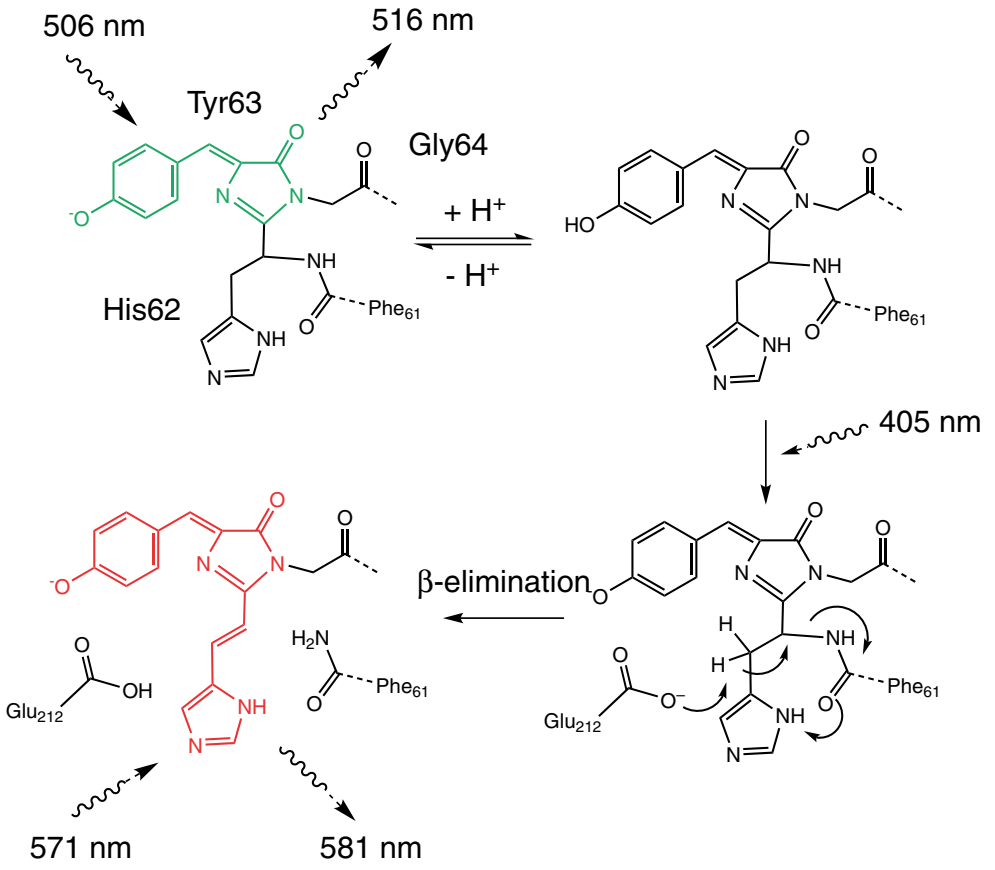


Figure 3.6. Proposed mechanism for the photoconversion of EosFP from its green form to red form. (Adapted from Nienhaus et al. [156].)

carried out in fixed cells and the overlapping green emission of tdEosFP before conversion in Dronpa's channel does not allow simultaneous imaging of both proteins in densely labeled live cells, it demonstrated that Dronpa and photoconverted tdEosFP have little spectral crosstalk and that EosFP has the highest contrast ratio between its photoinduced and noninduced forms compared to other PI-FPs including Dronpa, PA-GFP, and PA-CFP, making it an attractive probe [158].

Dendra2

Dendra2, the improved version of the monomeric fluorescent protein Dendra, is a recent addition to the PI-FP family [131]. Dendra2 is similar to mEosFP in that it can be converted irreversibly from a green-emitting species (excitation and emission maxima at 490 and 507 nm, respectively) to a red-emitting one (excitation and emission maxima at 553 and 573 nm, respectively). Denra2 differs from mEosFP in that its photoconversion can be mediated both by intense blue light irradiation of 488 nm and violet light irradiation of 405 nm; the latter is usually the only workable light source for photoconversion of many other PI-FPs,

including EosFP. Because the fluorophore of Dendra2 is formed by the same His-Tyr-Gly tripeptide, the photoconversion mechanism mediated by 405-nm irradiation is thought to be similar to that of EosFP. However, it is not clear what the mechanism is underlying the 488-nm blue light–mediated photoconversion because no crystal structure of Dendra2 is available. The use of 488-nm light to induce the green-to-red conversion of Dendra2 poses less toxicity to the cells than the 405-nm light, but high power, usually in the range of a few hundred W/cm^2, is required. Dendra2 is reported to express well in both bacterial and mammalian cells at 37°C, indicating efficient folding and fluorophore maturation processes. Although the overall brightness of Dendra2 before and after conversion is slightly less than that of EosFP, it has been successful imaged in live bacterial cells at the single-molecule level [159].

3.6. Special Considerations for Live-Cell Imaging

To achieve single-molecule detection in live cells, the main challenge is to enhance the signal-to-noise ratio by decreasing the cellular autofluorescence background and/or increasing detectable fluorescent signal. In addition, photodamage caused by laser exposure has to be carefully minimized to ensure a normal physiology of cells throughout an experiment.

3.6.1. Autofluorescence

The main hurdle in detecting single molecules in live cells is the background due to the intrinsic cellular fluorescence (autofluorescence). In bacterial cells, autofluorescence is usually homogeneous and diffusive, while in mammalian cells autofluorescence is heterogeneous, with most of the autofluorescence coming from various organelles such as mitochondria in the cytoplasm, although it is absent from the nucleus [160]. The strength of autofluorescence is also dependent on cell type, with microphages, neurons, and sperm cells exhibiting particularly strong autofluorescence.

The main sources of autofluorescence are nicotinamide adenine dinucleotide phosphate [NAD(P)H] between 400 and 500 nm and flavins and flavoproteins between 500 and 600 nm. Their emission levels are dependent on the local environment, cell type, growth condition, and cell physiology [160–162]. To increase the detectability of a single fluorophore above the autofluorescence background in the visible region, the most important tactic is to choose a fluorophore with red-shifted spectra compared to that of flavins whenever possible. The top panel of Figure 3.3 shows the excitation and emission spectra of flavin adenine dinucleotide (FAD). It is evident that by red-shifting excitation beyond 500 nm, autofluorescence from flavins can be greatly reduced. Therefore, YFPs and RFPs are much more suitable than GFPs for single-molecule imaging in live cells. One should also keep in mind that flavins exist at high concentrations inside cells (10^7 molecules/cell [162]); even excitation at the tail region of its excitation spectrum will generate a large fluorescence background and could easily overwhelm the signal from a single FP molecule. Therefore, it is necessary to further minimize the autofluorescence background of cells. Because NAD(P)H, flavins, and flavoproteins are essential cellular components, it is difficult to achieve background reduction by reducing their cellular levels and still maintain good cell physiology. Despite these difficulties, there are a few ways to alleviate the problem and allow single-molecule detection.

First, defined rather than complex medium usually works better for reducing cellular autofluorescence. For example, for bacterial cells, M9 instead of LB medium was found to reduce autofluorescence significantly [163]. For mammalian cells, Leibovitz's L15 medium

or Dulbecco's Modified Eagle Medium (DMEM) buffered with *N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid (HEPES) (no phenol red) has been used. For some cell types, certain media may induce high autofluorescence; therefore, different growth media should be tested prior to experiments. Second, cells under stressful, crowded, or otherwise poor growth conditions often have increased cellular autofluorescence [160, 163]; thus, keeping cells in a healthy growth condition is essential. Third, using a small excitation volume through the application of total internal reflection (TIR), confocal illumination or shrinking the wide-field illumination area can effectively reduce the background by reducing the number of autofluorescence-emitting molecules in the excitation volume. Fourth, cells can be prephotobleached to eliminate autofluorescence background before the collection of data. This is only applicable for short-term experiments because extensive prephotobleaching usually produces severe photodamage to cells. Finally, the emission fingerprinting mode offered by the Zeiss LSM META confocal microscope allows one to collect and resolve the spectra of fluorophores with closely overlapped emission at a relatively fast speed [164,165]. Although the detection of single molecules in this mode has not been realized (possibly due to the low detection efficiency of this commercial system), the basic principle of spectral instead of intensity separation holds promise for rejecting autofluorescence background from imaging fluorophores. In addition, it should be noted that although fluorescence lifetime gating has been suggested and successfully applied to separate autofluorescence and imaging fluorophores [76,166,167], this technique does not provide much advantage if the imaged fluorophores are FPs. This is because the fluorescence lifetime of FPs (~3 ns [76]) is not significantly different from that of flavins or flavoproteins (~1–5 ns [168, 169]).

3.6.2. Fluorescence Signal Enhancement

In addition to focusing on minimizing autofluorescence background, one should also consider enhancing the fluorescent signal from a given fluorophore. One way to achieve this is to spatially confine a fluorophore in a small region of the cell so that its diffusion is slowed down and its fluorescence is not spread over an area greater than a diffraction-limited spot during the acquisition time (usually in the range of a few tens or hundreds of milliseconds) [84,92]. An extension of this idea is the stroboscopic detection of a single RFP freely diffusing in the cytoplasm of *E. coli* cells using very high excitation power (~50 kW/cm^2) and short laser exposure time (0.3 ms) [116,170]. The high excitation power pumps the molecule as fast as possible so that a maximal number of photons are collected within the short period of laser exposure, during which the molecule has not moved so far as to produce a fluorescence image that is too spatially diffuse. Although it provides exciting possibilities for probing protein dynamics directly in the cytoplasm, brighter and more photostable fluorophores are needed because the intense pulsing of excitation often drives FP molecules into a dark state (blinking or photobleaching) [68,74] before enough photons are collected.

Another way to amplify a fluorescent signal is through the use of multifluorophore labeling of single macromolecules or enzymatic conversion of fluorogenic substrates [3,9,10,171,172]. For example, to observe the dynamics of a single chromosome molecule, an array of the binding site of the transcription factor LacI (96–256 copies) is integrated into the chromosome. On the binding of hundreds of GFP-LacI fusion protein molecules, a bright fluorescent locus is formed and serves as a marker to track the dynamics of the chromosome inside cells [6–8]. Similarly, by integrating 96 MS2 RNA-binding loops into an mRNA molecule and expressing MS2-GFP fusion proteins in the same cell, transcription can be monitored in real time in live cells and the subsequent transportation of mRNA

transcripts followed [3–5,173]. Another method is to use an enzyme that converts a fluorogenic substrate into fluorescent molecules at a high turnover rate, so that each enzyme molecule generates thousands of copies of fluorescent molecules, enabling the detection of single enzyme molecules in live cells [9,10]. These methods convert the challenging single-molecule detection into multifluorophore detection, making it easy for a biological lab equipped with a regular lamp-excitation fluorescence microscope to conduct investigations at the single-molecule level. However, the tagging of hundreds of copies of FP molecules on a chromosome or mRNA molecule may cause severe perturbation of the normal functions and cellular fates of these macromolecules [4,172], requiring extreme care for meaningful interpretation of experimental data. In addition, the enzymatic amplification method does not allow the imaging of the localization of enzyme molecules because the small fluorescent molecules diffuse quickly inside the cells and the spatial information is lost.

3.6.3. Laser-Induced Photodamage of Cells

Unlike fixed cell imaging, in which photobleaching is of major concern, in live-cell imaging, photodamage must be limited. Photodamage of live cells caused by laser illumination (Figure 3.7) is often observed as abnormal cell morphology, a granular appearance of the cell surface, a condensed nucleus, cell detachment, excessive vacuole formation, necrosis, arrested cell division, and, in some cases, sudden significantly enhanced cellular autofluorescence. Eventually, photodamage leads to cell death and prevents proper interpretation of imaging data. This is particularly detrimental in time-lapse imaging experiments, in which cells need to maintain their normal physiology for hours or even days.

The phenomenon of photodamage has been investigated, but its detailed mechanism is far from being conclusively determined. In general, the severity of photodamage of live cells is dependent on illumination wavelength, dose, and cell type—the shorter the wavelength, the higher the illumination dose, and the more fragile the cell type, the more severe is the damage. In the UV region, where nucleic acids and proteins absorb strongly, pyrimidine dimer formation, DNA strand breaks, and DNA–protein cross-linking caused by photon–DNA interactions have been reported [174–176]. In the near-infrared region (700–1100 nm), where high laser power density in the range of 10^7–10^8 W/cm^2 in optical-trapping and two-photo microscopy is used, laser-induced transient cell heating, two-photon absorption, and generation of reactive oxygen species (ROS) have been suggested to cause photodamage [177–180]. In the visible light region (500–600 nm), generation of ROS such as singlet oxygen and free radicals during photobleaching of excited biomolecules and fluorescent molecules has been documented [181–183]. This is especially problematic when high concentrations of fluorophores are used (Figure 3.7B, bottom panel). ROS may directly attack the lipid membrane, proteins, and nucleic acids through oxidative reactions [184] and perturb the cellular redox homeostasis through the accumulation of free radicals [185].

Because most single-molecule, live-cell imaging experiments are carried out in the visible region using a continuous wave (CW) laser that offers relatively low excitation power density (<1 kW/cm^2), and because FPs are usually expressed at very low levels, photodamage of cells is often minimal. Indeed, studies found that illumination of human lymphocyte cell cultures between 540 and 650 nm at low laser dose produces negligible photodamage [186] and that only when power exceeds 3.5 kW/cm^2 with 532-nm excitation do *E. coli* cells show appreciable damage [187].

One should still keep in mind that all cells are photosensitive, and free radicals generated through the photochemistry of excited fluorophores cannot be completely eliminated.

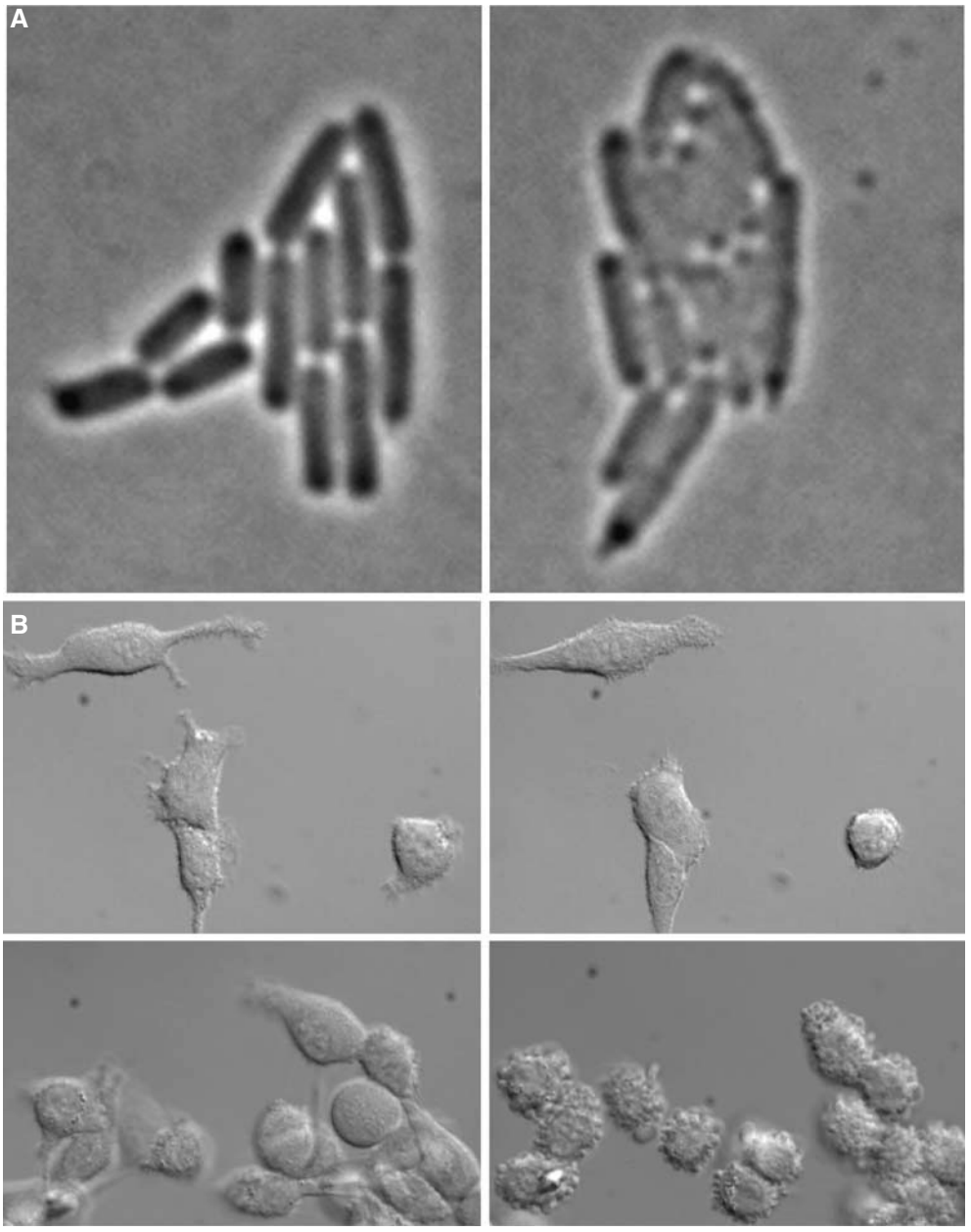


Figure 3.7. Effect of photodamage on live cells. Bright-field images of *Escherichia coli* cells (MG1655) (**A**) and mammalian cells (9L/LacZ, ATCC CRL-2200) (**B**) before (*left*) and after (*right*) prolonged laser illumination at 514 and 635 nm, respectively. Before laser illumination, cells appeared to have smooth cell surface and clearly defined edges. After laser illumination (514 nm, 0.6 kW/cm^2, 600 ms every 3 min for 3 hr), the photodamaged *E. coli* cells appeared to have blurred edges, dark spots, and rugged cell surfaces. The 9L/LacZ cells (**B, top**) did not show much morphology difference before and after laser illumination at 635 nm (0.4 kW/cm^2, 100-ms exposure every 20 sec for 120 min), except that cells were slightly bulged. In contrast, when cells (**B, bottom**) were loaded with 10 μM of dimethylacridinone (DDAO)-Gal (a fluorogeneic substrate of β-lactosidase) that was converted to highly fluorescent DDAO at high concentrations, cells showed severe photodamage, as evidenced by detachment, blebs, and shrunken cell sizes after illumination at the same condition for 70 min. This demonstrates the phototoxicity caused by the photochemistry of excited dyes.

Prolonged laser illumination, even at low power density and in the visible wavelength region, will produce enough free radicals that they will eventually exceed the buffering capacity of the redox system of cells and lead to severe photodamage of cells. A compromise between the best image quality and the health of cells must be reached. The most effective method of reducing photodamage is to use the lowest possible light intensity and to separate laser exposures widely during acquisition. This is because cell damage and free radical generation show a nonlinear relationship to the excitation light intensity, and high excitation intensities with brief exposure times promote greater damage to cells than low excitation intensities with long exposure times [183,188]. One should carefully adjust excitation light level and sampling rate to minimize the adverse effect of photodamage, and at the end of each experiment always compare cells that have and have not been subjected to laser exposures to determine whether significant photodamage has occurred.

3.7. Instrumentation

3.7.1. Illumination Source

Although single-molecule detection has been achieved using lamp excitation [189], a monochromatic and highly collimated laser beam is much preferred for single-molecule imaging experiments. Laser illumination offers advantages that cannot be matched by lamp excitation: It provides an extremely pure excitation wavelength, delivers high excitation power, and diverges very little during travel. The lasers most commonly used for single-molecule imaging in live cells are CW lasers, which produce a continuous beam output that has constant amplitude and frequency in time. CW lasers emitting in the blue, green, and yellow regions include gas lasers such as argon and krypton ion lasers, solid-state lasers such as diode lasers, and liquid lasers such as dye lasers. The major manufacturers for these lasers include Coherent, Inc., Spectra-Physics, and CrystaLaser. In Figure 3.3 the major lines of commercially available lasers are plotted along with the excitation spectra of commonly used fluorophores. Liquid dye lasers offer the flexibility of tuning excitation wavelengths continuously to match the excitation maxima of a fluorophore, but they generally require more maintenance than the gas or solid-state lasers because the dye needs to be changed frequently to ensure a constant power output. Pulse lasers, which generate intense laser pulses in milliseconds or femtoseconds, are used in multiphoton microscopy and are not discussed here.

The guidelines for selecting a suitable laser of excitation wavelength for single-molecule, live-cell imaging is not much different from that for *in vitro* single-molecule imaging. First, the laser wavelength should be close to the excitation maxima of a chosen fluorophore; in two-color imaging, the criterion is to minimize cross-talk of the two fluorophores while maximizing their excitation. Second, the laser power should be sufficiently high so that when a large illumination area is used an average excitation power density between 0.1 and 1.0 kW/cm^2 can still be achieved. Third, the laser should have a stable single mode with a Gaussian beam profile (TEM_{00}) so that the beam can be focused down to a diffraction-limited spot or expanded to illuminate a large area uniformly. Lastly, the laser should also have high pointing and power output stabilities, which is especially critical for long-time-lapse imaging of live cells. In addition, because the most popular fluorophores used in live-cell imaging are FPs, which are much larger than small-molecule organic dyes, their rotational movement inside cells may be slow and comparable to the imaging time scale. Therefore, the intensities of individual molecules will have large variations due to the relative orientations of their dipoles with respect to that of a polarized laser beam. These variations are larger than what

one would expect from pure photon statistics and sometimes make interpretation of single-molecule data difficult. To circumvent this problem, the linearly polarized incident laser beam can be converted to circularly polarized light by inserting a quarter-wave plate in the optical pathway. In some cases, the coherence of the laser beam is also scrambled to avoid speckle patterns in the imaging area due to the interference of the coherent laser beam by dust or surfaces of optics in the pathway [190].

3.7.2. Illumination Mode

To reduce background and enhance single-to-noise ratio in single-molecule, live-cell imaging, besides carefully prepared cell samples and selected optics, minimization of excitation volume is essential. Three illumination modes—wide-field, confocal, and TIR—are often used and discussed next (Figure 3.8). For the general principles of fluorescence microscopy, the reader should consult the textbook by Murphy [191].

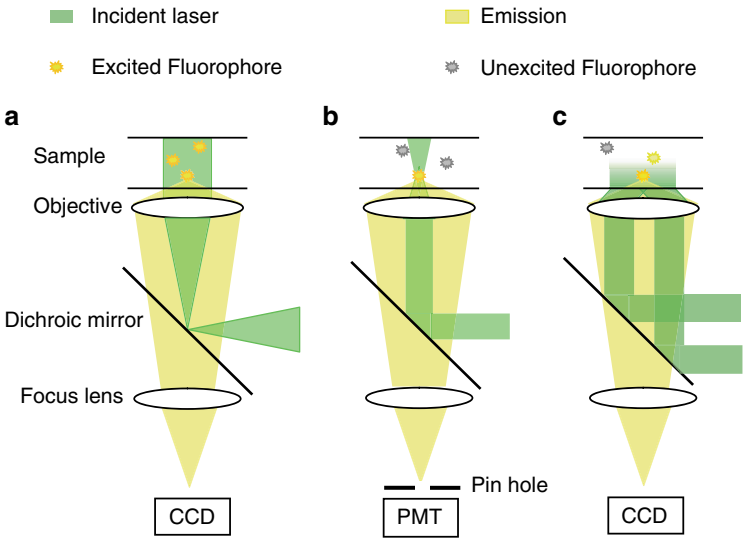


Figure 3.8. Schematic drawing of three illumination modes: wide field (**A**), laser scanning confocal (**B**), and objective-type total internal reflection (TIR) (**C**). The excitation beam is shown by green shading and the emitted fluorescence by yellow shading. Fluorophores that are in the excitation volume are shown as yellow stars, and the ones outside are shown in gray. In the TIR mode, the evanescent field is shown in gradient green color.

Wide-Field Illumination

Conventional wide-field illumination, in which a beam of collimated laser light is used to illuminate a large area of sample, is the most commonly used mode in fluorescence microscopy. It is easy to set up and often cost-effective compared to other illumination modes. It offers a large view field (up to the full size of a CCD chip) and allows simultaneous imaging of multiple molecules, improving data throughput. In single-molecule imaging, however, a smaller illumination area on the range of a few to a few tens of μm^2 instead of the full chip size is often used to minimize the background. In practice, a combination of lenses with different focal lengths is used to first expand and focus the laser beam to the back aperture of the

objective at a desired cone angle to reach an illumination area of 100 μm^2. Then an iris is used to cut off the fringe of the laser incident beam so that only the center area of about 20 μm^2 (or larger if cell size is larger) is used for imaging. This efficiently cuts down the background and ensures homogeneous excitation throughout the illumination area. The tradeoff of this configuration is obviously the reduced number of cells and molecules in the view field, which is easily compensated for by imaging multiple areas sequentially. A more serious drawback of wide-field illumination is that, for high–numerical aperture objectives, the depth of field where objects stay in focus is less than 1 μm; thus, for thick cells (5–30 μm) other fluorescent molecules above this 1 μm in the excitation path of the collimated laser beam will generate out-of-focus light and contribute to high background. Therefore, wide-field illumination is usually used for small and flat cells such as bacterial and epithelial cells.

Confocal Illumination

Confocal illumination effectively minimizes background signal by exciting a sample with a focused laser beam that has the dimension of a diffraction-limited spot, the smallest excitation volume achievable in far-field optical microscopy, and using a pinhole that is in an optically conjugate plane in front the detector to reject light below or above the focal plane of the objective. Because of the point excitation, illumination of a large area is achieved by scanning the area pixel by pixel through computer-controlled movement of the sample or laser beam, and fluorescence emission at each point is collected by a photomultiplier tube (PMT) through the pinhole. The output from the PMT is then built into an image according to their spatial positions and displayed by a computer.

Confocal illumination offers excellent ability in reducing background and generating high-contrast three-dimensional images, but it is not the best choice for single-molecule imaging in live cells. First, the point-scanning format renders a slow frame rate, typically at a few frames per second for a commonly used laser scanning confocal microscope (LSCM). Therefore, fast cellular dynamics at the millisecond time scale are not accessible. Shrinking the scanning area or reducing the dwell time of the laser at each pixel increases the frame rate to 20–30 frames/sec. However, this usually does not permit single-molecule detection because a laser dwell time of 0.1–1.0 μs/pixel at a single emitter in rapid scanning mode does not generate enough photons to distinguish the signal from background noise. Second, the PMT used in a confocal microscope has low quantum efficiency (0.5–0.6) for detecting photons, rendering single-molecule detection difficult.

A new format of confocal microscopy, spinning-disk confocal, also called Nipkow disk confocal [192,193], demonstrates improvements over LSCM, in that it offers relatively fast imaging speed. Spinning-disk confocal uses one disk carrying an array of spirally arranged microlenses to focus the laser beam onto a second disk carrying an array of pinholes arranged in the same pattern. The microlenses and pinholes are arranged in such a pattern that when the two disks spin at high frequency, they effectively split the excitation laser beam into thousands of beams, simultaneously illuminating every point of the imaging area; thus, a high imaging speed at 1,000 frames/sec can be achieved. In addition, a CCD camera instead of a PMT is used, so that the detection efficiency is also improved. Based on these techniques, spinning-disk confocal is a better choice over LSCM for single-molecule imaging in live cells, although the fluorescence signal is still reduced due to the splitting of the excitation light.

Total Internal Reflection Illumination (TIR)

TIR illumination offers the best combination of background reduction, wide-field data collection, and high imaging speed through the use of an evanescent field as the excitation source (Figure 3.9). The evanescent field is generated by the excitation laser beam traveling at a high incident angle at the interface between the glass coverslip and water. At a particular angle determined by the difference between the refractive indexes of glass and water, the light is totally reflected from the interface but generates a very thin electromagnetic field at the same frequency as the incident light in the water phase. This electromagnetic field is called the evanescent wave, and its intensity decays exponentially perpendicular to the interface with a distance dependence of

$$I_z = I_0 e^{\frac{-z}{d}}$$

where I_z is the excitation intensity at a distance z from the interface, I_0 is the excitation intensity at the interface, and d is the characteristic penetration depth, which is defined as the distance at which the evanescent wave intensity drops to I_0/e. It can be calculated as

$$d = \frac{\lambda}{4\pi\sqrt{(n_1^2 \sin^2(\theta) - n_2^2)}}$$

where λ is the incident light wavelength, n_1 is the refractive index of the higher–refractive index medium such as glass (1.518), and n_2 is the refractive index of the lower–refractive index medium such as water (1.33) or cell cytosol (1.33–1.37). Here θ is the incident angle that is larger than the so-called critical angle θ_c, the incident angle at which total internal reflection of the light occurs. The angle θ_c is related to only the refractive indices of n_1 and n_2 by

$$\sin(\theta_c) = \frac{n_2}{n_1}$$

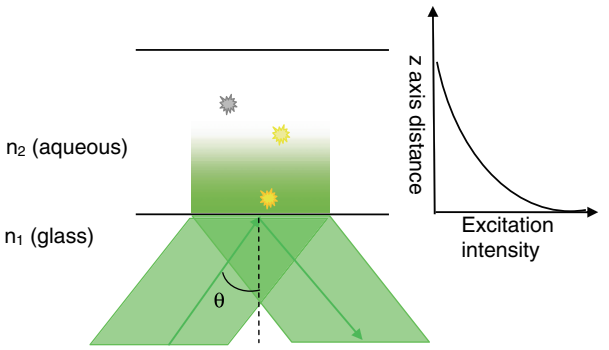


Figure 3.9. Schematic drawing of the concept of total internal reflection at the interface of glass and water. The laser beam passes through the glass and is reflected at the glass–water boundary at an angle larger than the critical angle θ_c, establishing an evanescent wave that penetrates into water in the sample chamber and decays exponentially along the z axis, therefore exciting only a few hundred nanometers from the interface.

Based on these equations, one can calculate that for a 514-nm excitation with an incident angle of 65°, the penetration depth d in cell cytosol is about 300 nm. Increasing the incident angle further decreases the penetration depth and allows only a thin layer above the coverslip to be effectively excited. Therefore, fluorescent molecules above the thin layer cannot be excited, and background is dramatically reduced. In addition, the illumination area of the incident laser beam can be adjusted to be as large as what is used in wide-field illumination, so that multiple molecules can be imaged simultaneously at a high frame rate. These properties render TIR illumination nearly the best choice for single-molecule imaging in live-cell imaging, although it limits observation to cellular events that occur close to the glass/cell interface. In addition, integrated fluorescent intensities of single molecules will have large variations, depending on their distances from the glass/cell interface; therefore care should be taken when determining whether a detected fluorescent spot is due to a single molecule or not.

There are two types of TIR illumination configurations, prism and objective type. In the prism-type configuration, cells are sandwiched between a glass slide and a coverslip, and the excitation laser beam is guided to the sample through a prism mounted on the top of the glass slide. The incident angle is adjusted by translating the beam relative to the prism. Fluorescence emitted by the sample is collected through the objective at the bottom of the coverslip. Because oil objectives with high numerical aperture are usually used to allow sensitive single-molecule detection, the thickness of the sample cannot exceed 15 μm, which often comprises the viability of cells during an extended period of imaging. Nevertheless, prism-type TIR has been successfully applied to imaging single GFP molecules in live fission yeast, *Dictyostelium discoideum*, and Jurkat cells [194].

The objective-type TIR is more suitable for live-cell imaging than the prism type. In this configuration, the excitation laser beam is focused to the periphery instead of the center of the back focal plane of the objective. Translating the beam in parallel to the center axis of the objective changes the incident angle of the light but not the position of illumination area. Tilting the beam relative to the center axis of the objective changes the incident angle as well, but the position of the illumination area is also changed; thus usually this configuration is not recommended. Because only an incident angle larger than the critical angle allows TIR illumination, and that maximal incident angle of the laser beam is limited by the numerical aperture (NA) of the objective, in that

$$\sin(\theta) = \frac{\mathrm{NA}}{n_3}$$

where θ is the maximal incident angle allowed by the objective and n_3 is the refractive index of the working medium (oil at 1.52, matching that of glass) of the objective, one can derive that NA needs to be greater than n_1, the refractive index of a cell (1.37), to achieve TIR illumination. Therefore, a high-NA objective, typically at 1.45, is essential in this type of configuration.

3.7.3. Camera-Based Detectors

Array detectors such as charged-coupled device (CCD) cameras instead of point detectors such as PMTs or APDs are usually used for single-molecule imaging in live cells. The selection criteria of those cameras are the same for single-molecule *in vitro* imaging. In

general these cameras are capable of single-photon detection and possess high quantum efficiency in converting photons into electrons, low readout noise, and a fast frame rate.

A CCD camera detects incoming photons on a two-dimensional array of tiny detector elements (or pixels) that are microns in size. When an incident photon hits one of the detector elements it has a probability of generating one electron from the element. This probability is the quantum efficiency of a CCD camera, and it can reach 0.95 in the newest generation of electron-multiplying (EM) CCD cameras. The electrons generated by the photons are then stored in the element until it is ready to be read out; charges are moved from each element of the array to an output register, and the contents of the output register are then amplified and read out. The speed to move the charges in each element and the output register limits the readout speed of the camera and hence the time resolution of an experiment.

Both intensified ICCD cameras and EM-CCDs have been used for single-molecule detection, but EM-CCDs possess clear advantages over ICCDs. EM-CCDs have high quantum efficiency (up to 0.95), while that of ICCDs is 0.2–0.5. In addition, EM-CCDs effectively minimize readout noise by amplifying the signal before readout noise is added by the output register; therefore, high frame rates of a few hundred hertz can be reached, and the readout noise is still negligible. By using small imaging area or binning pixels, one can further increase the frame rate but at the cost of decreased spatial resolution. A range of manufacturers such as Andor Technology, Hamamatsu, and Roper Scientific offer EM-CCDs that have quite comparable specifications for the key features such as quantum efficiency and pixel size but differ subtly in operating temperature, noise level, and so on. Therefore, the choice of which camera to use is dependent on personal preferences and specific applications.

3.7.4. Live-Cell Sample Preparation

A well-maintained cell chamber for optimal cell growth during a prolonged imaging period is essential for the success of experiments. In general, such a chamber has a coverslip at the bottom through which attached cells are imaged. Growth medium is added on top of cells, and the temperature of the chamber is maintained constantly. The detailed configuration of the chamber is dependent on the type of cells and the specific need of each experiment. The following paragraphs discuss the sample preparation for generic bacterial and mammalian cells.

For continuous growth of bacterial cells (or yeast cells) on the microscope stage, a sealed growth chamber is often used. The chamber can be homemade by spotting bacterial culture on top of a low–melting temperature agarose gel pad made with M9 growth medium, with the gel pad sandwiched between two coverslips separated by a spacer. The whole chamber is then sealed with glue to avoid dehydration over long periods of imaging. An objective heater can be employed to warm the chamber through the contact of the objective with the bottom coverslip of the chamber. Although this method is cost-effective and bacterial cells in the chamber can grow for multiple generations, the temperature of the chamber is not well regulated and often exhibits large fluctuations if the objective heater temperature is set much higher than the ambient room temperature. Therefore, this method is best suited for bacterial cell imaging at room temperature or when temperature fluctuations do not have significant impact on the cellular process of interest.

A commercially available cell growth chamber (Bioptechs, FCS2) has a similar geometry as the homemade chamber, but the temperature of the chamber can be accurately regulated through a specially designed glass slide at the top of the chamber. In addition, the FCS2 system allows the perfusion of growth media through the chamber, making it possible to study

the responses of cells to environmental stimulations. For a detailed description of how to use this chamber to monitor live bacterial cells, readers are referred to Xiao et al. [163] and the manufacturer's instructions.

The FCS2 system can also be used for mammalian cells. Cells are first cultured on coverslips immersed in growth media as usual, and then the coverslip is assembled into the closed chamber for observation. An open-dish format, in which the chamber is not sealed by the use of a top glass slide, is also popular in live-cell imaging. The easiest open-dish format is essentially a culture dish with a coverslip mounted at the bottom. This format allows easy addition of chemical stimuli but, again, suffers from unregulated temperature and media evaporation during long experiments. One important consideration for mammalian culture on a microscope stage is that if the medium is not perfused with CO_2 during the experiment, a growth medium with its pH not dependent on CO_2 such as Leibovitz's L15 should be used.

Besides these basic cell culture formats for live-cell imaging, many companies manufacture chambers with different geometries to meet specific needs and that address some of the concerns in the foregoing. Readers are referred to a collection of those companies (and other considerations for maintaining live cells on a microscope stage) at http://www.microscopyu.com/articles/livecellimaging/chamberresources.html. In addition, there is an excellent review describing the methods and considerations in live-cell imaging in much more detail [195].

The foregoing descriptions are for general live-cell imaging. To achieve single-molecule detection, however, there are some special considerations. The coverslip used in these growth chambers has to be thoroughly cleaned using cycles of ethanol, potassium hydroxide, and acetone [163]. This is to avoid dirt and background fluorescence from the coverslip. For this reason, commercially available culture dishes with a coverslip mounted at the bottom are not suitable for single-molecule imaging. If this format has to be used, it is recommended to manually mount cleaned coverslips to culture dishes with a hole drilled in the middle. In addition, defined, colorless growth medium needs to be used to minimize background fluorescence from the medium. It is also not recommended to use a large box to encapsulate the entire microscope and blow warm air into the box to maintain culture temperature. The air flow inside the box creates instability in the imaging systems, which is especially critical if sub–diffraction-limited resolution of single molecules is to be achieved. In addition, the temperature in those large boxes is usually not homogeneous, and a slight change of temperature even of 1°C could cause the optics in the microscope or box to move and cause focus or laser alignment drift.

3.8. Applications

In the last few years, single-molecule imaging in live cells have been employed to study gene expression [4,23,84], transcription factor–DNA interactions [196], cell signaling [31,33,197], protein complex composition [198–200], membrane protein dynamics [91,94,95,123,201–204], and cytoskeletal protein dynamics [92,96,159,205]. In addition, by amplifying the fluorescent signal by multifluorophore labeling or using a fluorogenic substrate, it has been possible to examine the dynamics in gene expression [4,10] and the movement of individual DNA and mRNA transcripts [3,5–8]. The following subsections discuss a few representative studies of single-molecule imaging in live cells. Studies in membrane and cytoskeletal protein dynamics share similar methods in that single-molecule tracking is used; they are discussed in Chapters 1 and 8and are not included here.

3.8.1. Gene Expression

Gene expression is inherently a stochastic process due to the small copy numbers of the participants such as DNAs, mRNAs, and transcription factors [11–13,206,207]. However, this stochasticity is often masked in ensemble studies and underestimated when overexpression of a gene has to be used to achieve detectable signal. If each single protein or mRNA molecule produced from a gene can be observed and counted in real time inside a live cell, the stochasticity of gene expression can be quantified and the molecular mechanisms of translation and transcription revealed.

To achieve this goal, Yu and Xiao et al. developed a single-molecule gene-expression reporter that allows the counting of protein molecules generated one at a time for multiple generations in live *E. coli* cells (Figure 3.10A) [84]. The reporter is generated by fusing a membrane-targeting sequence (MTS) to the YFP variant Venus. On expression, the reporter quickly folds and is targeted to the membrane so that its diffusion is slowed down. Therefore, the emitted photons of the reporter can be collected from a diffraction-limited spot, greatly enhancing the signal-to-background ratio. Using this reporter, the authors monitored the leaky expression of a repressed *lac* promoter, revealing that protein molecules are generated in randomly spaced bursts, corresponding to a Poissonian transcription process, and that the number of protein molecules generated from one mRNA molecule follows a geometric distribution due to the competition between RNaseE and ribosome binding of the mRNA molecule. This experiment is a good example of how a particular FP is selected to meet the particular experimental design—The FP needs to mature fast enough to report the production of a protein molecule as soon as it is generated and to be photoliable enough so that following its detection the FP can be photobleached quickly to avoid background accumulation.

In a different work by the same research group, the induction mechanism of the *lac* operon was investigated [23]. The authors found that in the presence of an inducer, a significant portion of the uninduced cells contain a few *lac* permease (LacY) molecules, disproving the long-held view in which a single LacY molecule, coupled with the positive feedback of the operon, is sufficient for induction [208]. Without the ability to visualize and count single permease molecules inside the cells, this conclusion would be difficult to reach. In addition, the authors monitored the real-time expression of *lac* permease using the same single-molecule gene expression reporter. They discovered that rare (once every many cell cycles) and large expression bursts of permease molecules (a few hundred), likely caused by the stochastic, complete dissociation of *lac* repressor from the operators, is the prerequisite for induction. This suggests a novel mechanism by which that an *E. coli* cell's decision to be induced or not hinges on the stochastic behavior of a single repressor molecule.

At the transcription level, Golding et al. used a MS2-GFP fusion protein to monitor the real-time production of single mRNA molecules that have 96 repeats of MS2 RNA-binding loops incorporated (Figure 3.10B) [4]. Under induced condition, Golding et al. found that mRNA molecules are produced in clustered bursts and are separated by periods of inactivity, corresponding to stochastic transitions of the promoter between active and inactive periods. This indicates that even in bacterial cells that lack a chromatin structure, there are other mechanisms that cause large fluctuations in transcription activity similar to those observed in eukaryotic cells [209,210].

3.8.2. Transcription Factor Dynamics

Single-molecule techniques can be used to address transient molecular interactions, especially for cellular components that exist at very low copy numbers. For example, the

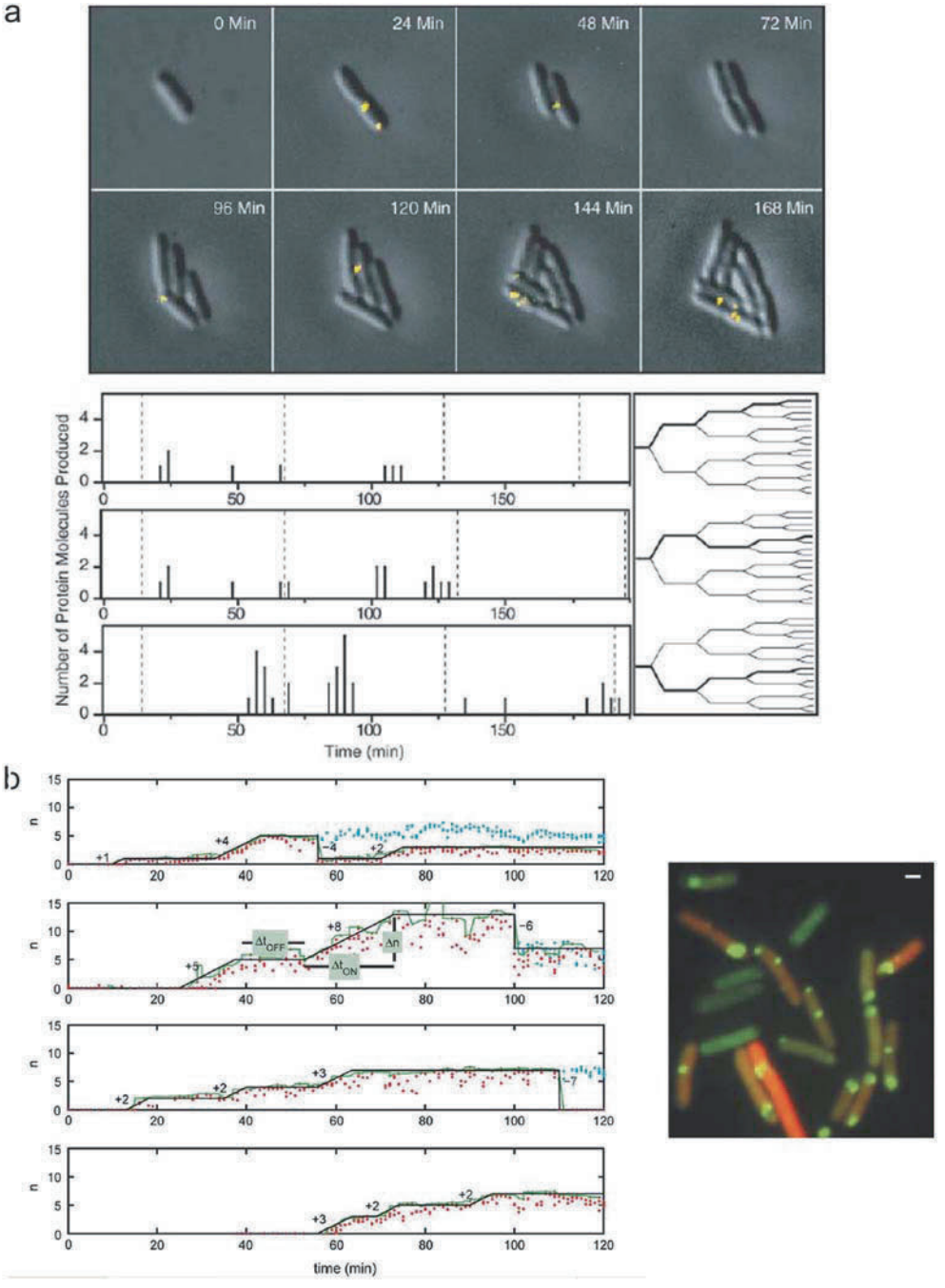


Figure 3.10. A. Single protein molecules counted in live *Escherichia coli* cells as they are generated one at a time [84]. **Top.** Sequences of an *E. coli* colony growing on a microscope stage with expressed single Tsr-Venus proteins shown in yellow. **Bottom.** Time traces of expressed Tsr-Venus molecules counted in real time. **B**. Single mRNA molecules counted in live *E. coli* cells [4]. *Left*: Time traces of mRNA production counted in real time. *Right*: Fluorescence image of *E. coli* cells expressing mRNA molecules tagged with a 96-ms^2 looped bound by MS2-GFP fusion proteins. Each green dot corresponds to one or a few mRNA molecules. (Figures reprinted with permission.)

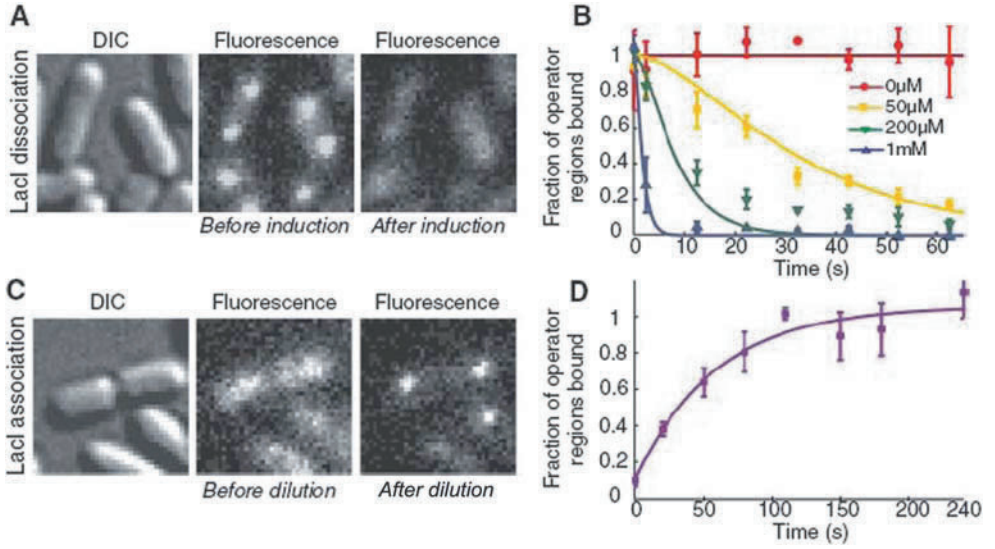


Figure 3.11. Imaging of *lac* repressor dynamics in live *Escherichia coli* cells [196]. **A.** *E. coli* cells before and 40 sec after addition of the inducer isopropyl-β-D-thiogalactopyranoside (IPTG). The disappearance of the bright fluorescent spot in the cells corresponds to the dissociation of LacI repressor from DNA. **B.** Fraction of *lac* operators that are LacI bound plotted as a function of time after induction by various concentrations of IPTG. **C.** *E. coli* cells before and 1 min after dilution of IPTG from 100 to 2 μM and with the addition of 1 mM of the IPTG-binding inhibitor orthonitrophenylfucoside (ONPF. **D.** Fraction of the operators that are LacI-bound as a function of time after rapid dilution of IPTG from 100 to 2 nM. (Figures reprinted with permission.)

binding of transcription factors to their corresponding DNA sitesplays an important role in regulating gene expression and has been subject to intensive genetic, biochemical, and biophysical studies, but the direct observation of the specific and nonspecific DNA binding of a single transcription factor inside live *E. coli* cells has just become possible (Figure 3.11) [196]. Using a LacI-Venus fusion, Elf et al. [196] monitored the binding of a single *lac* repressor LacI to its operator site *lacO* in response to the inducer isopropyl-β-D-thiogalactopyranoside (IPTG) in live *E. coli* cells. In the absence of IPTG, the repressor bound tightly to *lacO* and was detected as a diffraction-limited spot. When IPTG was added, the clear fluorescent spot disappeared within seconds, indicating the quick dissociation of the repressor from the operator site. If IPTG was removed subsequently, the fluorescent spot reappeared within 1 min. From this it was calculated that it takes one repressor molecule up to ∼360 sec to find a *lac*O site inside the cell. During the 360 sec, the repressor spends 90% of its time to diffuse one-dimensionally on nonspecific DNA for up to 5 ms before it hops to another DNA segment to start a new round of searching again. This is the first quantitative demonstration of the search of a transcription factor for its binding site in live cells. This would be impossible without the ability to image single molecules inside live cells.

Similarly, the binding of Venus-tagged T7 RNA polymerase (RNAP) molecules on an engineered T7 promoter was observed in live *E. coli* cells [116]. The T7 promoter is controlled by the *lac* repressor and integrated on the *E. coli* chromosome. After addition of IPTG, multiple T7 RNAP molecules were found to bind to the promoter region, indicating multiple active transcription processes. This is exciting, in that it allows direct observation of the promoter

state, which could possibly provide a mechanism for the transcription bursts observed previously [4,23].

3.8.3. Cell Signaling

Cellsrespond to environmental cues through vast signal transduction networks. Using single-molecule techniques, it is possible to directly observe the signal flow in this process and the dynamics of molecular interactions in real time. Sako et al. labeled the epidermal growth factor (EGF) with Cy3 and found that Cy3-EGF binds to the receptor (EGFR) within a few seconds after its addition into the culture medium [31]. Of interest, during this process, they observed that the fluorescence intensity of one spot, corresponding to one bound Cy3-EGF molecule, suddenly increased twofold, indicating the binding of another CY3-EGF molecule (Figure 3.12). This suggests that the dimerization of EGFR, a prerequisite for the downstream Ca^{2+} response, occurs at least before the binding of the second EGF molecule. In addition, the binding of the second Cy3-EGF to an EGF-bound EGFR dimer in the presence of a high concentration of unbound EGFR monomers (or dimers) indicates that the binding of the first EGF may enhance the binding of the second EGF through a cooperative mechanism.

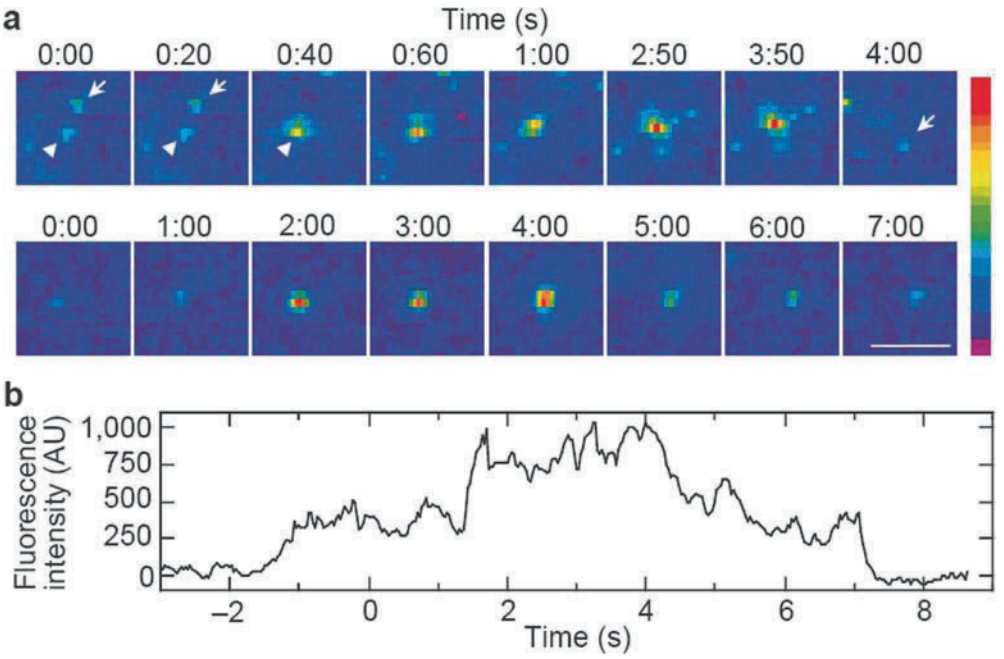


Figure 3.12. Imaging of the dimerization of epidermal growth factor (EGF)–EGF receptor complexes at the single-molecule level [31]. **A.** In the upper panel, two spots (*arrow and arrowhead*) collided at time 0:40 sec and then moved together. The fluorescence intensity increased after the collision. At 4:00 sec, the intensity of the spot decreased to about half (*arrow*), probably because of photobleaching of one Cy3-EGF molecule. In the second panel, a fluorescence spot with the intensity of a single molecule was observed until 1:00 sec; the intensity of the spot increased suddenly between 1:00 and 2:00 sec, then decreased between 4:00 and 5:00 sec. The scale bar represents 5 mm. **B.** Time trace of the fluorescence intensity of the spot shown in the second panel of panel A. A two-step increase and bleaching of the intensity can be seen. The time trace is noisy because of slow, lateral diffusive movement of the spot. (Figures reprinted with permission.)

Conformational fluctuations between the two monomers within one EGFR dimer were also observed by labeling one monomer with Cy3 and the other with Cy5 and monitoring the FRET signal between the pair.

Another study monitored the interaction of the small G protein Ras with its cofactor GTP in live cells using single-molecule FRET (sm-FRET) [33]. In this work, Ras is expressed as a fusion protein to YFP, and the dye-conjugated Bodipy-GTP is microinjected into the cell. On stimulation by the addition of EGF, within 1 min, YFP-Ras molecules became stationary, BodipyTR-GTP molecules appeared as a clear spot on the membrane at exactly the positions of YFP-Ras molecules, and FRET between YFP and Bodipy was observed. However, the BodipyTR dye is not photostable enough to allow the imaging of its turnover by Ras, which would be interesting to correlate with the recruitment of downstream proteins in the signal transduction pathway. This highlights the importance of developing new photostable dyes to allow such applications.

In both studies, sm-FRET was used to observe the protein–protein or protein–ligand interactions. Organic dyes were used as the FRET acceptor not only because the molecules they labeled are small molecules, but also because their spectral properties are better suited for sm-FRET in live cells than are those of available RFPs. This limits the use of sm-FRET in bacterial cells because dye-labeled ligand or proteins cannot be introduced into cells by microinjection. In addition, in eukaryotic cells some ligands or proteins of interest may be hard to modify to incorporate the dye. There is no report of live-cell sm-FRET based on two FPs, possibly due to the fact that current RFPs can also be directly excited around 488 nm, the wavelength that excites the potential donor GFP. CFP and YFP have been proven to be a good FRET pair at the ensemble level [79], but CFP is not detectable at the single-molecule level due to the high autofluorescence background in that wavelength range [76]. Developing a bright far-red FP to pair with GFP or YFP will significantly enhance current research capacity for probing protein–protein interactions in live cells at the single-molecule level.

3.8.4. Protein Complex Composition

In addition to monitoring gene expression and molecular interactions, the multimeric state and composition of a protein complex in live cells can be measured using single-molecule imaging. Ulbrich and Isacoff reported a single-molecule assay to count the copy numbers of subunits of proteins in live cell membranes by taking advantage of the discrete steps of photobleaching of single molecules [200]. This method was used to determine the numbers of subunits NR1 and NR3 in the glutamate-gated receptor *N*-methyl-D-aspartate (NMDA) [211]. In another study, the copy numbers of seven kinetochore proteins tagged with GFP in budding yeast were determined, not by discrete photobleaching, but by the integrated fluorescence intensity normalized against the known single-GFP-molecule intensity [212]. The limitation of these methods is that the copy numbers cannot exceed more than a few molecules per spot because the higher the copy number, the more likely it is that two molecules are photobleached at the same time and the greater is the error in using integrated intensity to determine the total copy number of molecules due to the stochastic nature of photobleaching.

Recent advances in single-molecule superresolution imaging have enabled the application of these highly promising techniques in live cells [123,137,158,213]. These techniques not only allow the quantification of the composition of a protein complex, but also provide nanometer resolution of the spatial organization of the complex. Their potentials cannot be

underestimated. Because they are discussed in detail in Chapter 10, they are not included here, and readers are encouraged to see the related information and references there.

3.9. Outlook

Recent advances in single-molecule imaging techniques and FP development have enabled the direct visualization of cellular processes and molecular interactions as they occur in real time inside live cells. This opens a wide window for investigating many fundamental biological questions at a level that was inaccessible by population studies before. Yet this field is still young. Development of better FPs and more powerful imaging techniques, such as spectrally separable FP pairs for sm-FRET imaging in live cells, fast-maturing FPs for correlating the expression of multiple genes, and multicolor superresolution imaging in live cells, promises to bring in even more exciting discoveries. With the continuing efforts in pushing current technical limits, in the near future it may no longer be a fantasy for a biologist to directly visualize the dynamics of individual molecules in multiplex with high temporal and spatial resolution in various live cells, tissues, and even animals, discovering unprecedented molecular details at a systems level.

References

1. Moerner, W. E., and L. Kador (1989). Optical detection and spectroscopy of single molecules in a solid. Phys Rev Lett, **62**(21): 2535–8.
2. Orrit, M., and J. Bernard (1990). Single pentacene molecules detected by fluorescence excitation in a *p*-terphenyl crystal. Phys Rev Lett, **65**(21): 2716–19.
3. Shav-Tal, Y., X. Darzacq, et al. (2004). Dynamics of single mRNPs in nuclei of living cells. Science, **304**(5678): 1797–800.
4. Golding, I., J. Paulsson, et al. (2005). Real-time kinetics of gene activity in individual bacteria. Cell, **123**(6): 1025–36.
5. Fusco, D., N. Accornero, et al. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Curr Biol, **13**(2): 161–7.
6. Bates, D., and N. Kleckner (2005). Chromosome and replisome dynamics in E. coli: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. Cell, **121**(6): 899–911.
7. Lemon, K. P., and A. D. Grossman (2000). Movement of replicating DNA through a stationary replisome. Mol Cell, **6**(6): 1321–30.
8. Viollier, P. H., M. Thanbichler, et al. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. Proc Natl Acad Sci USA, **101**(25): 9257–62.
9. Rotman, B. (1961). Measurement of activity of single molecules of beta-D-galactosidase. Proc Natl Acad Sci USA, **47**: 1981–91.
10. Cai, L., N. Friedman, and X. S. Xie (2006). Stochastic protein expression in individual cells at the single molecule level. Nature, **440**(7082): 358–62.
11. Ghaemmaghami, S., W. K. Huh, et al. (2003). Global analysis of protein expression in yeast. Nature, **425**(6959): 737–41.
12. Guptasarma, P. (1995). Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia *coli*? Bioessays, **17**(11): 987–97.
13. Freund, J. A., and T. Poschel (2005). *Stochastic Processes in Physics, Chemistry, and Biology*. Springer: Berlin.
14. Elowitz, M. B., A. J. Levine, et al. (2002). Stochastic gene expression in a single cell. Science, **297**(5584): 1183–6.
15. Ozbudak, E. M., M. Thattai, et al. (2002). Regulation of noise in the expression of a single gene. Nat Genet, **31**(1): 69–73.

16. Weinberger, L. S., J. C. Burnett, et al. (2005). Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell, **122**(2): 169–82.
17. Huang, C. Y., and J. E. Ferrell, Jr. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci USA, **93**(19): 10078–83.
18. Braun, H. A., H. Wissing, et al. (1994). Oscillation and noise determine signal transduction in shark multimodal sensory cells. Nature, **367**(6460): 270–3.
19. Wang, W., and Z. D. Wang (1997). Internal-noise-enhanced signal transduction in neuronal systems. Phys Rev E, **55**(6): 7379–84.
20. Bezrukov, S. M., and I. Vodyanoy (1995). Noise-induced enhancement of signal transduction across voltage-dependent ion channels. Nature, **378**(6555): 362–4.
21. Arkin, A., J. Ross, and H. H. McAdams (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. Genetics, **149**(4): 1633–48.
22. Suel, G. M., J. Garcia-Ojalvo, et al. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. Nature, **440**(7083): 545–50.
23. Choi, P. J., L. Cai, et al. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. Science, **322**(5900): 442–6.
24. Maamar, H., A. Raj, and D. Dubnau (2007). Noise in gene expression determines cell fate in *Bacillus subtilis.* Science, **317**(5837): 526–9.
25. Kussell, E., R. Kishony, et al. (2005). Bacterial persistence: a model of survival in changing environments. Genetics, **169**(4): 1807–14.
26. Berland, K. M., P. T. So, and E. Gratton (1995). Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment. Biophys J, **68**(2): 694–701.
27. Brock, R., M. A. Hink, and T. M. Jovin (1998). Fluorescence correlation microscopy of cells in the presence of autofluorescence. Biophys J, **75**(5): 2547–57.
28. Schwille, P., U. Haupts, et al. (1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. Biophys J, **77**(4): 2251–65.
29. Komarova, Y., J. Peloquin, and G. Borisy (2005). Microinjection of fluorophore-labeled proteins. In *Live Cell Imaging, a Laboratory Manual*, R. D. Goldman and D. L. Spector, Editors. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, pp. 67–86.
30. O'Hare, H. M., K. Johnsson, and A. Gautier (2007). Chemical probes shed light on protein function. Curr Opin Struct Biol, **17**(4): 488–94.
31. Sako, Y., S. Minoghchi, and T. Yanagida (2000). Single-molecule imaging of EGFR signalling on the surface of living cells. Nat Cell Biol, **2**(3): 168–72.
32. Schütz, G. J., G. Kada, et al. (2000). Properties of lipid microdomains in a muscle cell membrane visualized by single molecule microscopy. EMBO J **19**: 892–901.
33. Murakoshi, H., R. Iino, et al. (2004). Single-molecule imaging analysis of Ras activation in living cells. Proc Natl Acad Sci USA, **101**(19): 7317–22.
34. Guignet, E. G., J. M. Segura, et al. (2007). Repetitive reversible labeling of proteins at polyhistidine sequences for single-molecule imaging in live cells. Chemphyschem, **8**(8): 1221–7.
35. Lord, S. J., N. R. Conley, et al. (2008). A photoactivatable push-pull fluorophore for single-molecule imaging in live cells. J Am Chem Soc. (2008) Jul 23; **130**(29): 9204–5.
36. Lord, S. J., Z. Lu, et al. (2007). Photophysical properties of acene DCDHF fluorophores: long-wavelength single-molecule emitters designed for cellular imaging. J Phys Chem A, **111**(37): 8934–41.
37. Michalet, X., F. F. Pinaud, et al. (2005). Quantum dots for live cells, *in vivo* imaging, and diagnostics. Science, **307**(5709): 538–44.
38. Zhou, M., and I. Ghosh (2007). Quantum dots and peptides: a bright future together. Biopolymers, **88**(3): 325–39.
39. Davenport, D., and J. A. C. Nicol (1955). Luminescence of Hydromedusae. Proc R Soc B, **144**: 399–411.
40. Shimomura, O., F. H. Johnson, and Y. Saiga (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea.* J Cell Comp Physiol, **59**: 223–39.
41. Hastings, J. W., and J. G. Morin (1969). Comparative biochemistry of calcium photoproteins from the ctenophore, *Mnemiopsis* and the coelenterates *Aequorea*, *Obelia*, *Pelagia* and *Renilla.* Biol Bull, **137**: 402.
42. Morin, J. G., and J. W. Hastings (1971). Energy transfer in a bioluminescent system. J Cell Physiol, **77**(3): 313–8.
43. Morise, H., O. Shimomura, et al. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea.* Biochemistry, **13**(12): 2656–62.

44. Mills, C. E. (2001). Jellyfish blooms: are populations increasing globally in response to changing ocean conditions? Hydrobiologia, **451**: 55–68.
45. Prasher, D. C., V. K. Eckenrode, et al. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. Gene, **111**(2): 229–33.
46. Chalfie, M., Y. Tu, et al. (1994). Green fluorescent protein as a marker for gene expression. Science, **263**(5148): 802–5.
47. Heim, R., D. C. Prasher, and R. Y. Tsien (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proc Natl Acad Sci USA, **91**(26): 12501–4.
48. Inouye, S., and F. I. Tsuji (1994). *Aequorea* green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. FEBS Lett, **341**(2–3): 277–80.
49. Heim, R., A. B. Cubitt, and R. Y. Tsien (1995). Improved green fluorescence. Nature, **373**(6516): 663–4.
50. Yang, F., L. G. Moss, and G. N. Phillips, Jr. (1996). The molecular structure of green fluorescent protein. Nat Biotechnol, **14**(10): 1246–51.
51. Ormo, M., A. B. Cubitt, et al. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. Science, **273**(5280): 1392–5.
52. Cubitt, A. B., F. Carrel, et al. (1992). Molecular genetic analysis of signal transduction pathways controlling multicellular development in *Dictyostelium.* Cold Spring Harb Symp Quant Biol, **57**: 177–92.
53. Reid, B. G., and G. C. Flynn (1997). Chromophore formation in green fluorescent protein. Biochemistry, **36**(22): 6786–91.
54. Rosenow, M. A., H. A. Huffman, et al. (2004). The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. Biochemistry, **43**(15): 4464–72.
55. Rosenow, M. A., H. N. Patel, and R. M. Wachter (2005). Oxidative chemistry in the GFP active site leads to covalent cross-linking of a modified leucine side chain with a histidine imidazole: implications for the mechanism of chromophore formation. Biochemistry, **44**(23): 8303–11.
56. Zhang, L., H. N. Patel, et al. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. J Am Chem Soc, **128**(14): 4766–72.
57. Barondeau, D. P., C. J. Kassmann, et al. (2005). Understanding GFP chromophore biosynthesis: controlling backbone cyclization and modifying post-translational chemistry. Biochemistry, **44**(6): 1960–70.
58. Ward, W. W., H. J. Prentice, et al. (1982). Spectral perturbations of the *Aequorea* green fluorescent protein. Photochem Photobiol, **35**: 803–8.
59. Voityuk, A. A., A. D. Kummer, et al. (2001). Absorption spectra of the GFP chromophore in solution: comparison of theoretical and experimental results. Chem Phys, **269**(1–3): 83–91.
60. Chattoraj, M., B. A. King, et al. (1996). Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. Proc Natl Acad Sci USA, **93**(16): 8362–7.
61. van Thor, J. J., T. Gensch, et al. (2002). Phototransformation of green fluorescent protein with UV and visible light leads to decarboxylation of glutamate 222. Nat Struct Biol, **9**(1): 37–41.
62. Cubitt, A. B., R. Heim, et al. (1995). Understanding, improving and using green fluorescent proteins. Trends Biochem Sci, **20**(11): 448–55.
63. Cubitt, A. B., L. A. Woollenweber, and R. Heim (1999). Understanding structure–function relationships in the *Aequorea victoria* green fluorescent protein. Methods Cell Biol, **58**: 19–30.
64. Remington, S. J. (2006). Fluorescent proteins: maturation, photochemistry and photophysics. Curr Opin Struct Biol, **16**(6): 714–21.
65. McAnaney, T. B., W. Zeng, et al. (2005). Protonation, photobleaching, and photoactivation of yellow fluorescent protein (YFP 10C): a unifying mechanism. Biochemistry, **44**(14): 5510–24.
66. Shaner, N. C., M. Z. Lin, et al. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. Nat Methods, **5**(6): 545–51.
67. Swaminathan, R., C. P. Hoang, and A. S. Verkman (1997). Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. Biophys J, **72**(4): 1900–7.
68. Peterman, E. J. G., S. Brasselet, and W. E. Moerner (1999). The fluorescence dynamics of single molecules of green fluorescent protein. J Phys Chem A, **103**(49): 10553–60.
69. Pierce, D. W., N. Hom-Booher, and R. D. Vale (1997). Imaging individual green fluorescent proteins. Nature, **388**(6640): 338.
70. Greenbaum, L., C. Rothmann, et al. (2000). Green fluorescent protein photobleaching: a model for protein damage by endogenous and exogenous singlet oxygen. Biol Chem, **381**(12): 1251–58.

71. Bulina, M. E., D. M. Chudakov, et al. (2006). A genetically encoded photosensitizer. Nat Biotechnol, **24**(1): 95–9.
72. Lakowicz, J. (2006). *Principles of Fluorescence Spectroscopy*, 3rd ed. Springer: New York.
73. Heikal, A. A., S. T. Hess, et al. (2000). Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: coral red (dsRed) and yellow (Citrine). Proc Natl Acad Sci USA, **97**(22): 11996–2001.
74. Garcia-Parajo, M. F., G.M. Segers-Nolten, et al. (2000). Real-time light-driven dynamics of the fluorescence emission in single green fluorescent protein molecules. Proc Natl Acad Sci USA, **97**(13): 7237–42.
75. Shaner, N. C., P. A. Steinbach, and R. Y. Tsien (2005). A guide to choosing fluorescent proteins. Nat Methods, **2**(12): 905–9.
76. Harms, G. S., L. Cognet, et al. (2001). Autofluorescent proteins in single-molecule research: applications to live cell imaging microscopy. Biophys J, **80**(5): 2396–408.
77. Dickson, R. M., A. B. Cubitt, et al. (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. Nature, **388**(6640): 355–8.
78. Moerner, W. E., E. J. Peterman, et al. (1999). Optical methods for exploring dynamics of single copies of green fluorescent protein. Cytometry, **36**(3): 232–8.
79. Miyawaki, A., and R. Y. Tsien (2000). Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. Methods Enzymol, **327**: 472–500.
80. Nifosi, R., A. Ferrari, et al. (2003). Photoreversible dark state in a tristable green fluorescent protein variant. J Phys Chem B, **107**(7): 1679–84.
81. Haupts, U., S. Maiti, et al. (1998). Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. Proc Natl Acad Sci USA, **95**(23): 13573.
82. Evdokimov, A. G., M. E. Pokross, et al. (2006). Structural basis for the fast maturation of Arthropoda green fluorescent protein. EMBO Rep, **7**(10): 1006–12.
83. Nagai, T., K. Ibata, et al. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol, **20**(1): 87–90.
84. Yu, J., J. Xiao, et al. (2006). Probing gene expression in live cells, one protein molecule at a time. Science, **311**(5767): 1600–3.
85. Shaner, N. C., R. E. Campbell, et al. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. Nat Biotechnol, **22**(12): 1567–72.
86. Tsien, R. Y. (1998). The green fluorescent protein. Annu Rev Biochem, **67**: 509–44.
87. Kennis, J. T., D. S. Larsen, et al. (2004). Uncovering the hidden ground state of green fluorescent protein. Proc Natl Acad Sci USA, **101**(52): 17988–93.
88. Cormack, B. P., G. Bertram, et al. (1997). Yeast-enhanced green fluorescent protein (yEGFP)a reporter of gene expression in *Candida albicans.* Microbiology, **143**(Pt 2): 303–11.
89. Brejc, K., T. K. Sixma, et al. (1997). Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. Proc Natl Acad Sci USA, **94**(6): 2306–11.
90. Wachter, R. M., D. Yarbrough, et al. (2000). Crystallographic and energetic analysis of binding of selected anions to the yellow variants of green fluorescent protein. J Mol Biol, **301**(1): 157–71.
91. Iino, R., I. Koyama, and A. Kusumi (2001). Single molecule imaging of green fluorescent proteins in living cells: E-cadherin forms oligomers on the free cell surface. Biophys J, **80**(6): 2667–77.
92. Watanabe, N., and T. J. Mitchison (2002). Single-molecule speckle analysis of actin filament turnover in Lamellipodia. Science, **295**(5557): 1083–6.
93. Pedelacq, J. D., S. Cabantous, et al. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol, **24**(1): 79–88.
94. Deich, J., E. M. Judd, et al. (2004). Visualization of the movement of single histidine kinase molecules in live *Caulobacter* cells. Proc Natl Acad Sci USA, **101**(45): 15921–6.
95. Harms, G. S., L. Cognet, et al. (2001). Single-molecule imaging of l-type Ca(2+) channels in live cells. Biophys J, **81**(5): 2639–46.
96. Kim, S. Y., Z. Gitai, et al. (2006). Single molecules of the bacterial actin MreB undergo directed treadmilling motion in *Caulobacter crescentus.* Proc Natl Acad Sci USA, **103**(29): 10929–34.
97. Chesler, M., and K. Kaila (1992). Modulation of pH by neuronal activity. Trends Neurosci, **15**(10): 396–402.
98. Mansoura, M. K., J. Biwersi, et al. (1999). Fluorescent chloride indicators to assess the efficacy of CFTR cDNA delivery. Hum Gene Ther, **10**(6): 861–75.
99. Zacharias, D. A., J. D. Violin, et al. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science, **296**(5569): 913–6.
100. Griesbeck, O., G. S. Baird, et al. (2001). Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. J Biol Chem, **276**(31): 29188–94.

101. Shaner, N. C., G. H. Patterson, and M. W. Davidson (2007). Advances in fluorescent protein technology. J Cell Sci, **120**(Pt 24): 4247–60.
102. Rekas, A., J. R. Alattia, et al. (2002). Crystal structure of Venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity. J Biol Chem, **277**(52): 50573–8.
103. Nguyen, A. W., and P. S. Daugherty (2005). Evolutionary optimization of fluorescent proteins for intracellular FRET. Nat Biotechnol, **23**(3): 355–60.
104. Ohashi, T., S. D. Galiacy, et al. (2007). An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins. Protein Sci, **16**(7): 1429–38.
105. Vinkenborg, J. L., T. H. Evers, et al. (2007). Enhanced sensitivity of FRET-based protease sensors by redesign of the GFP dimerization interface. Chembiochem, **8**(10): 1119–21.
106. Mishin, A. S., F. V. Subach, et al. (2008). The first mutant of the *Aequorea victoria* green fluorescent protein that forms a red chromophore. Biochemistry, **47**(16): 4666–73.
107. Elowitz, M. B., M. G. Surette, et al. (1997). Photoactivation turns green fluorescent protein red. Curr Biol, **7**(10): 809–12.
108. Sawin, K. E., and P. Nurse (1997). Photoactivation of green fluorescent protein. Curr Biol, **7**(10): R606–7.
109. Matz, M. V., A. F. Fradkov, et al. (1999). Fluorescent proteins from nonbioluminescent *Anthozoa* species. Nat Biotechnol, **17**(10): 969–73.
110. Baird, G. S., D. A. Zacharias, and R. Y. Tsien (2000). Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. Proc Natl Acad Sci USA, **97**(22): 11984–9.
111. Gross, L. A., G. S. Baird, et al. (2000). The structure of the chromophore within DsRed, a red fluorescent protein from coral. Proc Natl Acad Sci USA, **97**(22): 11990–5.
112. Wall, M. A., M. Socolich, and R. Ranganathan (2000). The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. Nat Struct Biol, **7**(12): 1133–8.
113. Yarbrough, D., R. M. Wachter, et al. (2001). Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution. Proc Natl Acad Sci USA, **98**(2): 462–7.
114. Tubbs, J. L., J. A. Tainer, and E. D. Getzoff (2005). Crystallographic structures of Discosoma red fluorescent protein with immature and mature chromophores: linking peptide bond trans-cis isomerization and acylimine formation in chromophore maturation. Biochemistry, **44**(29): 9833–40.
115. Shu, X., K. Kallio, et al. (2007). Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies. Biochemistry, **46**(43): 12005–13.
116. Xie, X. S., P. J. Choi, et al. (2008). Single-molecule approach to molecular biology in living bacterial cells. Annu Rev Biophys, **37**: 417–44.
117. Karasawa, S., T. Araki, et al. (2004). Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. Biochem J, **381**(Pt 1): 307–12.
118. Shu, X., N. C. Shaner, et al. (2006). Novel chromophores and buried charges control color in mFruits. Biochemistry, **45**(32): 9639–47.
119. Vavylonis, D., J. Q. Wu, et al. (2008). Assembly mechanism of the contractile ring for cytokinesis by fission yeast. Science, **319**(5859): 97–100.
120. Merzlyak, E. M., J. Goedhart, et al. (2007). Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nat Methods, **4**(7): 555–7.
121. Petersen, J., P. G. Wilmann, et al. (2003). The 2.0-A crystal structure of eqFP611, a far red fluorescent protein from the sea anemone *Entacmaea quadricolor.* J Biol Chem, **278**(45): 44626–31.
122. Shcherbo, D., E. M. Merzlyak, et al. (2007). Bright far-red fluorescent protein for whole-body imaging. Nat Methods, **4**(9): 741–6.
123. Manley, S., J. M. Gillette, et al. (2008). High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat Methods. 2008 Feb; **5**(2): 155–7.
124. Niu, L., and J. Yu (2008). Investigating intracellular dynamics of FtsZ cytoskeleton with photoactivation single-molecule tracking. Biophys J, **95**(4): 2009–16.
125. Verkhusha, V. V., and A. Sorkin (2005). Conversion of the monomeric red fluorescent protein into a photoactivatable probe. Chem Biol, **12**(3): 279–85.
126. Lukyanov, K. A., A. F. Fradkov, et al. (2000). Natural animal coloration can Be determined by a nonfluorescent green fluorescent protein homolog. J Biol Chem, **275**(34): 25879–82.
127. Ando, R., H. Mizuno, and A. Miyawaki (2004). Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. Science, **306**(5700): 1370–3.
128. Ando, R., H. Hama, et al. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. Proc Natl Acad Sci USA, **99**(20): 12651–6.

129. Tsutsui, H., S. Karasawa, et al. (2005). Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. EMBO Rep, **6**(3): 233–8.
130. Wiedenmann, J., S. Ivanchenko, et al. (2004). EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc Natl Acad Sci USA, **101**(45): 15905–10.
131. Gurskaya, N. G., V. V. Verkhusha, et al. (2006). Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat Biotechnol, **24**(4): 461–5.
132. Chudakov, D. M., V. V. Verkhusha, et al. (2004). Photoswitchable cyan fluorescent protein for protein tracking. Nat Biotechnol, **22**(11): 1435–9.
133. Lukyanov, K. A., D. M. Chudakov, et al. (2005). Innovation: Photoactivatable fluorescent proteins. Nat Rev Mol Cell Biol, **6**(11): 885–91.
134. Patterson, G. H. (2008). Photoactivation and imaging of photoactivatable fluorescent proteins. Curr Protoc Cell Biol, **2008**(March): Chapter 21: Unit 21.6.
135. Lippincott-Schwartz, J., and G. H. Patterson (2008). Fluorescent proteins for photoactivation experiments. Methods Cell Biol, **85**: 45–61.
136. Cinelli, R. A., V. Tozzini, et al. (2001). Coherent dynamics of photoexcited green fluorescent proteins. Phys Rev Lett, **86**(15): 3439–42.
137. Biteen, J. S., M. A. Thompson, et al. (2008). Super-resolution imaging in live *Caulobacter crescentus* cells using photoswitchable EYFP. Nat Methods, **5**(11): 947–9.
138. Habuchi, S., R. Ando, et al. (2005). Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. Proc Natl Acad Sci USA, **102**(27): 9511–6.
139. Andresen, M., A. C. Stiel, et al. (2007). Structural basis for reversible photoswitching in Dronpa. Proc Natl Acad Sci USA, **104**(32): 13005–9.
140. Stiel, A. C., S. Trowitzsch, et al. (2007). 1.8 A bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. Biochem J, **402**(1): 35–42.
141. Wilmann, P. G., K. Turcic, et al. (2006). The 1.7 A crystal structure of Dronpa: a photoswitchable green fluorescent protein. J Mol Biol, **364**(2): 213–24.
142. Dedecker, P., J. Hotta, et al. (2006). Fast and reversible photoswitching of the fluorescent protein Dronpa as evidenced by fluorescence correlation spectroscopy. Biophys J, **91**(5): L45–7.
143. Fron, E., C. Flors, et al. (2007). Ultrafast excited-state dynamics of the photoswitchable protein Dronpa. J Am Chem Soc, **129**(16): 4870–1.
144. Mizuno, H., T. K. Mal, et al. (2008). Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. Proc Natl Acad Sci USA, **105**(27): 9227–32.
145. Habuchi, S., P. Dedecker, et al. (2006). Photo-induced protonation/deprotonation in the GFP-like fluorescent protein Dronpa: mechanism responsible for the reversible photoswitching. Photochem Photobiol Sci, **5**(6): 567–76.
146. Chisari, M., D. K. Saini, et al. (2007). Shuttling of G protein subunits between the plasma membrane and intracellular membranes. J Biol Chem, **282**(33): 24092–8.
147. Wiegert, J. S., C. P. Bengtson, and H. Bading (2007). Diffusion and not active transport underlies and limits ERK1/2 synapse-to-nucleus signaling in hippocampal neurons. J Biol Chem, **282**(40): 29621–33.
148. Aramaki, S., and K. Hatta (2006). Visualizing neurons one-by-one *in vivo*: optical dissection and reconstruction of neural networks with reversible fluorescent proteins. Dev Dyn, **235**(8): 2192–9.
149. Dedecker, P., J. Hotta, et al. (2007). Subdiffraction imaging through the selective donut-mode depletion of thermally stable photoswitchable fluorophores: numerical analysis and application to the fluorescent protein Dronpa. J Am Chem Soc, **129**(51): 16132–41.
150. Betzig, E., G. H. Patterson, et al. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science, **313**(5793): 1642–5.
151. Eggeling, C., M. Hilbert, et al. (2007). Reversible photoswitching enables single-molecule fluorescence fluctuation spectroscopy at high molecular concentration. Microsc Res Tech. 2007 DEC; **70**(12): 1003–9.
152. Egner, A., C. Geisler, et al. (2007). Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. Biophys J. NOV 1; **93**(9): 3285–90.
153. Stiel, A. C., M. Andresen, et al. (2008). Generation of monomeric reversibly switchable red fluorescent proteins for far-field fluorescence nanoscopy. Biophys J. SEP 15; **95**(6): 2989–97.
154. Campbell, R. E., O. Tour, et al. (2002). A monomeric red fluorescent protein. Proc Natl Acad Sci USA, **99**(12): 7877–82.
155. Nienhaus, G. U., K. Nienhaus, et al. (2006). Photoconvertible fluorescent protein EosFP: biophysical properties and cell biology applications. Photochem Photobiol, **82**(2): 351–8.

156. Nienhaus, K., G. U. Nienhaus, et al. (2005). Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. Proc Natl Acad Sci USA, **102**(26): 9156–9.
157. Mizuno, H., T. K. Mal, et al. (2003). Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. Mol Cell, **12**(4): 1051–8.
158. Shroff, H., C. G. Galbraith, et al. (2007). Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. Proc Natl Acad Sci USA, **104**(51): 20308–13.
159. Niu, L., and J. Yu (2008). Investigating intracellular dynamics of FtsZ cytoskeleton with photo-activation single-molecule tracking. Biophys J, **95**(4): 2009–16, 15 Aug 2008.
160. Aubin, J. E. (1979). Autofluorescence of viable cultured mammalian cells. J Histochem Cytochem, **27**(1): 36–43.
161. Chance, B., and B. Thorell (1959). Localization and kinetics of reduced pyridine nucleotide in living cells by microfluorometry. J Biol Chem, **234**: 3044–50.
162. Benson, R. C., R. A. Meyer, et al. (1979). Cellular autofluorescence—is it due to flavins? J Histochem Cytochem, **27**(1): 44–8.
163. Xiao, J., J. Elf, et al. (2007). Imaging gene expression in living cells at the single-molecule level. In *Single Molecules: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, pp. 149–69.
164. Lansford, R., G. Bearman, and S. E. Fraser (2001). Resolution of multiple green fluorescent protein color variants and dyes using two-photon microscopy and imaging spectroscopy. J Biomed Opt, **6**: 311.
165. Dickinson, M. E., G. Bearman, et al. (2001). Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy. Biotechniques, **31**(1272): 1274–6.
166. Connally, R., D. Veal, and J. Piper (2004). Flash lamp-excited time-resolved fluorescence microscope suppresses autofluorescence in water concentrates to deliver an 11-fold increase in signal-to-noise ratio. J Biomed Opt, **9**(4): 725–34.
167. Wilkerson, C. W., Jr., P. M. Goodwin, et al. (1993). Detection and lifetime measurement of single molecules in flowing sample streams by laser-induced fluorescence. Appl Phys Lett, 62(17): 2030–32.
168. Qin, J., Y. Fung, et al. (2004). Native fluorescence detection of flavin derivatives by microchip capillary electrophoresis with laser-induced fluorescence intensified charge-coupled device detection. J Chromatogr A, **1027**(1–2): 223–29.
169. Yang, H., G. Luo, et al. (2003). *Protein Conformational Dynamics Probed by Single-Molecule Electron Transfer*. American Association for the Advancement of Science: Washington, DC: pp. 262–6.
170. Xie, X. S., J. Yu, and W. Y. Yang (2006). Living cells as test tubes. Science, **312**(5771): 228–30.
171. Straight, A. F., A. S. Belmont, et al. (1996). GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr Biol, **6**(12): 1599–1608.
172. Lau, I. F., S. R. Filipe, et al. (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. Mol Microbiol, **49**(3): 731–43.
173. Bertrand, E., P. Chartrand, et al. (1998). Localization of ASH1 mRNA particles in living yeast. Mol Cell, **2**(4): 437–45.
174. Averbeck, D., and S. Averbeck (1998). DNA photodamage, repair, gene induction and genotoxicity following exposures to 254 nm UV and 8-methoxypsoralen plus UVA in a eukaryotic cell system. Photochem Photobiol, **68**(3): 289–95.
175. Pfeifer, G. P., R. Drouin, et al. (1992). Binding of transcription factors creates hot spots for UV photoproducts *in vivo*. Mol Cell Biol, **12**(4): 1798.
176. Jones, C. A., E. Huberman, et al. (1987). Mutagenesis and cytotoxicity in human epithelial cells by far-and near-ultraviolet radiations: action spectra. Radiat Res, **110**(2): 244–54.
177. Mohanty, S. K., M. Sharma, and P. K. Gupta (2006). Generation of ROS in cells on exposure to CW and pulsed near-infrared laser tweezers. Photochem Photobiol Sci, **5**(1): 134–9.
178. Neuman, K. C., E. H. Chadd, et al. (1999). Characterization of photodamage to *Escherichia coli* in optical traps. Biophys J, **77**(5): 2856–63.
179. Liu, Y., G. J. Sonek, et al. (1996). Physiological monitoring of optically trapped cells: assessing the effects of confinement by 1064-nm laser tweezers using microfluorometry. Biophys J, **71**(4): 2158.
180. Ashkin, A., and J. M. Dziedzic (1989). Internal cell manipulation using infrared laser traps. Proc Natl Acad Sci USA, **86**(20): 7914–18.
181. Moan, J., K. Berg, et al. (1989). Intracellular localization of photosensitizers. In *Photosensitizing Compounds: Their Chemistry, Biology and Clinical Use*, G. Bock and S. Harnett, Editors. John Wiley and Sons: New York, pp. 95–111.

182. Bensasson, R. V., E. J. Land, and T. G. Truscott (1993). *Excited States and Free Radicals in Biology and Medicine*. Oxford University Press: Oxford.
183. Dixit, R., and R. Cyr (2003). Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy. Plant J, **36**(2): 280–90.
184. Halliwell, B., and J. M. C. Gutteridge (1989). Protection against oxidants in biological systems: the superoxide theory of oxygen toxicity. Free Rad Biol Med. Edited by Halliwell B, Gutteridge JMC, Oxford Clarendun Press pp 86–179.
185. Schafer, F. Q., and G. R. Buettner (2001). Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. Free Rad Biol Med, **30**(11): 1191–1212.
186. de With, A., and K. O. Greulich (1995). Wavelength dependence of laser-induced DNA damage in lymphocytes observed by single-cell gel electrophoresis. J Photochem Photobiol B Biol, **30**(1): 71–76.
187. Tripathi, A., R. E. Jabbour, et al. (2008). Waterborne pathogen detection using Raman spectroscopy. Appl Spectrosc, **62**: 1–9.
188. Benson, D. M., J. Bryan, et al. (1985). Digital imaging fluorescence microscopy: spatial heterogeneity of photobleaching rate constants in individual cells. J Cell Biol, **100**(4): 1309–23.
189. Adachi, K., K. Kinosita, Jr., and T. Ando (1999). Single-fluorophore imaging with an unmodified epifluorescence microscope and conventional video camera. J Microsc, **195**(2): 125–32.
190. Gratton, E., and M. J. vandeVen (1995). *Laser sources for confocal microscopy*, in *Handbook of Biological Confocal Microscopy*, J. B. Pawley, Editor. Plenum Press: New York, pp. 65–98.
191. Murphy, D. B. (2002). *Fundamentals of Light Microscopy and Electronic Imaging*. Wiley-Liss: New York.
192. Nipkow, P. (1884). German Patent 30,105.
193. Inoue, S., and K. R. Spring (1997). *Video Microscopy*. Plenum Press: New York.
194. Stuurman, N., and R. D. Vale (2006). Imaging single molecules using total internal reflection fluorescence microscopy. In *Live Cell Imaging, A Laboratory Manual*, R.D. Goldman and D.L. Spector, Editors. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, pp. 585–601.
195. Swedlow, J. R., P. D. Andrews, and M. Platani (2005). *In vivo* imaging of mammalian cells. In *Live Cell Imaging: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, pp. 329–43.
196. Elf, J., G. W. Li, and X. S. Xie (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. Science, **316**(5828): 1191–4.
197. Ueda, M., Y. Sako, et al. (2001). Single-molecule analysis of chemotactic signaling in *Dictyostelium* cells. Science, **294**(5543): 864–7.
198. Haggie, P. M., and A. S. Verkman (2008). Monomeric CFTR in plasma membranes in live cells revealed by single-molecule fluorescence imaging. J Biol Chem. AUG 2008; **283**: 23510–23513.
199. Fu, G., C. Wang, et al. (2008). Heterodimerization of integrin Mac-1 subunits studied by single-molecule imaging. Biochem Biophys Res Commun, **368**(4): 882–6.
200. Ulbrich, M. H., and E. Y. Isacoff (2007). Subunit counting in membrane-bound proteins. Nat Methods, **4**(4): 319–21.
201. Groc, L., M. Heine, et al. (2004). Differential activity-dependent regulation of the lateral mobilities of AMPA and NMDA receptors. Nat Neurosci, **7**(7): 695–6.
202. Lommerse, P. H., G. A. Blab, et al. (2004). Single-molecule imaging of the H-ras membrane-anchor reveals domains in the cytoplasmic leaflet of the cell membrane. Biophys J, **86**(1 Pt 1): 609–16.
203. Lommerse, P. H., B. E. Snaar-Jagalska, et al. (2005). Single-molecule diffusion measurements of H-Ras at the plasma membrane of live cells reveal microdomain localization upon activation. J Cell Sci, **118**(Pt 9): 1799–809.
204. Sonnleitner, A., L. M. Mannuzzu, et al. (2002). Structural rearrangements in single ion channels detected optically in living cells. Proc Natl Acad Sci USA, **99**(20): 12759.
205. Cai, D., K. J. Verhey, and E. Meyhofer (2007). Tracking single Kinesin molecules in the cytoplasm of mammalian cells. Biophys J, **92**(12): 4137–44.
206. Holland, M. J. (2002). Transcript abundance in yeast varies over six orders of magnitude. J Biol Chem, **277**(17): 14363–6.
207. Velculescu, V. E., L. Zhang, et al. (1997). Characterization of the yeast transcriptome. Cell, **88**(2): 243–51.
208. Novick, A., and M. Weiner (1957). Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci USA, **43**(7): 553–66.
209. Blake, W. J., M. Kærn, et al. (2003). Noise in eukaryotic gene expression. Nature, **422**(6932): 633–7.
210. Raser, J.M., and E. K. O'Shea (2004). Control of stochasticity in eukaryotic gene expression. Science, **304**(5678): 1811–4.

211. Kohout, S. C., M. H. Ulbrich, et al. (2008). Subunit organization and functional transitions in Ci-VSP. Nat Struct Mol Biol, **15**(1): 106–8.
212. Joglekar, A. P., D. C. Bouck, et al. (2006). Molecular architecture of a kinetochore-microtubule attachment site. Nat Cell Biol, **8**(6): 581–5.
213. Hess, S. T., T. J. Gould, et al. (2007). Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. Proc Natl Acad Sci USA, **104**(44): 17370–5.
214. Patterson, G. H., et al. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys J, **73**(5): 2782–90.
215. Iwane, A. H., et al. (1997). Single molecular assay of individual ATP turnover by a myosin-GFP fusion protein expressed *in vitro*. FEBS Lett, **407**(2): 235–8.
216. Cognet, L., et al. (2002). Fluorescence microscopy of single autofluorescent proteins for cellular biology. C R Physique, **3**(5): 645–56.
217. Kremers, G. J., et al. (2007). Improved green and blue fluorescent proteins for expression in bacteria and mammalian cells. Biochemistry, **46**(12): 3775–83.
218. Hendrix, J., et al. (2008). Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. Biophys J, **94**(10): 4103.

4

Fluorescence Imaging at Sub-Diffraction-Limit Resolution with Stochastic Optical Reconstruction Microscopy

Graham T. Dempsey,* Wenqin Wang,* and Xiaowei Zhuang

Abstract Fluorescence microscopy is an essential tool in biological research. One major drawback of conventional light microscopy, however, is its relatively low resolution, which is limited by the diffraction of light to several hundreds of nanometers. In recent years, a number of fluorescence imaging techniques with sub-diffraction-limit resolution have been developed, achieving a spatial resolution of tens of nanometers in both the lateral and axial dimensions. This chapter focuses on one of these methods, stochastic optical reconstruction microscopy (STORM), which utilizes photoswitchable flourescent probes to separate spatially overlapping images of individual fluorophores in time and construct superresolution images from the precise positions of these fluorophores determined from the single-molecule images. Application of this technique has been extended to imaging fluorophores of different colors simultaneously, in three dimensions, and in living cells. This chapter describes the implementation of multicolor and three-dimensional STORM to imaging cellular structures. It begins by discussing the choice of photoswitchable fluorescent probe and the scheme with which to label a cellular target of interest. The instrumentation and methods for performing a STORM experiment are then described, followed by an outline of the analysis routines used for creating a STORM image. Applications of the technique along with general protocols and troubleshooting are given at the conclusion of the chapter.

G. T. Dempsey • Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA
W. Wang • Department of Physics, Harvard University, Cambridge, MA 02138, USA
X. Zhuang • Departments of Physics, Chemistry and Chemical Biology, Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138, USA

*These authors contributed equally to this work.

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_4,

4.1. Introduction

Optical microscopy has long been limited by diffraction to a spatial resolution of approximately 200–300 nm in the lateral dimensions and 500–800 nm in the axial dimension. This limit leaves many biological structures too small to be examined in detail, considering that subcellular structures span a range of length scales from nanometers to microns. With the recent advent of superresolution imaging techniques, however, the classical resolution limit of far-field light microscopy has been surpassed, allowing biological processes to be probed with unprecedented levels of detail [1–6]. One such technique relies on the use of single-molecule imaging and photoswitchable fluorescent probes whose fluorescent state can be switched from a dark to a bright state using color-specific activation [3–5]. This methodology was developed by three groups and has been referred to as stochastic optical reconstruction microscopy (STORM), photoactivated localization microscopy (PALM), and fluorescence photoactivation localization microscopy (FPALM). Although all three terms describe the same imaging method, in this chapter we use the term STORM for simplicity. We describe the methodology of STORM and its application to both multicolor and three-dimensional (3D) cellular imaging.

4.1.1. Basic Principle of STORM

The basic principle of STORM is shown in Figure 4.1A [3]. In a typical experiment, a sample of interest, such as specific molecular structures in a cell, is densely labeled with photoswitchable probes. Only a sparse subset of these molecules is activated at any time, such that their images are nonoverlapping. The positions of these molecules are then determined by fitting the image of each fluorophore (which corresponds to the point spread function [PSF] of the microscope) to determine its centroid position. The localization precision is limited by the number of photons collected before returning to the dark state, along with the pixel size and the background fluorescence in the image [7–16]. This process is repeated, and each iteration switches on and localizes the positions of a statistically different set of fluorophores until the structure of interest is sufficiently sampled. The final STORM image is then constructed by plotting the localizations obtained from all switching cycles.

4.1.2. Multicolor STORM

STORM can also be extended to image multiple types of fluorescent probes with distinct colors simultaneously [17–19]. Multicolor STORM can be implemented by using different colored photoswitchable dyes (e.g., cyanine dyes [17,20], caged fluorescein [Invitrogen], Q-rhodamine [Invitrogen], photochromic rhodamines [21,22]) and/or fluorescent proteins (FPs) (e.g., PA-GFP [23], Dronpa [24], Kaede [25], mEosFP, d2EosFP [26], Dendra2 [27], PS-CFP2 [28], rsFastLime [29], KikGR [30]). These probes can be distinguished based on the wavelength of light that they emit. Table 4.1 gives a summary of selected photoswitchable dyes and FPs. It is also possible to distinguish different photoswitchable probes by the wavelength of light that activates these probes. In some cases, the activation and the emission color scheme can be combined to obtain even more colors. As a specific example, we discovered that red carbocyanine dyes, such as Cy5, Cy5.5, and Cy7 (from GE Healthcare), reversibly switch between a fluorescent and a dark state [17]. These dyes, referred to as reporters, can be switched to a dark state by excitation with a red laser. When paired with another fluorophore, which we refer to as an activator (e.g., Alexa 405, Cy2, and Cy3), these

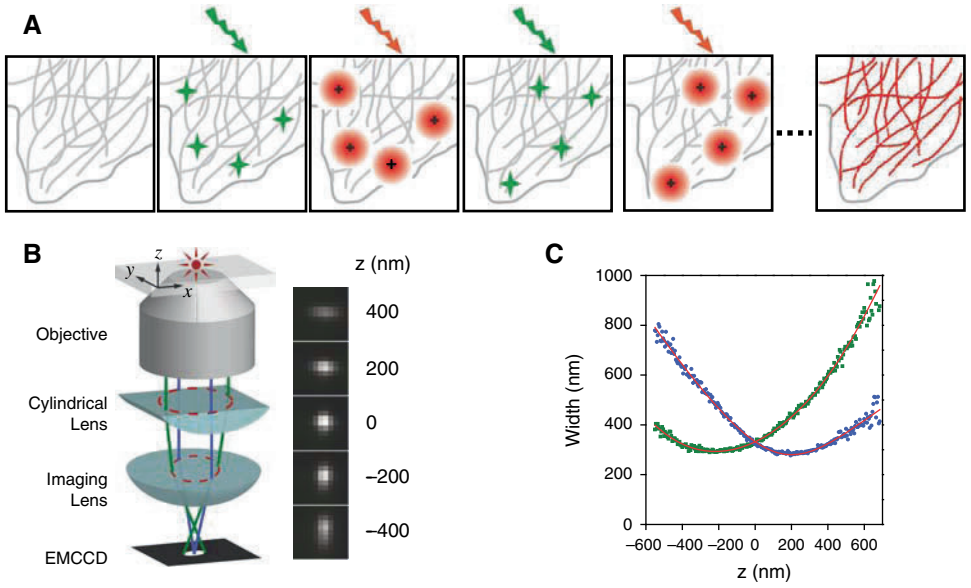


Figure 4.1. The principles of stochastic optical reconstruction microscopy (STORM). **A.** A region of a cell in which the structure of interest is represented as gray filaments that are labeled densely with photoswitchable probes (not shown). The fluorophores are initially in or switched to a nonfluorescent state. In the next panel, an activation pulse switches on a small subset of fluorophores (*green stars*) such that their images are resolvable from each other. Fluorescence images from these molecules (*red circles*) are fit to determine the centroid position of each peak, shown as a black cross. This process is repeated, and a new subset of molecules is switched on and localized. After multiple iterations, a high-resolution map of the cellular structure is constructed by plotting the positions for all fluorophores (*red dots*). The spatial resolution of the STORM image is not limited by diffraction, but by the accuracy of each localization. **B.** The use of optical astigmatism introduced in the imaging pathway enables high-resolution imaging of molecules in three dimensions. A cylindrical lens will cause elongation of the image in either the x or y direction, depending on the emitter's location above or below the imaging plane. Thus, z-position information can be encoded within the shape of the image of each identified molecule. (Adapted from Huang et al. [31].) **C.** To determine the z position for a given peak within a STORM image, a calibration curve is first determined that relates the widths of the image in x and y as a function of z. The plot shows an example calibration, where each data point is an average of several Alexa 647 molecules. The red curve shows the fit to the data, whose functional form is described in the text. The x and y widths determined from fluorophores within a sample of interest are compared with this calibration curve to determine the z positions of the fluorophores. (Adapted from Huang et al. [31].)

red cyanine dyes can be switched back to the fluorescence state by excitation with a light source whose wavelength matches the activator absorption. Multicolor STORM images can be obtained with identical reporters paired to different activators by distinguishing the activation light wavelength [17] or with different reporters paired to the same type of activator by distinguishing emission wavelength. Combination of both approaches will allow an even larger number of colors to be obtained in a single image.

4.1.3. 3D STORM

Extension of STORM to 3D imaging can be done using optical astigmatism or multifocal plane imaging [31,32]. Figure 4.1B shows a schematic for implementing 3D STORM using astigmatism imaging [31]. A cylindrical lens inserted into the imaging path creates different focal planes for the x and y directions, causing x and y to defocus differently, depending on

Table 4.1. Photoswitchable fluorescent probes.

	Commercial availability	References	Reversibile?	Before activation		After activation		Oligomeric state
				Absorption	Emission	Absorption	Emission	
PA-FPS				maximum (nm)	maximum (nm)	maximum (nm)	maximum (nm)	
Dronpa	MBL Int	[24]	+	390	–	503	518	monomer
Kaede	MBL Int	[25]	–	508	518	572	580	tetramer
mEosFP	–	[26]	–	505	516	569	581	monomer
d2EosFP	–	[26]	–	506	516	569	581	dimer
Dendra2	Evrogen	[27]	–	490	507	553	573	monomer
PA-GFP	–	[23]	–	400	515	504	517	monomer (weak dimer)
PS-CFP2	Evrogen	[28]	–	400	468	490	511	monomer
Organic Dyes								
Cy5	GE Healthcare	[17]	+	–	–	649	670	–
Cy5.5	GE Healthcare	[17]	+	–	–	675	694	–
Cy7	GE Healthcare	[17]	+	–	–	747	776	–
Alexa 647	Invitrogen	[17]	+	–	–	650	665	–
Alexa 680	Invitrogen	–	+	–	–	679	702	–
HiLyte 680	Anaspec	–	+	–	–	678	699	–
Caged Fluoroscein	Invitrogen	–	–	–	–	497	516	–
Photochromic Rhodamine	–	[22]	+	230	–	570	580	–
Azido-DCDHF Flurogen	–	[38]	–	424	552	570	613	–

the molecule's position relative to its focal plane [33,34]. The shape (ellipticity) of the defocused spot carries depth information and is fit to an elliptical Gaussian function to extract the z position, while the centroid of the defocused image provides the x-y position, as done with two-dimensional STORM. This method attains a lateral resolution of 20–30 nm and an axial resolution of 50–60 nm over a range of several hundred nanometers in z without scanning the sample [31] or over a range of several microns in z when scanning the sample in the axial direction [35]. Another approach to obtaining the z-dimension information utilizes bi-focal-plane imaging [32]. By imaging a fluorophore simultaneously at two different focal planes, the two images can be fit with an experimentally determined 3D PSF to measure its 3D position. This technique achieved a resolution $\sim$75 nm in z and also has a working range of several hundred nanometers without scanning the sample, or several microns if scanning the sample in z.

4.1.4. Applications

The use of photoswitchable probes for sub–diffraction limit imaging has been demonstrated on a number of cellular targets. For example, the meshwork of microtubule filaments [17,18], the cortical actin [4], adhesion complexes [19,36], clusters of membrane proteins on the cell surface [37], and the 3D morphology of clathrin-coated pits [31] can now be probed by fluorescence imaging with a resolution much better than that of conventional confocal microscopy and approaching that of electron microscopy. In general, the technological developments achieved for this method of imaging have created an exciting new tool with the potential for studying previously unseen cellular features and dynamics.

In this chapter we describe the implementation of multicolor and 3D STORM for imaging cellular structures. First, the choice of fluorescent probe, such as photoswitchable organic dyes and fluorescent proteins, is discussed, along with methods for specific labeling of cellular proteins. Next, the instrumentation is described, along with the procedure for performing a STORM experiment. Data analysis routines are then explained for reconstructing a STORM image. The chapter concludes with example applications and, finally, specific protocols.

4.2. Labeling Cellular Targets with Photoswitchable Fluorescent Probes

A variety of photoswitchable fluorescent probes, summarized in Table 4.1, are available for use with STORM. In general, the choice of probe is specific to the desired application. Proteins can be labeled with an assortment of small-molecule photoswitchable dyes. For example, carbocyanine dyes such as Cy5, Cy5.5, Cy7 (GE-Healthcare), Alexa 647, and Alexa 680 (Invitrogen) or HiLyte 680 (Anaspec) can be easily conjugated to a protein target and will undergo multiple cycles of photoswitching, either through specific reactivation or by spontaneous blinking. Caged dyes that undergo a single photoswitching event, such as fluorescein, are also available from Invitrogen and can be used for STORM imaging. Additional dyes have recently been specifically synthesized to create photoswitchable probes, such as photoswitchable rhodamine and 2,7-dichlorodihydrofluorescein (DCDHF) fluorogen [21,38]. In general, however, any dye that can undergo photoswitching is acceptable. Several photophysical properties should be considered, depending on the specific application. These include, but are not limited to, the brightness, the contrast ratio, and the number of switching cycles before photobleaching. The brightness, namely the number of photons detected per switching cycle,

directly correlates with the localization precision and ultimately the image resolution. The contrast ratio is defined as the ratio of the fluorescence intensity in the activated state to that in the dark state, or equivalently the fraction of time the fluorophore spends in the dark state versus the fluorescent state in the absence of specific activation. This parameter has a significant effect on background fluorescence during the experiment and thus on the localization precision and density. Finally, the number of switching cycles determines in general how long a dynamic process can be followed, although in some special cases, dynamic imaging may also be obtained from probes that can only be activated once.

Alternatively, photoactivatable fluorescent proteins (PA-FPs) may also be used [23–30]. While these probes are less bright compared to some of the photoswitchable dye molecules and give lower localization precision, and thus potentially lower image resolution, the ability to directly fuse fluorescent proteins to the target of interest provides a powerful means for labeling proteins in cells [39]. Plasmid DNA encoding the PA-FPs can be obtained from several vendors (e.g., Evrogen and MBL International) or requested from the research groups that developed them. In choosing a PA-FP, a number of properties in addition to brightness, contrast ratio, and number of switching cycles should also be considered, including the absorption and emission spectra pre- and post-activation, the chromophore maturation time, and the oligomeric state. See Table 4.1 for a list of selected PA-FPs and some of their properties.

4.2.1. Labeling Proteins with Photoswitchable Dyes

A number of different strategies for linking a dye to a protein target have been developed. Typically, commercial dyes can be attached by thiol- or amine-reactive chemistry either directly to the protein of interest or to antibodies against that target. Immunofluorescence imaging is one such example, in which the antibody is labeled with a photoswitchable dye and the protein is subsequently stained with the antibody. A variety of primary and secondary antibodies can be purchased from companies such as Abcam and Invitrogen, and many protocols for immunostaining have been developed. As an example, amine-reactive, photoswitchable dyes can be conjugated to antibodies for use in immunofluorescence-based STORM (see Protocol 1).

Alternatively, a number of small-molecule labeling strategies for direct coupling of dyes to protein targets in cells with a genetic approach have recently been developed and may prove advantageous for sub–diffraction limit imaging [40–43]. In one set of methods, a genetically encoded peptide tag is attached to the N- or C-terminus of the protein of interest and serves as a specific target for dye conjugation. The tag can then be site-specifically labeled in two ways. One approach is to use a specific chemical group on the dye to bind with high affinity to the tag. For example, a tetracysteine motif has been used to bind to the biarsenical dye FLASH or ReASH for imaging proteins within living cells [44]. Alternatively, fluorescent probes consisting of a nickel-ion–containing moiety and a chromophore have been used to specifically chelate with an oligohistidine peptide for labeling proteins, also in live cells [45]. A second strategy is to use the peptide sequence as a substrate for enzymatic attachment, such that a fluorescent dye with a specific chemical group can be covalently conjugated to the tag [46–50]. A number of different enzymes have been used for this purpose, including biotin ligase [47], lipoic acid ligase [48], a transpeptidase such as sortase [49], and phosphopantetheinyl transferases [50]. In general, there exists a large array of techniques for attaching STORM-compatible dyes to a cellular target of interest. The approach of using specific peptide tags labeled with small organic dyes combines the advantages of bright and small fluorescent

labels with the high specificity of a genetic labeling strategy and may prove very useful for achieving true molecular-scale image resolution.

4.2.2. Immunofluorescence Staining of Cells

Targeting fluorescently labeled antibodies to cellular proteins can be done by standard immunofluorescence protocols. The temperature and fixatives used, along with their concentrations and fixation time durations, may affect the structure of the cellular target and the efficacy of antigens binding to the antibodies. Readers are referred elsewhere [51–53] for a detailed description of fixation chemistry. Fixation conditions can be optimized for specific applications. Formaldehye and glutaraldehyde, either individually or in combination, are the two most commonly used fixatives. Fixation reagents containing glutaraldehyde typically preserve the cellular structure better, especially at the ultrastructure level, as previously shown in electron microscopic imaging [54]. This is also found to be true for STORM imaging of fixed cells (see Protocol 2). Before antibody staining, the fixed cells are permeabilized with detergent, such as Triton X-100 or Tween-20, if intracellular proteins need to be labeled. For indirect immunofluorescence labeling of intracellular proteins, permeabilized cells are first incubated with primary antibodies followed by dye-labeled secondary antibodies against the primaries. Alternatively, primary antibodies that are labeled with photoswitchable probes can be used directly, omitting the secondary antibody-staining step.

4.2.3. Labeling Cellular Structures with Photoswitchable Proteins

PA-FPs can be genetically fused to the target protein of interest and introduced to the cell with a variety of expression methods. Making fusion DNA usually involves the preparation of insert DNA by polymerase chain reaction (PCR) and/or restriction enzyme digestion and ligation of the insert to a vector by the use of DNA ligase. For mammalian cell lines, the fusion construct can be transiently transfected or stably expressed in the cell. Protocol 3 provides condensed guidelines for transiently transfecting cells. Molecular cloning books and lab manuals, such as *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press), describe procedures for making fusion DNA and transfecting cells. For expression of multiple generations of cells, stable incorporation of the fusion DNA to the genome must be implemented. One approach is to include in the vector that encodes the fusion protein another gene that provides the cell with a certain selection advantage, such as resistance to antibiotics (e.g., neomycin). After transfection is performed, negative selection pressure corresponding to the encoded drug resistance gene (e.g., G418, also called Geneticin) is then applied to the cell culture. Viable cell colonies with an appropriate expression level are then selected and cultured as stable cell lines. Besides transfection with antibiotic resistance, retroviruses can also be used to introduce the gene of interest into the genome of the cell for stable expression (see http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=rv) [55,56]. Retrovirus systems (e.g., RetroMax System from Imgenex) provide higher incorporation efficiency and hence a shorter timeline for selection than transfection. For nondividing cells, lentiviral expression systems (e.g., ViraPower from Invitrogen) can be used to incorporate the gene of interest [57]. Expressing genetically encoded fusion proteins may lead to disruption of normal cellular localization and/or function. Thus, control experiments need to be done to assess these effects, and different expression levels, promoters, and linkages between the FPs and the target proteins may need to be tried to minimize the adverse effect of expression (or overexpression) of FP-fused proteins.

4.3. Instrumentation for Storm Imaging

STORM imaging can be performed using both epi- and total internal reflection fluorescence (TIRF) microscopes. Depending on the sample, TIRF microscopy may offer a distinct advantage in reducing out-of-focus background fluorescence, thus enabling more precise acquisition of single-molecule localizations. For imaging deeper into a sample, normal epifluorescence or epifluorescence with a high oblique incident angle close to the critical angle for total internal refection may be used. In this section we describe a typical STORM setup, which is shown schematically in Figure 4.2A.

4.3.1. Excitation Pathway

Light Sources Used for Various Multicolor STORM Imaging Schemes

Illumination of a sample is typically done with laser light of various colors, with the laser wavelength specific for the choice of fluorophore. One can select from a wide variety of solid-state, diode, or gas lasers to optimally match the illumination wavelength to the absorption spectra of the fluorescent probe(s) in question. Table 4.1 gives information on the peak absorption of a number of different photoswitchable probes. As an example, one can consider imaging photoswitchable carbocyanine dyes, such as Cy5, Cy5.5, and Cy7, by absorption of a red photon. Alexa 647, a structural analog to Cy5 with similar switching and spectral properties, can also be used. These fluorophores, which we call reporter dyes, can then be reactivated specifically with a different-color light if a second dye, such as Cy3, is positioned in close proximity [17]. We call this second dye an activator. Specific activation is accomplished by first illuminating the Cy3–Cy5 pair with green laser light and then exciting fluorescence from the Cy5 reporter with red light. A red laser source, such as a 657-nm diode-pumped solid-state laser (e.g., RCL-200-656, CrystaLaser), a 633-nm HeNe laser (e.g., 25-LHP-928-249, Melles Griot), or the 647-nm line from the Ar-Kr laser, can be used for excitation of the reporter cyanine dyes. For reactivation by Cy3, a 532-nm diode-pumped solid-state laser (e.g., GCL-200-L, CrystaLaser) or the 514568-nm line from the Ar-Kr laser can be used. Multiline lasers, such as the Innova 70C Spectrum Ar-Kr laser from Coherent or the Stabilite 2018 Ar-Kr laser from Spectra-Physics, provide many of the necessary laser lines needed for STORM imaging [17]. The photoswitchable cyanine reporters Cy5, Cy5.5, and Cy7 can also be spontaneously reactivated by the red imaging laser without pairing with the activator dye, albeit at a much lower rate. It is thus also possible to perform STORM imaging with only the red laser source.

The imaging scheme described above would allow one to do multicolor STORM by distinguishing different reporter dyes by their emission color. Alternatively, a second approach for multicolor STORM is to use spectrally dissimilar activator dyes paired with the same reporter. Since the laser with a wavelength matching the activator's absorption spectrum activates the dye pair with much higher efficiency (or rate) than those with nonmatching wavelengths, different probes can be identified by the color of light that activates them, [17]. For example, Cy3, Cy2, and Alexa 405 (A405) can all be used to activate Cy5. In addition to the aforementioned laser lines to activate the Cy3–Cy5 pair, one can use the 457-nm line of an Ar ion laser (e.g., 35-LAL-030-208, Melles Griot) to activate the Cy2–Cy5 pair and a 405-nm diode laser (e.g., CUBE 405, Coherent) to activate the A405–Cy5 pair. Using this approach, we have achieved three-color STORM imaging [17]. Three-color imaging has also been demonstrated with spectrally distinct photochromic rhodamines. In this case, these dyes

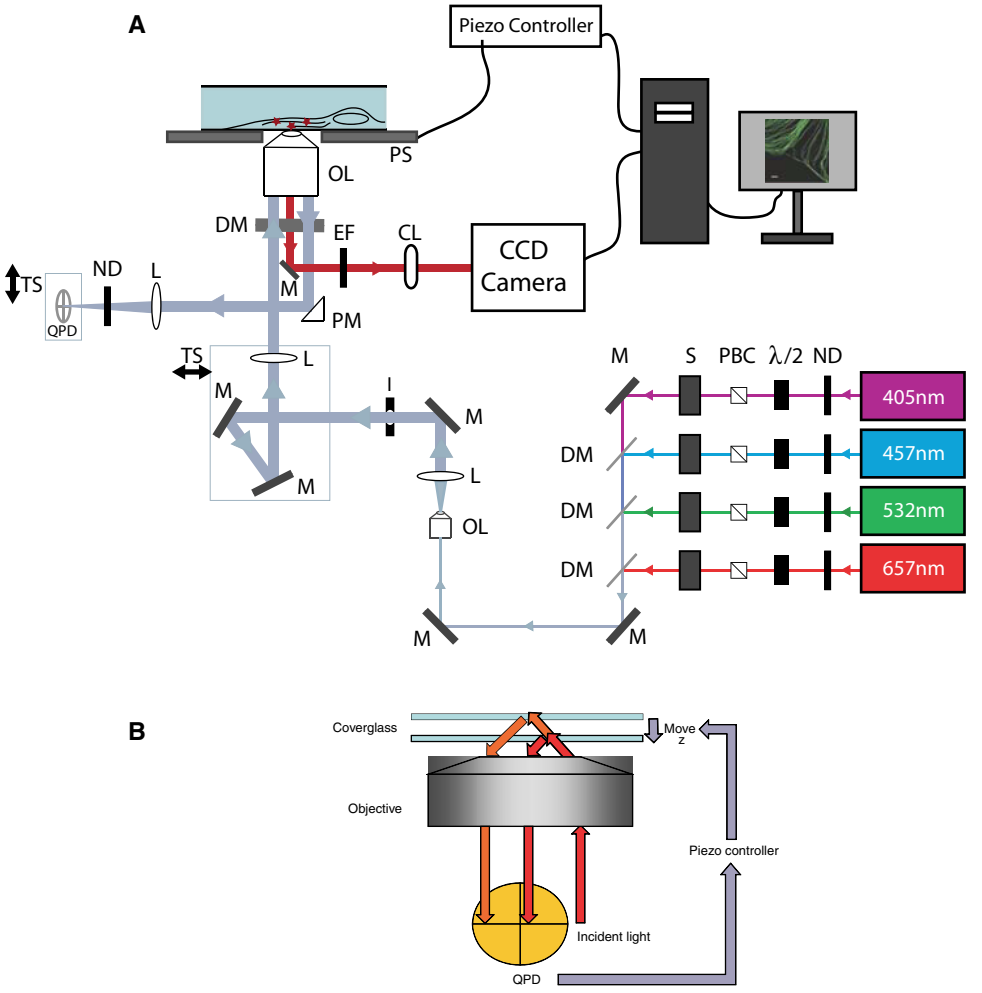


Figure 4.2. A stochastic optical reconstruction microscopy (STORM) setup with a focus lock system. **A.** The imaging and activation lasers each pass through optical components used to shutter and control the output intensity of each line before being combined via several dichroic mirrors. The combined beam passes first through a telescope composed of an objective lens and an achromat lens and then through an iris at the conjugate plane of the image plane. The beam then reflects off of two steering mirrors before being focused to the back focal place of the objective through the backport of the microscope. These components are placed on a translation stage to switch between total internal reflection fluorescence (TIRF) and epifluorescence. The laser reflects off of a dichroic mirror and passes through the objective to the sample, which is mounted on a piezo stage. The reflected laser from the coverglass–sample interface is directed to a quadrant photodiode (QPD) by a prism mirror for the focus lock system (see panel B). Fluorescence emission passes through the dichroic mirror, emission filters, and a cylindrical lens if one is performing three-dimensional STORM before being imaged onto the electron-multiplying charge-coupled device (EM-CCD) camera. Abbreviations: 405 nm, 457 nm, 532 nm, 657 nm, respective laser wavelengths; λ/2, half-wave plate; CL, cylindrical lens; DM, dichroic mirror; EF, emission filter; I, iris; L, lens; M, mirror; ND, neutral-density filter; OL, objective lens; PBC, polarizing beamsplitter cube; PM, prism mirror; PS, piezo sample stage; S, shutter; TS, translation stage. **B.** Schematic of the focus-lock working principle. An incident laser (shown in *red*) is reflected at the coverglass–sample interface. This reflected light hits the QPD. When the coverglass drifts upward, the reflected laser beam (*orange*) hits a different position on the QPD. The difference of signals between the left and right halves of the QPD normalized by the total signal is a measure of the coverglass *z*-position information, which is fed to a piezo controller. The piezo stage on which the coverglass and sample are mounted then moves to compensate for the coverglass movement. Alternatively, a piezo-controlled objective positioner can also be used for the focus lock.

can all be activated with a common ultraviolet light at 375 nm, and then the fluorescence can be excited using 532-nm light. The dyes are then distinguished by their emission color [22].

Other schemes to achieve multicolor imaging using PA-FPs or a combination of a PA-FP and a photoswitchable dye have also been demonstrated. For example, PA-FP rsFastLime has been combined with Cy5 in two-color imaging. 488- and 633-nm lasers were used to excite rsFastLime and Cy5, respectively [18]. Two different PA-FPs, Dronpa and EosFP, have been used for two-color imaging as well [19]. In this case, a 405-nm laser (405-50C, Coherent) was used to activate each of these PA-FPs. A 561-nm laser (GCL-150-561, CrystaLaser) was used to excite the activated form of EosFP, and a 488-nm laser (Cyan Scientific Laser, Spectra-Physics) was used to excite native-state EosFP and the activated form of Dronpa or PS-CFP2. However, since the emission of EosFP before activation spectrally overlaps with the emission of the postactivation state of Dronpa, two-color imaging was only obtained by imaging EosFP and Dronpa in a sequential manner, that is, Dronpa was imaged after all EosFPs were imaged and photobleached. This sequential imaging procedure is not compatible with recording dynamic processes, which requires simultaneous or quasi-simultaneous imaging of both probes at different time points. This deficiency is overcome by the recent development of a Dronpa variant, bsDronpa, with a blue-shifted emission, which allows simultaneous two-color imaging of Dronpa and bsDronpa according to their different emission wavelengths [58]. Both forms of Dronpa are activated with 375-nm light (iPulse-375, Toptica Photonics), and the fluorescence of the activated molecules is excited with the 476-nm line from an argon-ion laser (Innova 70C, Coherent).

Power Control and Temporal Modulation of Excitation Light

For each laser source, fine control of the output power is essential. For example, in a typical experiment with photoswitchable cyanine dyes, the red laser source is maintained at $\sim$0.5–2 kW/cm^2 to ensure that most reporter dyes are switched off sufficiently quickly. The activation lasers are initially set to a few W/cm^2 such that an optimal subset of molecules is switched on in a given field of view, and a high density of single molecules with nonoverlapping images can be identified. During the course of data acquisition, the power of the activation light sources may be increased to maintain a sufficient density of activated fluorophores as more dye molecules are photobleached. For coarse control of the laser power, the light initially passes through a neutral-density filter wheel with a number of different optical density adjustments. Fine control can then be accomplished through the use of a half-wave plate and a polarizing beamsplitter cube.

To temporally modulate the light source reaching the sample, mechanical shutters may be used. Some STORM experiments require periodic activation, which is accomplished by computer-controlled shuttering synchronized with the camera. Shutters are available from Vincent Associates with a number of different aperture size and speed options, but are currently limited to millisecond switching speeds. Alternatively, an acousto-optical tunable filter (AOTF) or acousto-optical modulator (AOM) may be used to control both the light intensity and the shuttering at significantly faster speeds.

Excitation Light Path

Before illuminating the sample, the light sources are combined and directed to the microscope. This can be accomplished by the use of dichroic mirrors that reflect one light

source but allow the others to be transmitted. As seen in Figure 4.2A, each laser has a dichroic mirror to reflect its wavelength and pass the longer (or shorter) wavelength light entering through the opposing side. These mirrors are supplied by companies such as Chroma, Omega Optical, and Semrock and can be custom made for various applications. Once combined, the excitation beam needs to be expanded to evenly illuminate the field of view. This can be done by passing the light through a commercial beam expander or a homemade telescope composed of either two achromatic lenses or an objective and an achromat lens positioned in such a way that the two lenses share a common focal plane. The magnification of the objective and the focal length of the lenses are chosen depending on the degree of beam expansion required.

Once the beam is expanded, an iris diaphragm is placed at an imaging plane conjugate to the sample plane for precise control of the illumination area. The beam is sent into the microscope by two mirrors on adjustable mounts, allowing for control of the laser direction, and then through a lens that focuses the light to the back focal plane of the objective. For convenience of switching between epifluorescence and TIRF illumination, these three optical elements can be placed on a single-axis lateral translation stage so that the illumination can be shifted off-axis toward the edge of the objective lens for TIRF illumination. Many commercially available microscopes, such as the Olympus IX-71, Zeiss AxioObserver, and Nikon Eclipse Ti inverted microscopes, can be used. The illumination beam is reflected off of a dichroic mirror in the microscope filter cube turret, which directs light of all incoming wavelengths to the objective and transmits fluorescence emission from the sample. For example, a 650-nm long-pass dichroic mirror (650DRLP, Omega Optical) works well for reflecting 405-, 457-, 532-, and 647-nm activation/excitation light sources typically used. Alternatively, a polychroic beamsplitter (z458/514/647rpc, custom made from Chroma, reflecting 488-, 532-, 633-, and 657-nm lasers efficiently and to some extent the 405-nm laser as well, or 400/535/635TBDR, Omega Optical, reflecting 405-, 488-, 532-, and 633-nm lasers, but not the 657-nm laser) can be used. The excitation laser light reflected off of this dichroic mirror then passes through an objective lens (e.g., 100x Olympus UPlanSApo oil immersion with a numerical aperture of 1.4) and illuminates the sample.

4.3.2. Emission Pathway

Emission Filters

The fluorescence emission from the sample is collected by the same objective used for excitation and is passed through the appropriate emission filters. A number of different bandpass and long-pass emission filters can be used, depending on the fluorophore being imaged. Most microscope systems offer an additional slot for a slider emission filter to be used. One can use a 51007m (Chroma) and a 595-700DBEM filter (Omega Optical) for Cy3 and Cy5 simultaneous detection, or a HQ710/70m (Chroma) stacked with a HQ665LP (Chroma) for more efficient detection of Cy5. For imaging Cy5, Cy5.5, and/or Cy7, a HQ710/80m bandpass filter can be used (Chroma). For other probes, such as those mentioned in Table 4.1, emission filters should be chosen according to excitation and emission spectra, the lasers involved, and whether a combination of probes is being used.

3D Imaging

STORM in 3D can be implemented by determining both the lateral and axial positions of the fluorophores simultaneously with high precision. This can be accomplished by using

optical astigmatism, in which a cylindrical lens is inserted into the imaging path to create two different focal planes for the x and y directions, resulting in an elliptical rather than a circular image for individual molecules [31]. When the z position of a fluorophore changes, the widths of its image in x and y change, and the ellipticity of the image can then be used to identify its z position (see Figure 4.1B), whereas the centroid of the image can be used to determine the x and y positions. Alternatively, 3D localization can be achieved by bi-focal-plane imaging, in which the fluorescence emission is split with a 50:50 beamsplitter cube in two paths, each with a slightly different path length, before being imaged onto nonoverlapping areas of the camera. By imaging two different focal planes simultaneously, the positions of the fluorophores can be determined in 3D by fitting their images in the two planes with a 3D PSF [32].

In the following, we give a detailed description of 3D STORM using astigmatism imaging. The fluorescence image is relayed through a pair of lenses with a weak cylindrical lens (e.g., 1-m–focal length achromat) inserted in between. In general, the working range and thus the sensitivity of determining the z position is dependent on the chosen focal length of the cylindrical lens. Inserting the cylindrical lens in the light path between the relay lens pair minimizes unwanted image distortion. The choice of the focal length of the cylindrical lens depends on the focal lengths of the other lenses used in the imaging pathway. The principle is to created a sizable ellipticity of the single-molecule images that is sensitive to the z position of the molecule, but is not so large as to degrade the x and y localization precision. When the fluorophore moves away from the focal plane, its image gets wider, and thus the intensity per unit area gets dimmer, which limits the z working range without scanning to within several hundred nanometers of the average focal plane. Scanning the focal plane in z, however, can increase the imaging depth to at least several microns in aqueous samples. Additionally, precise z-position determination needs correction of spherical aberrations if index mismatch occurs between the sample and immersion fluid of the objective (see Section 4.5.2 for a discussion of spherical aberrations caused by refractive index mismatch).

Detection of Fluorescence Emission

Finally, the fluorescence emission is projected onto an electron-multiplying charge-coupled device (EM-CCD) camera controlled by data acquisition software. EM-CCD cameras can be purchased from Andor, Roper Scientific, and Hamamatsu and are ideal for STORM imaging since electron multiplication enables greater sensitivity for single-molecule fluorescence detection. CCD cameras come with different chip sizes and data transfer speeds, and some are also back-illuminated, which provides higher quantum efficiency at a higher cost. In general, the typical camera used for STORM experiments has a 512×512 pixel chip and a pixel size of 16 μm, such as the Andor Ixon DV897DCS-BV, for which the full image frame can be read out at a maximum speed of 30 Hz. Alternatively, the Andor Ixon 860 has been used to achieve imaging speeds of $\sim$500 Hz, with a chip size of 128×128 pixels and a pixel size of 24 μm. The desired field of view and data acquisition rates vary, depending on the sample being imaged. To keep file sizes manageable, data may also be collected from only a subregion of the CCD. For example, a 256×256 region of a 512×512 CCD may be used. The optimal frame rate for STORM imaging depends on the rate at which the fluorophores are switching on and off. Maximum signal-to-noise ratio is achieved when the exposure time for a single frame is close to the average time that a fluorophore remains in the activated fluorescent state during a single switching event. The fluorophore's switching kinetics is in turn

governed by the illumination intensity and other factors. The magnification of the imaging optics is also an important parameter for STORM imaging. If the magnification is too high or too low, the accuracy of each localization is reduced [7]. In general, if the PSF of the microscope is treated as a Gaussian spot with standard deviation σ, the optimal magnification sets the sample area imaged by one pixel to be σ^2.

4.3.3. Focus Lock for Axial Stability

During the course of data acquisition, the distance between the objective and the sample may drift significantly, causing a shift in the image plane. For 2D STORM measurements, the change of focus compromises the image sharpness and hence the precision of localizing single molecules. Furthermore, it makes maintaining a specific image plane difficult. For 3D STORM, the focal drift causes incorrect identification of the z position.

This problem is remedied by incorporating a real-time feedback system to stabilize the focus ("focus lock") (Figure 4.2B). To achieve this, one can take advantage of the reflection of the imaging laser at the coverglass–sample (often biospecimen in aqueous medium) interface. The reflected laser beam can be redirected by a mirror or a reflective prism onto a quadrant photodiode (QPD), position-sensitive detector (PSD), or line CCD. For this system to work, the incident angle of the excitation light needs to be relatively large to easily separate the incident and reflected beams, as well as to achieve optimal sensitivity of the reflected beam position. The QPD, PSD, or CCD reads the position of the beam, which is sensitive to the distance between the objective and the coverglass. Taking a QPD as an example, the difference signal from the left and right halves of the QPD can be normalized over the sum signal from all quadrants to account for power fluctuations. Axial drift in the focus results in a change in the normalized difference signal of the QPD. The position information is then fed to a z-axis piezo stage that moves either the sample stage (e.g., Nano View-M, Mad City Labs) or the objective (e.g., F-100, Mad City Labs) via custom software (e.g., Labview) to compensate for the distance change. The z position of the focus is maintained within a $\sim$40-nm range, and residual z drift can be corrected during data analysis, which is described in a later section. For the aforementioned focus lock to work properly, it is important to maintain good pointing stability of the laser beam. Some lasers may be insufficiently stable, causing false corrections by the focus lock when the beam direction fluctuates. A solution is to couple the laser into an optical fiber to reduce directional fluctuations.

When the incident angle of the excitation light is small, this method is no longer effective, as explained earlier. Alternatively, an independent light source, preferably with an infrared wavelength to avoid affecting fluorescence excitation, decoupled from the excitation pathway can be used. Several microscope manufacturers (e.g., Olympus, Nikon, Zeiss) have also made commercially available units for maintaining the focus position based on similar principles. When considering purchasing a focus lock with the microscope, the correction rate, step size, and residual drift should be considered. Lateral drift of the sample also occurs during image acquisition, but can be corrected for as described in a later section.

4.4. Performing a Storm Experiment

In performing a STORM experiment, a number of steps must be taken to ensure proper sample preparation and imaging conditions for the desired application. These steps are described below. Depending on the cellular targets, photoswitchable probes must be linked to

the protein of interest either through immunofluorescence or direct labeling. The appropriate buffer conditions for optimal imaging must also be chosen depending on the fluorophore(s) being used. If one is performing 3D STORM, a calibration curve of the z position must also be taken. In all cases, a method of drift correction must be chosen, and depending on the method of choice, fiducial markers may need to be introduced in the sample preparation step. The sample is then imaged as described later.

4.4.1. Preparation of Cells

In the following, we use mammalian cell imaging as an example to describe the STORM imaging procedures. Cells such as BS-C-1 or HeLa cells are typically plated at 50%–75% confluency on glass coverslips or other chambered wells (e.g., from LabTek) that have a standard #1.5 glass substrate attached. Once plated, cells can be either immunostained with photoswitchable probe-labeled antibodies or transfected with the relevant photoswitchable fusion proteins. Example protocols for labeling antibodies, immunostaining of cells with antibodies, and transfecting with PA-FP fusion proteins are given in Protocols, 1, 2 and 3, respectively.

For immuno-STORM imaging with photoswitchable antibodies, cells are fixed and then stained for the relevant protein target (see Protocol 2 for example fixation and staining procedures). For direct immunofluorescence imaging, primary antibodies labeled with photoswitchable dyes are used. For indirect immunofluorescence imaging, unlabeled primary antibodies are first added, followed by secondary antibodies labeled with photoswitchable dyes. Optionally, one can repeat fixation after staining to reduce antibody dissociation.

Prior to imaging, the appropriate imaging buffer must be added to the sample well, depending on the photoswitchable probe being used. For fixed-cell imaging with photoswitchable cyanine dyes, a buffered solution of pH 7–8 is made, typically with 50 mM Tris and 10 mM NaCl. This buffer contains an oxygen-scavenging system, such as 0.5 mg/mL of glucose oxidase and 40 μg/mL of catalase with 1%–10% glucose, to reduce the effects of photobleaching. Furthermore, a primary thiol at a concentration of ~100 mM or less, such as β-mercaptoethanol (βME) or mercaptoethylamine, is included in the solution for photoswitching of cyanine dyes to occur [17]. In general, the switching behavior is sensitive to the pH of the solution, which should be roughly 7 or greater for efficient switching. When using the glucose oxidase/catalase scavenging system, protons will be released in converting oxygen and glucose into gluconic acid. Thus, over time, the pH will drop, which can adversely affect switching behavior, requiring strong buffering conditions or the buffer to be exchanged. When imaging a living cell, medium (phenol red–free Minimum Essential Medium) with an oxygen scavenger system with reduced strength (~0.5 mg/mL of glucose oxidase and 40 μg/mL of catalase, 1% glucose) and lower βME concentration (~10 mM) should be used to keep cells viable [17]. For imaging PA-FPs in fixed or living cells, the thiol reagent is not required. Photobleaching of FPs, however, happens quickly, and the aforementioned oxygen scavenger system may also be used to prolong imaging time, for example, in recording dynamic processes over time.

4.4.2. Calibration of *z* Position for 3D Imaging

To determine the z position of single fluorophores from the x and y widths of their images, a calibration curve must be measured first [31]. Calibration should be done before each experiment since it may vary depending on the optical alignment. Our calibration

method is to record the fluorescence from single molecules/particles within a single layer, such as the coverglass surface, while scanning in the z direction with a piezo stage. Both photoswitchable and nonswitchable dyes can be used for this purpose. Fluorescent beads or quantum dots can also be used. In general, however, the size and emission wavelength of the fluorescent object used for calibration should be matched to those used for STORM imaging. For each z position, averaged ellipticity information from all molecules in the field of view, namely the average x and y widths ($w_{x,\mathrm{calib}}$, $w_{y,\mathrm{calib}}$) of the images of individual fluorescent molecules, are determined as described in the data analysis section. The obtained ($w_{x,\mathrm{calib}}$, $w_{y,\mathrm{calib}}$) versus z curve constitutes the calibration curve (Figure 4.1C) for determining the z position of the molecules in the actual sample from their fluorescent images.

4.4.3. Choosing a Method for Drift Correction

Due to temperature variations in the room or other mechanical instabilities of the optical setup, the sample may drift relative to the objective. Although the focus lock reduces the axial drift substantially, the residual axial and lateral drift still need to be corrected for in the STORM image. One approach to correcting for the stage drift is to use fiducial markers, such as fluorescence beads (e.g., 580/605-nm, 200-nm diameter microspheres, Molecular Probes), to track the x, y, and z movement of the sample during image acquisition. Preferably, the fluorescent emission from these markers should be identifiable within each activation frame. Fiducial marker tracking is advantageous, in that it is independent of the cellular dynamics for live-cell STORM imaging. One disadvantage, however, is that fiducial markers adsorbed on the coverglass surface are useful only when one is imaging a thin sample of several hundred nanometers, because the images of the markers will be severely defocused and difficult to track with high precision when imaging thick samples significantly further away from the coverglass. Fiducial markers imbedded in a thicker matrix could be used for imaging thicker samples.

As an alternative approach, correlation between the STORM images during different time periods can also be used for drift correction. For example, the raw movie data can be divided into ten time windows, and the STORM image derived from the first window can be cross-correlated with the STORM image derived from the second, third,. . ., and tenth windows. The cross-correlation function of a given window with the first window is a function of x, y, and z. The function is then fit to find the peak, which represents the spatial displacement of the given time window relative to the first window, that can be used for drift correction. Alternatively, with proper configuration, periodically recorded conventional fluorescence images of the sample can also be used for this purpose. Use of STORM images for computing the correlation can be limited by the number of localizations obtained during a given time window. In contrast, the conventional fluorescence image correlation is independent of the number of acquired localizations. As an example, when imaging using the Cy3–Cy5 pair as the fluorescent probe, the conventional image obtained in the Cy3 channel can be used for drift correction. In general, all reversible photoswitchable probes should allow drift correction with conventional images. STORM image correlation is typically more effective for z drift correction than tracing fiducial markers or correlating conventional fluorescence images. The disadvantage of using image correlation between different time windows for drift correction is that it does not apply to live samples in which the spatial distribution of molecule changes with time.

4.4.4. Imaging a Sample

For STORM imaging of cells that are expressing photoactivatable FPs or stained with photoswitchable dye-labeled antibodies, the lasers and optics being used must be aligned properly to center the illumination light onto the sample. Choice of laser and laser power will vary depending on the probe(s) being imaged, as described earlier in the section Light Sources Used for Various Multicolor STORM Imaging Schemes. The *z*-calibration data are first acquired as described previously if one is imaging in 3D. The sample can then be mounted on the microscope stage. Next, the focus is adjusted to the desired position. The QPD is then repositioned such that the reflected laser is centered on the QPD, and the feedback loop for the focus lock is turned on in the data acquisition software.

Once the desired field of view is found, data can be taken with the proper laser sequence, imaging speed, and camera gain setting for optimal acquisition, depending on the probes being used. During the course of data acquisition, it may be desired to increase the activation laser powers to maintain an optimal density of activated molecules. In addition, *z*-stacking for thick-sample imaging can be accomplished by stepping the piezo stage to change the focal plane. A more detailed procedure for performing a STORM imaging experiment is described in Protocol 4.

4.5. Data Analysis

The data acquired in a STORM experiment typically consist of a stack of thousands to tens of thousands of image frames acquired at a rate of 10–500 Hz, with each frame an array of intensities from the CCD. The overall goal of the analysis routines is to find the location of every fluorophore in each frame with the highest possible accuracy and then to determine for each identified fluorophore the number of collected photons, its color identity, and a number of other important parameters, such as the peak's ellipticity. In a 2D STORM experiment, the positions of all identified peaks are then plotted in a single plane. For 3D STORM, the shape of each identified single molecule can be used to determine the axial position, and thus a final rendered image gives the full *x*, *y*, and *z* coordinates of all peaks. In both cases, a high-resolution map of the structure of interest is produced. The analysis procedure for producing a final image can be accomplished in four steps, each of which is explained below. The algorithms described later were written in either IDL or Visual C++ computing languages, which can be distributed to interested users upon request.

4.5.1. Peak Finding

For each frame of acquired data, images of individual single molecules must first be identified above the background signal, the sources of which include cellular autofluorescence and CCD camera noise. The stack of image frames is first loaded into custom software that enables one to look at each individual frame of the movie. To estimate the background fluorescent counts of the data, a frame with few activated molecules is selected, and the median of the pixel intensities within the frame is computed. The standard deviation (S.D.) of this frame is also computed to identify the background noise level. For a given background, a threshold value is set in terms of the number of S.D.'s above that background and can be used to screen for molecules with a sufficient signal-to-noise ratio. For example, a pixel with an intensity that is 3 S.D.'s above the background may be a candidate peak from a single molecule with a sufficient signal-to-noise ratio. This threshold is first tested by the user for

identifying individual peaks in a single frame and then adjusted and retested until the value enables optimal peak selection before running automated peak finding on the entire data set.

During peak identification, each frame of the stack is spatially filtered to reject low-frequency autofluorescence and high-frequency noise spikes associated with camera read noise that may be incorrectly identified as peaks. To accomplish this, a normalized two-dimensional Gaussian function with zero mean and standard deviation approximately equal to the microscope PSF is convolved with each image in the stack [59]. By doing this, the background signal is subtracted from the image and low-frequency signals are reduced in amplitude. High-frequency noise can later be filtered based on the sharpness criterion to be described. A pixel with a resulting intensity that is greater than or equal to the aforementioned threshold is identified as a peak and selected for further analysis. Specifically, a region around each local maximum whose size is defined by the PSF of the microscope is then fit to a 2D Gaussian as described in the next section and used to compute the x and y center positions, the sharpness, and the roundness of the peak. The sharpness parameter of a peak is computed as (the peak intensity of the center pixel – the average intensity of the surrounding pixels)/the height of the best-fitting Gaussian function. High-frequency noise spikes in the data that are falsely selected as peaks can be filtered out since these spikes typically span a single pixel, giving a sharpness parameter close to unity. The roundness of a peak is determined as the difference of the peak widths in the x and y directions divided by the sum of the two peak widths and can be used later to reject peaks that are distorted in the x or y direction, often an indication of overlapping molecules in 2D STORM. In 3D STORM with astigmatism imaging, this parameter will be used to determine the z positions of molecules because different z positions give different ellipticity values of their images.

4.5.2. Localizing Molecules in *x,y,* and *z* by Fitting

Identified peaks are fit with an elliptical Gaussian function:

$$G(x,y) = h\exp\left(-2\frac{(x-x_0)^2}{{w_x}^2} - 2\frac{(y-y_0)^2}{{w_y}^2}\right) + b$$

where h is the peak height, b is the background, (x_0, y_0) are the center coordinates, and (w_x, w_y) are the widths of the image in the x and y directions, respectively [31]. The photons can then be computed by taking the total number of counts collected in the peak, $(\pi w_x w_y h)/2$. This value is converted to photoelectrons and then photons using the camera manufacturer's calibrated curve for the electron multiplication and analog-to-digital converter (ADC) gain settings used during image acquisition. The z position of a molecule can be derived by searching the z-calibration curve (i.e., widths $w_{x,\text{calib}}$, $w_{y,\text{calib}}$ vs. z curve, Figure 4.1C) to find a best match. A best match is defined as when the expression

$$\sqrt{\left(w_x^{1/2} - w_{x,\text{calib}}^{1/2}\right)^2 + \left(w_y^{1/2} - w_{y,\text{calib}}^{1/2}\right)^2}$$

is minimized. It is recommended that a maximum value of this expression be set to reject irregular localizations, such as localizations derived from more than one molecule at a nearby x and y position activated at the same time.

The analysis procedure for generating a z-calibration curve is as follows. As described previously and in Protocol 4, a scanning movie is obtained by recording fluorescent molecules/particles within a single layer while scanning in z. The software then identifies

and fits the images of individual molecules/particles in each frame to determine their widths while the z position of each frame is read out from the piezo stage. The averaged widths from multiple molecules/particles then give a scatter plot of x and y widths versus z. The data points are then fit to the empirical defocusing curve

$$w_{x,y}(z) = w_0\sqrt{1 + \left(\frac{z-c}{d}\right)^2 + A\left(\frac{z-c}{d}\right)^3 + B\left(\frac{z-c}{d}\right)^4}$$

where w_0 is the image width for a molecule at the focal plane, c is the offset of the x or y focal plane from the average focal plane (the average focal plane being the z position where the image is spherical and symmetric in both x and y directions, defined as $z = 0$), d is the focal depth of the microscope, and A and B are coefficients of higher-order terms accounting for the nonideality of the imaging optics (ideal defocusing curves would have A and $B = 0$). An example of a calibration curve is shown in Figure 4.1C. We note that the fit of the z-calibration data to a functional curve is not essential, but is convenient when searching for the z position that best matches z position dependence determined by the experimentally measured w_x and w_y values. The foregoing functional expression of the calibration curve is empirical, and other expressions could also be used. The image gets wider (i.e., more defocused) when the sample moves away from the focal planes, yielding lower localization accuracy [7]. For 3D STORM, this sets a practical limit on the working range in z without scanning that depends on the resolution desired [31] Larger imaging depth can be achieved by combining the above approach with z-scanning [35].

An important factor to consider is the effect of spherical aberration, especially that associated with imaging an aqueous sample with an oil immersion objective. A typical oil immersion objective with high numerical aperture is designed to work ideally in the situation where the refractive index (RI) of all materials between the focus and the objective lens matches that of the immersion oil to avoid spherical aberrations [60]. For imaging samples in an aqueous medium, such as cells adhered to coverglass and immersed in a buffer, index mismatch occurs between the coverglass/immersion oil (RI = 1.515–1.518) and the imaging buffer (e.g., RI = 1.35 for a 10% glucose solution). This index mismatch causes the apparent z position of the molecule to shift away from the coverglass due to the bending of light rays that occurs at the glass/buffer interface. This distortion in the z position can be corrected by a linear rescaling when the molecule is close to the glass surface [31,60] [35]. For instance, for a numerical aperture of 1.4, if the glass RI = 1.515 and imaging buffer RI = 1.35, the rescaling factor is 1.39 or 0.72 ($z_{\text{measured}} = 1.39z_{\text{actual}}$, or $z_{\text{actual}} = 0.72z_{\text{measured}}$, where z_{measured} is the z position determined from the calibration curve and z_{actual} is the actual z position of the fluorophore). This rescaling works with high precision for molecules that are within several hundred nanometers of the interface. As the imaging depth increases, spherical aberration arises because the refracted light rays no longer converge perfectly, distorting the point spread function and making it asymmetric in the axial direction. Additional measures need to be taken to reduce and/or correct for the spherical aberration, such as using a water immersion objective that index-matches with the aqueous buffer, using an imaging medium that index-matches with the oil immersion objective, or accepting only the localizations below the focal plane in the case of index mismatch, because the difference between z_{measured} and z_{actual} was found to be much smaller for molecules below the focal plane than for those above [35]. Spherical aberration does not affect x, y localization as significantly.

For fixed-sample imaging, images from neighboring frames can be used in combination to boost localization precision. Once all identified peaks have been fit and all relevant peak parameters have been computed, the data are then organized to link peaks across frames. Due

to the low activation density, a fluorescent spot that appears in the same pixel or in one of the neighboring pixels in successive frames most likely arises from the same fluorophore. The centroid positions of these peaks are connected as a "string" in which the initial frame in which the peak is found is defined as the start of the string, and the last frame in which the same peak appears is the end of the string. For a given string, the final rendered position can be calculated as the average centroid position weighted by the number of photons collected in each frame. This procedure typically improves the localization precision because more photons are used to determine the position of each fluorophore. The use of this string average for live-cell imaging may be limited since the molecule may move a substantial amount between frames.

4.5.3. Color Identification

Color identification of a photoswitchable fluorophore can be performed by distinguishing the wavelength of the light used to activate its fluorescence or by the emission wavelength of the fluorophore itself. In the first case, typically, a known repetitive sequence of alternating activation and imaging laser exposures is used where each frame can be identified by the color of activation. Thus, a fluorophore that is found to be activated in an imaging frame immediately following an activation pulse is assigned a color corresponding to the activation laser color. Fluorophores that are spontaneously activated in other frames are identified as nonspecific events because their activation color is not defined; these localizations may also be used in some cases, especially in single-color imaging. In the second case, the color of the fluorophore can be identified by the wavelength of its fluorescence emission. This can be achieved by separating the emission into two paths with a dichroic mirror, imaging both paths onto a camera, and determining the ratio of intensities in the two paths [22,61].

4.5.4. Drift Correction

Several drift correction methods, including the use of fiducial markers and image correlations, can be used to correct for lateral and axial drift of the sample that occurs during imaging. If fiducial markers are used, an average marker position versus time trace is generated by identifying the peaks corresponding to these markers in each image frame. Typically, only a few fiducial markers within a single field of view are used because the bright fluorescence of the fiducial markers prevents STORM imaging in the region immediately surrounding the marker. For tracking the lateral position of the markers, one can use fitting to find the x, y centroid position and track it over the course of data acquisition. The average x, y position trace of the markers as a function of time can then be used for drift correction. The z position can also be tracked, for example, by determining the ellipticity of the image in astigmatism imaging. Alternatively, one can use the localizations of the photoswitchable molecules for drift correction. For this purpose, the data are divided into multiple time windows. A 2D (x and y) and/or 1D (z) correlation function is calculated between the localizations in the first time window and those in each of the subsequent windows, which provides a time series of cross-correlation functions. For each time window, a peak position, (x, y) or z, of the correlation function is found by fitting to a Gaussian, giving a time trace of the drift. The time trace is then interpolated to all imaging frames assuming continuous drift, and the drift is then subtracted from the localizations in each frame. For live-cell STORM imaging, only the method of tracking fiducial markers is appropriate because the distribution of molecules changes over time.

4.5.5. Cross-Talk Subtraction

For multicolor imaging, one can optionally consider correcting the color cross-talk, which can arise from false activation by the incorrect activation laser (e.g., 532-nm activation of Cy2), false identification of the reporter dye (e.g., a Cy5 identified as a Cy5.5), or nonspecific activation of the probes by the imaging laser. A cross-talk subtraction/correction procedure can be used to computationally correct false activations statistically. In the following, we give a specific example for cross-talk subtraction in the case of using the Alexa 405–Cy5, Cy2–Cy5, and Cy3–Cy5 pairs for three-color imaging. For example, the Cy3–Cy5 pair may be falsely activated by 457- and 405-nm light. Thus, the cross-talk ratios for misidentifying Cy3 as Cy2 or Alexa 405 will be the fraction of Cy3–Cy5 molecules that will switch on when activated by 457- or 405-nm light, respectively, as compared to the fraction that will switch on when activated by 532-nm light. These ratios can be determined for all different pairs and have been reported previously [17]. For a given location within the STORM image the density of localizations observed, D_1, D_2, and D_3, for each of the three colors can be expressed as

$$\begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \begin{pmatrix} 1 & C_{2\to1} & C_{3\to1} \\ C_{1\to2} & 1 & C_{3\to2} \\ C_{1\to3} & C_{2\to3} & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

where $C_{i\to j}$ is the fraction of color i incorrectly identified as color j and d_1, d_2, and d_3 are the actual respective densities of the three different colored molecules. Knowing the cross-talk ratios and the observed density of molecules from the image, we can compute the probability, P, that a given localization in the image is falsely identified. For color i, the probability that a given localization within the image is incorrectly identified can be written as

$$P = 1 - \frac{d_i}{D_i}$$

In the cross-talk subtraction algorithm, P is computed for each localization, and it is stochastically retained or removed according to this probability. This procedure of color-cross-talk subtraction can be generally applied to other probes, provided that the cross-talk ratio can be determined. Some degradation of resolution may occur due to cross-talk subtraction because the local density of localizations D_i and d_i are estimated in a finite-sized area.

4.5.6. Displaying the Image

To display a STORM image, the position of each of the previously identified localizations is plotted. Typically, the image can be rendered with each single molecule displayed as a marker to identify its center of mass or as a normalized Gaussian whose width, $\sigma_{x,y}$, scales with the localization accuracy. For example, one can use the theoretical localization accuracy determined by the equation

$$\sigma_{x,y} = \left(\frac{s^2 + a^2/12}{N} + \frac{8\pi s^4 b^2}{a^2 N^2} \right)^{1/2}$$

where s is the standard deviation of the imaging setup PSF, a is the pixel size after magnification of the microscope, b is the background noise level, and N is the number of photons collected [7]. In general, these parameters should be determined for each microscope. This Gaussian rendering scheme displays molecules in a way such that peaks with smaller $\sigma_{x,y}$ are rendered brighter and sharper than those with larger $\sigma_{x,y}$. In practice, the experimentally

determined localization accuracy typically does not reach the theoretical value. Although the theoretical localization accuracy is often used for image rendering to indicate which localizations should be considered with more weight, it should not be considered a quantitative measure of the actual localization accuracy because it does not account for other factors that reduce this accuracy, such as finite probe size, sample drift, and so on. For multicolor imaging, localizations of different probes will be presented in different colors.

4.5.7. Additional Filtering of the Image

One can use the number of parameters obtained from the peak identification and fitting to remove localizations that do not meet certain criteria that often indicate a high probability of being an erroneous localization. Peaks that are misshapen or too elliptical can be filtered by the roundness criterion described earlier. Localizations derived from fluorescent spots containing a very small or large number of photons can also be removed. Peaks that are too bright in the image may indicate multiple activated fluorophores in close proximity or other fluorescent contaminants. Peaks that are too dim will give poor localization accuracy. The final STORM image is a collection of all localizations in either 2D or 3D that satisfy the foregoing criteria.

4.6. Example Applications

STORM has been used to demonstrate sub–diffraction-limit resolution in mammalian cells in both two and three dimensions, as well as with multiple colors simultaneously. In the following, we describe a few examples from our lab that have provided insight into some of the new and unique information that can be obtained from a STORM image as compared to a conventional fluorescence image.

One of the first model systems that we imaged with STORM is the microtubule network in mammalian cells. For this purpose, BS-C-1 cells were fixed and immunostained with photoswitchable probe–labeled antibodies against microtubules [17]. The staining protocol is given in Protocol 2. The photoswitchable probe used was the Cy3–Alexa 647 pair, in which Alexa 647 is a photoswitchable cyanine dye structurally analogous to Cy5, and Cy3 serves as the activator. A conventional image of the microtubule network is shown in Figure 4.3A and an expanded region in Figure 4.3C. In the diffraction-limited image it is not possible to completely disambiguate all of the microtubules within the region. A STORM image of the same area (Figure 4.3B and D), however, shows significantly greater detail and makes identification of each microtubule filament possible. We then extended this technique to image multiple different color probes simultaneously. For example, we performed immunofluorescence-based STORM on both microtubules and clathrin-coated pits. Figure 4.4A shows a conventional image of the microtubules in green and clathrin in red. The STORM image of the same region (Figure 4.4B) and an expanded region (Figure 4.4C) shows unique information that cannot be obtained from diffraction-limited light microscopy. As before, all individual microtubules can now be identified. In addition, in regions of the conventional image where clathrin-coated pits appear to be colocalized with microtubules, the STORM image reveals that many are in fact not in direct contact. Furthermore, coated pits appear as diffraction-limited spots in the conventional image with no discernible morphology. The STORM image, however, reveals the spherical structure of the pits. To determine the intrinsic optical imaging resolution, we identified nonspecifically bound antibodies, each appearing as a cluster of localizations. The localization distribution was then fit to a Gaussian whose full-width at half-maximum

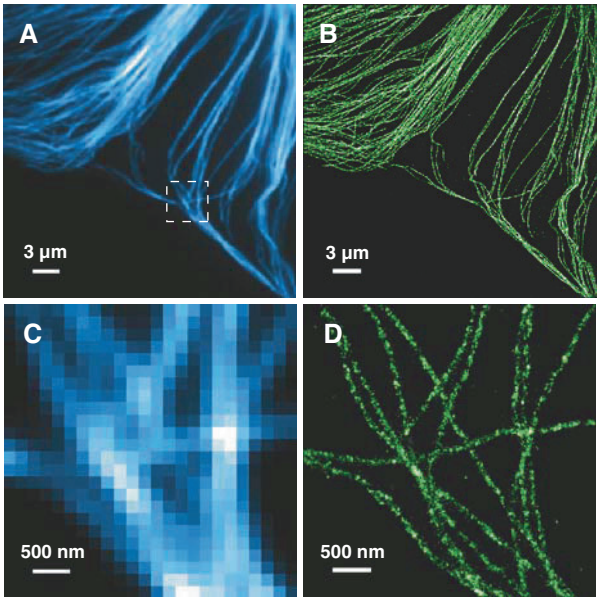


Figure 4.3. Single-color stochastic optical reconstruction microscopy (STORM) imaging of microtubules. **A, B.** Conventional and STORM images, respectively, of a region of a BS-C-1 cell. **C, D.** The corresponding images of the region identified by the white box in panel A. (Adapted from Bates et al. [17].)

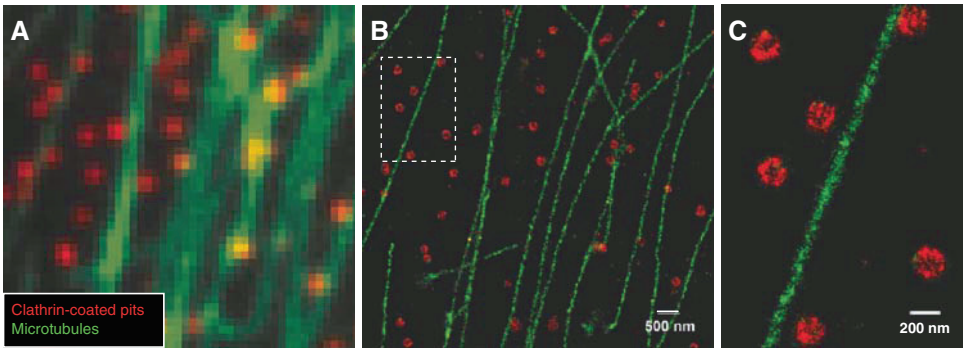


Figure 4.4. Multicolor stochastic optical reconstruction microscopy (STORM) imaging of microtubules and clathrin-coated pits. **A.** A conventional two-color image of microtubules and clathrin-coated pits. **B.** A STORM image of the same region in panel A shows significantly greater detail than what can be seen in the conventional image. **C.** A closer look at the region identified by the white box in panel B. (Adapted from Bates et al. [17].)

(FWHM) of 25 nm represents the intrinsic optical resolution. The effective resolution also depends on the label size and the label density. The sub-diffraction-limited lateral resolution provided by STORM imaging and its ability to use a diversity of different colored photoswitchable probes can provide previously unseen details of protein interactions within a cell.

We then applied 3D STORM to image the same cellular structures described [31]. Microtubules and clathrin-coated pits in BS-C-1 cells were immunostained with Cy3-Alexa 647–labeled antibodies indirectly with dye-labeled secondary antibodies or directly with

dye-labeled primary antibodies, respectively. The conventional image of microtubules of a BS-C-1 cell is shown in Figure 4.5A. The corresponding STORM image is shown in Figure 4.5B and is color coded based on *z* position. Cross sections of a smaller region in Figure 4.5B are shown in Figure 4.5C, illustrating the structure of five microtubules from different vantage points. Clathrin-coated pit conventional and STORM images are shown in Figure 4.6. Serial *x*-*y* and *x*-*z* cross sections (Figure 4.6F and G, respectively) clearly show the hollow shape of the pit, which would be otherwise indiscernible with conventional fluorescence microscopy. Besides the improved resolution, another unique advantage of 3D STORM is that structures can be probed without scanning the excitation source. When combined with *z* scanning, thicker samples can be imaged. We determined our optical resolution to be 20–30 nm in the lateral directions and 50–60 nm in the axial direction. Beyond the immunofluorescence STORM examples mentioned above, live-cell imaging has also been demonstrated with photoswitchable FPs [36,37] and cyanine dyes (unpublished results). The ability of STORM to determine ultrastructural features in cells at physiological conditions with nanometer-scale resolution opens a new window for bioimaging, providing important new insights into biological systems by unveiling previously unseen details of molecular structures, cells, and tissues.

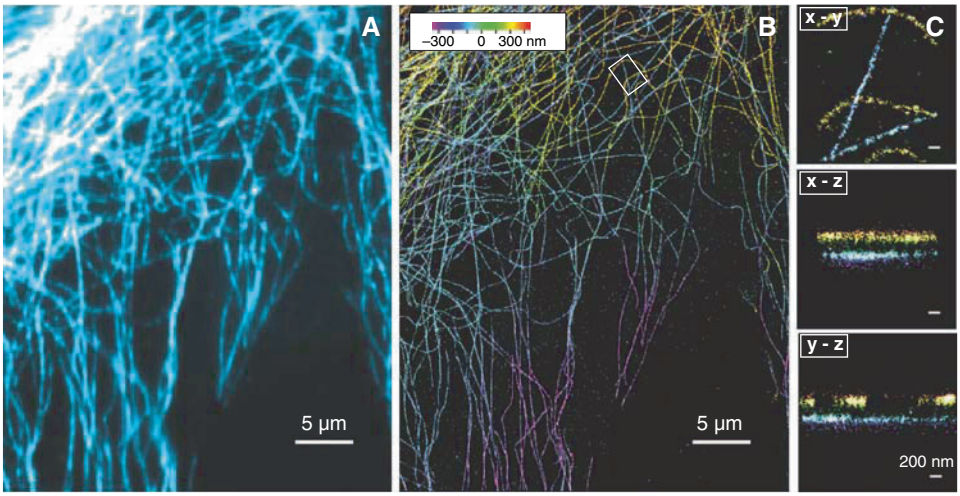


Figure 4.5. Three-dimensional (3D) stochastic optical reconstruction microscopy (STORM) imaging of microtubules. **A, B.** Conventional and 3D STORM images, respectively, of immunofluorescence-stained microtubules of a BS-C-1 cell. The position of each localization in the STORM image is color-coded based on its position in *z*. **C.** The *x*–*y*, *x*–*z*, and *y*–*z* cross sections of the region highlighted in panel B. (Adapted from Huang et al. [31].)

4.7. Protocol 1

4.7.1. Labeling Antibodies or Other Proteins with Organic Dyes

This protocol describes the labeling and purification of primary or secondary antibodies with photoswitchable organic dyes. Activator–reporter pairs are used here as examples, but the protocol can also be used for labeling with other photoswitchable dyes or can be adapted for labeling other proteins with amine-reactive dyes. For other labeling chemistries, refer to the manufacturers' protocols.

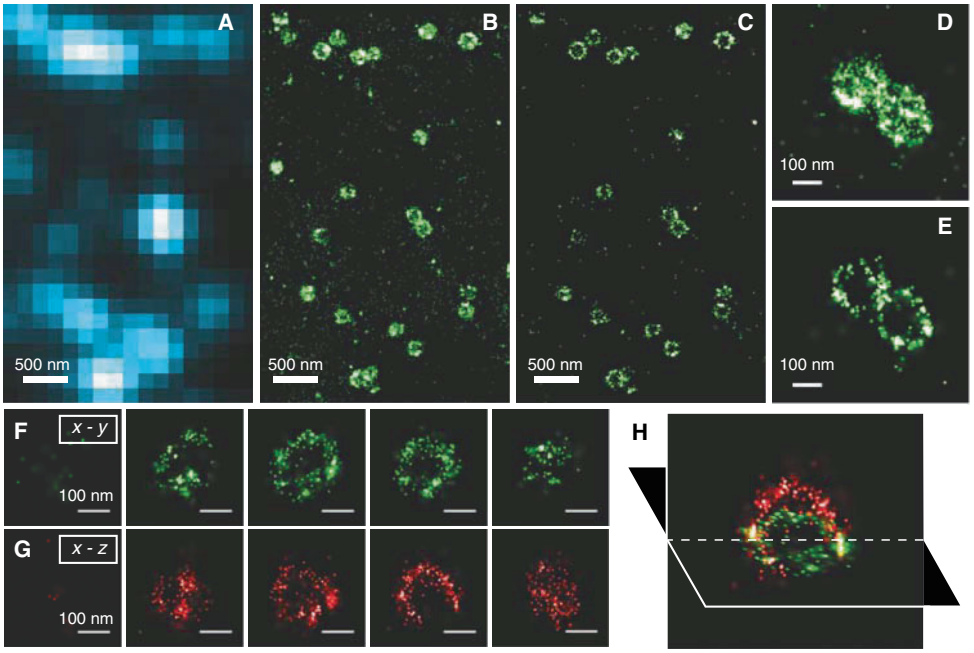


Figure 4.6. 3D stochastic optical reconstruction microscopy (STORM) imaging of clathrin-coated pits. **A, B.** Conventional and 2D STORM images, respectively, of directly immunostained clathrin-coated pits in a cell. **C.** A 50-nm-thick *x–y* cross section of the same area showing clathrin localized to the periphery of the pits. **D.** An enlarged view of two adjacent pits in the STORM image, with all *z* localizations included. **E.** The 100-nm-thick *x–y* cross section of the same two pits in panel D, showing the ringlike structure. **F.** The serial 50-nm-thick (in *z*), *x–y* cross sections of a clathrin-coated pit. **G.** For the same pit, the serial *x–z* cross sections with a 50-nm thickness in *y*. **H.** One of the *x–y* and *x–z* cross sections presented in 3D perspective to show the half-spherical cage structure of the clathrin-coated pit. (Adapted from Huang et al. [31].)

Materials

Reagents

Amine-reactive dyes (Cy2, Cy3, Cy5, Cy5.5, Cy7, GE-Healthcare; Alexa 405, Alexa 647, Invitrogen)
Dimethyl sulfoxide (DMSO) <!>
Sodium bicarbonate ($NaHCO_3$) 1 M stock solution
Phosphate-buffered saline (PBS)
Primary or secondary antibody

Equipment

Gel-filtration columns (Nap5 columns, GE-Healthcare)
Shaking platform
Ultraviolet/visible spectrophotometer

Methods

1. Prepare the amine reactive dyes by resuspending 1 mg of dried dye in the manufacturer's recommended solvent (e.g., DMSO for Cy5). Aliquot into 50 tubes and vacuum dry for a final aliquot size of 0.02 mg. Store at –20°C until use.
2. Dissolve one aliquot of the activator (e.g., Alexa 405, Cy2, or Cy3) and one of the reporter dyes (e.g., Cy5, Cy5.5, Cy7, or Alexa 647) each in 10 μL of anhydrous DMSO.
3. Prepare a single 50- to 60-μL labeling reaction containing ~1.0 mg/mL of antibody, 100 mM $NaHCO_3$, and different amounts of each of the two dye solutions, depending on the level of labeling desired.
   Ideally, one should aim for ~0.4–0.8 reporter and ~2–4 activators per antibody. Therefore, start with a ~3–5 times higher concentration of activator to reporter.
4. Agitate gently at room temperature for 30 min.
5. Remove the unreacted dyes by purification over a Nap5 gel-filtration column. First, equilibrate the column with three volumes of PBS. Bring the labeling reaction mixture up to 100 μL with PBS, load it onto the column, and elute with PBS.
6. Determine the antibody and dye concentrations using a spectrophotometer. Determine the absorption at 280 nm for the antibody and at the respective absorption peaks for the dyes. Use the known extinction coefficients for the antibodies and dyes to determine their concentrations. Be sure to correct for the dyes' absorbance at 280 nm. The labeling ratio of each dye per antibody can then be determined (see Troubleshooting).

Troubleshooting

Problem: The dye-per-antibody ratio is too low or too high.
Solution: When labeling an antibody with both a reporter and an activator dye, the appropriate dye concentrations needed to achieve ideal labeling ratios must be determined empirically. Once the dye-to-antibody ratios are measured, the labeling may need to be repeated with the dye concentrations adjusted appropriately.

4.8. Protocol 2

4.8.1. Cell Fixation and Staining

This protocol explains the procedure for fixation of cells and for immunostaining with switch-labeled antibodies.

Materials

Reagents

Cells to be fixed
Phosphate-buffered saline (PBS)
Fixation solution <!>

3% formaldehyde and 0.1% glutaraldehyde in PBS

Sodium borohydride ($NaBH_4$)

0.1% in PBS

Blocking buffer (BB) <!>

3% Bovine serum albumin (BSA) + 0.5% Triton X-100 in PBS

Primary antibody
Labeled secondary antibody
Washing buffer (WB) <!>

0.2% BSA + 0.1% Triton X-100 in PBS

Equipment

Eight-well glass-bottom chambers (LabTek)

Methods

1. Culture cells in eight-well glass-bottom chambers (LabTek II) to ~50%–60% confluency.
   All volumes subsequently described are the volumes to be used PER WELL after removing the previously added volume. Fixation and staining are to be done at room temperature.
2. Wash the cells once with 500 μL of PBS.
3. Fix the cells using 200 μL of fixation solution for 10 min, followed by a reducing step for 7 min with 200 μL of 0.1% $NaBH_4$.
   $NaBH_4$ should be prepared right before use to avoid hydrolysis.
4. Wash three times with 200 μL of PBS.
5. Permeabilize the cells in 200 μL of BB for 15 min.
6. Add 150 μL of the primary antibody, diluted in BB, and incubate for 30 min at room temperature.
   The appropriate concentration used will depend on the specific cellular target and antibody efficiency. For example, for imaging microtubules, use a 1:200 dilution of 0.5 mg/mL mouse anti–β-tubulin (ATN01, Cytoskeleton; or T4026, Sigma).
7. Wash three times with 200 μL of WB, with a 10-min incubation between washes.
8. Add 150 μL of the labeled secondary antibody diluted in BB, and incubate for 30 min at room temperature.
   The appropriate concentration used will depend on the specific cellular target and antibody efficiency.
9. Wash three times with 200 μL of WB, with a 10-min incubation between washes.
10. Optionally, at this point one can do a second fixation step by repeating steps 2–4, omitting the reducing step.
11. Fixed and stained cells can then be stored in 200 μL of PBS prior to and following imaging.

4.9. Protocol 3

4.9.1. Transient Transfection of Cells for Expression of Fusion Constructs of PA-FPs and Proteins of Interest

This is an abbreviated protocol describing the procedure for transfecting cells with fusion protein plasmids for transient expression. Refer to detailed instructions by the manufacturers (also see the previously cited *Molecular Cloning: A Laboratory Manual*).

Materials

Reagents

Cell medium (e.g., Eagle's Minimum Essential Medium, ATCC 30-2003) with serum
Serum-free cell medium (e.g., Minimum Essential Medium, Gibco 51200)
Cells to be transfected
Plasmids containing the PA-FP fusion protein
Transfection reagent (e.g., FuGene, Roche Applied Sciences)

Equipment

Routine light microscope
Fluorescence microscope
Incubator
Petri dishes (e.g., 50-mm glass bottom, MatTek)

Methods

1. Culture cells in glass-bottomed petri dishes to ~70% confluency.
   Confluency can be easily checked on a routine binocular microscope.
2. Per petri dish, add 3–6 μL of transfection reagent into a total of 100 μL of serum-free cell medium into a seperate tube.
   Directly pipet transfection reagent into medium to avoid contact with plastics walls of the tubes.
3. Add 1–2 μg of plasmid DNA to the diluted reagent from step 2. Mix and incubate for 15 min at room temperature.
4. For each petri dish, change the cell medium to 2 mL of fresh serum-containing medium before adding the reagent DNA complex to the cells.
5. Add the complex from step 3 dropwise onto the cells.
6. Depending on the transfection reagent being used, the cells may need to be rinsed with fresh medium after adding the transfection complex. Follow manufacturers' instructions.
7. Return the cells to the incubator with the appropriate preset conditions.
8. Check the expression level on a conventional fluorescence or STORM microscope 24 hr after transfection. If transfection efficiency or expression level is too low, return the cells to the incubator for longer time before imaging. Observe whether there is any morphological change using a fluorescence microscope, differential interference contrast, or phase contrast.

4.10. Protocol 4

4.10.1. Using STORM to Image a Sample

This protocol describes the procedure of a STORM experiment on cellular structures immunostained with photoswitchable dye–labeled antibodies or genetically tagged with PA-FPs.

Materials

Reagents

Cell fixation and staining reagents (Protocol 2) <!>
Phosphate-buffered saline (PBS)
STORM imaging buffer

> 50 mM Tris, pH 7.5–8.5, 10 mM NaCl, 0.5 mg/mL of glucose oxidase (G2133, Sigma-Aldrich), 40 μg/mL of catalase (106810, Roche Applied Science), 10% (w/v) glucose, and 1% (v/v) β-mercaptoethanol (βME) <!>
> *Catalase and* βME *are added immediately before use.*

3D calibration materials, for example, one of the following:

> Proteins (such as antibodies, protein A, BSA, or streptavidin) labeled with switchable or nonswitchable dyes

Equipment

STORM microscope (see Instrumentation section), mainly composed of:

> Optical table
> EM-CCD camera (e.g., Ixon DV897DCS-BV, Andor)
> Shutters (e.g., from Uniblitz) or AOTF/AOMs
> Lasers <!>
> Inverted microscope
> Mirrors, lenses
> Data acquisition software

Excitation filters for multiline lasers if needed
Emission filters
Cylindrical lens if imaging in 3D
z-axis piezo sample stage or objective positioner if imaging in 3D (e.g., Nano View-M or F-100, Mad City Labs)
Focus lock system if needed

> Optical fiber, if needed
> Quadrant photodiode, position-sensitive detector, or line CCD
> Mirror or reflective prism
> Focus lock control software

Glass slide, coverslip for calibration slide
Analysis software

Methods

1. Allow the lasers to warm up and the EM-CCD camera to cool. The piezo sample stage/objective positioner should be set to the middle of its working range (50 μm). Align the excitation and detection light path and set the appropriate laser powers (e.g., for Cy3–Alexa 647, a strong imaging 657-nm laser, e.g., 1 kW/cm^2 for Alexa 647, and a weak 532-nm laser, e.g., 1 W/cm^2, are appropriate).
   The filters to be used will vary depending on the laser source and fluorophore(s). See the Instrumentation section for details.
2. Insert the cylindrical lens into the detection path if imaging in 3D.
3. Take *z*-calibration data if planning to image in 3D:
   3.1 Prepare a coverslip with proteins conjugated with cyanine dyes nonspecifically bound on its surface.
   Surface density should be such that the field is full of distinguishable fluorescent molecules. Their images, however, should not overlap.
   3.2 Adjust the incident angle of the illumination laser to above the critical angle for total internal reflection (TIR). With the proper laser (e.g., 657-nm diode-pumped solid-state laser for Alexa 647) and microscope configurations, record the fluorescence of these molecules while scanning the piezo stage in *z*.
4. Change the sample buffer to the proper STORM imaging buffer as described in the Preparation of Cells section. Mount the petri dish/coverslip on the microscope. Adjust the focus and excitation incident angle.
   Unless the labeled structures of interest are all on the surface of the coverglass, it is advisable to shift the excitation beam away from the critical angle for TIR, slightly toward epi-illumination. This increases the excitation depth to several microns and still reduces fluorescence background from the sample [62].
5. If needed, align the focus lock system.
6. Record data using an illumination laser sequence depending on the probes being used and with the imaging speed and camera gain set for optimal signal-to-noise ratio.
   The laser sequence used defines the order in which the sample is illuminated by the various excitation and activation wavelengths. Computer-controlled laser shutters are synchronized with the camera frames and are opened and closed according to the predefined laser sequence. Cycles of short activation frames (typically 1 frame) are followed by longer imaging/photobleaching frames (typically 5–10 frames). When faster imaging speed is needed, increase the laser power and use fewer imaging frames per cycle. When necessary, the activation laser power is adjusted gradually over time to compensate for photobleaching and maintain an optimal density of activated fluorophores. An alternative imaging method is to expose the sample continuously to the imaging laser and simply record the spontaneous activation of the fluorophores, avoiding the need for shuttering and activation pulses.
7. During the 3D imaging of a thick sample, *z*-scanning can be done by stepping the piezo stage to bring the objective closer to the coverglass, hence imaging deeper into the sample.
   When stepping the piezo stage, make sure the focus lock feedback system is stopped. Otherwise, the focus lock will attempt to return to the previous z position. The QPD position may need to be readjusted if out of range. The step size is determined by the

working range of 3D STORM, which depends on the objective and the cylindrical lens used. The typical step size is 300–600 nm. The number of steps required to complete a whole-cell image depends on the cell thickness. The image quality may deteriorate when imaging a large distance away from the coverglass if the refractive indices of the sample solution (e.g., 1.33 for water, 1.35 for 10% glucose solution) and coverglass (1.515) are not matched.

Troubleshooting

Problem: The photoswitchable dye-pair labeled sample is no longer switching.
Solution: During the course of an experiment, the glucose oxidase/catalase scavenging system will convert glucose to gluconic acid, lowering the pH of the solution. If the pH becomes too low, the switching will be adversely affected. Thus, exchange fresh solution to the sample. To maintain switching longer, use strong pH buffers, maximize the solution volume, and seal the sample to reduce oxygen content.

Problem: Nonspecific activation is too high for single-molecule detection.
Solution: Large numbers of nonspecifically activated molecules may appear per frame if the sample is too densely labeled. This results in overlapping molecules within a diffraction-limited spot, making single molecules indistinguishable. To fix this, repeat the staining with lower antibody concentrations, with a lower dye:antibody ratio, or lower photoswitchable protein expression level. In the case of cyanine dye imaging, using mercaptoethylamine instead of β*ME* also helps to reduce nonspecific activation.

Problem: The labeling density is too low for sufficient image reconstruction.
Solution: Try using different fixation conditions, different antibody concentrations, or different antibodies until optimal labeling is achieved.

Problem: Significant background fluorescence is coming from the sample or coverglass surface.
Solution: If there is significant cellular autofluorescence, check that the excitation is adjusted to near-TIR or TIR configuration. Use proper emission filters to reduce the cell autofluorescence signal. If background is coming from the coverglass surface, it can be cleaned by sonication in 1 M KOH for 15 min, rinsed with milliQ water, followed by additional sonication for 15 min and rinsing with milliQ water.

Acknowledgments

We thank W. M. Bates for a critical reading of the manuscript and other members of the Zhuang laboratory for helpful advice and discussions. This work was supported in part by the National Institutes of Health. XZ is a Howard Hughes Medical Institute investigator.

References

1. Hell, S. W., and J. Wichmann. 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated emission depletion microscopy. Opt. Lett. 19:780–782.
2. Gustafsson, M. G. L. 2005. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proc. Natl. Acad. Sci. USA 102:13081–13086.

3. Rust, M. J., M. Bates, and X. Zhuang. 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods 3:793–796.
4. Betzig, E. 2006. Imaging intracellular fluorescent proteins at nanometer resolution. Science 313:1642–1645.
5. Hess, S. T., T. P. K. Girirajan, and M. D. Mason. 2006. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophys. J. 91:4258–4272.
6. Hell, S. W. 2007. Far-field optical nanoscopy. Science 316:1153–1158.
7. Thompson, R. E., D. R. Larson, and W. W. Webb. 2002. Precise nanometer localization analysis for individual fluorescent probes. Biophys. J. 82:2775–2783.
8. Gelles, J., B. J. Schnapp, and M. P. Sheetz. 1988. Tracking kinesin-driven movements with nanometre-scale precision. Nature 331:450–453.
9. Ghosh, R. N., and W. W. Webb. 1994. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. Biophys. J. 66:1301–1318.
10. van Oijen, A. M., J. Kohler, J. Schmidt, M. Muller, and G. J. Brakenhoff. 1998. 3-Dimensional super-resolution by spectrally selective imaging. Chem. Phys. Lett. 292:183–187.
11. Lacoste, T. D., X. Michalet, F. Pinaud, D. S. Chemla, A. P. Alivisatos, and S. Weiss. 2000. Ultrahigh-resolution multicolor colocalization of single fluorescent probes. Proc. Natl. Acad. Sci. USA 97:9461–9466.
12. Yildiz, A., J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin. 2003. Myosin V walks hand-over-hand: single fluorophore imaging with 1.5 nm localization. Science 300:2061–2065.
13. Gordon, M. P., T. Ha, and P. R. Selvin. 2004. Single-molecule high-resolution imaging with photobleaching. Proc. Natl. Acad. Sci. USA 101:6462–6465.
14. Qu, X. H., D. Wu, L. Mets, and N. F. Scherer. 2004. Nanometer-localized multiple single-molecule fluorescence microscopy. Proc. Natl. Acad. Sci. USA 101:11298–11303.
15. Lidke, K. A., B. Rieger, T. M. Jovin, and R. Heintzmann. 2005. Superresolution by localization of quantum dots using blinking statistics. Opt. Exp. 13:7052–7062.
16. Churchman, L. S., Z. Okten, R. S. Rock, J. F. Dawson, and J. A. Spudich. 2005. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. Proc. Natl. Acad. Sci. USA 102:1419–1423.
17. Bates, M., B. Huang, G. T. Dempsey, and X. Zhuang. 2007. Multicolor super-resolution imaging with photo-switchable fluorescent probes. Science 317:1749–1753.
18. Bock, H., C. Geisler, C. A. Wurm, C. Von Middendorff, S. Jakobs, A. Schonle, A. Egner, S. W. Hell, and C. Eggeling. 2007. Two-color far-field fluorescence nanoscopy based on photoswitchable emitters. Appl. Phys. B 88:161–165.
19. Shroff, H., C. G. Galbraith, J. A. Galbraith, H. White, J. Gillette, S. Olenych, M. W. Davidson, and E. Betzig. 2007. Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. Proc Natl Acad Sci USA 104:20308–20313.
20. Bates, M., T. R. Blosser, and X. Zhuang. 2005. Short-range spectroscopic ruler based on a single-molecule optical switch. Phys. Rev. Lett. 94:108101.
21. Folling, J., V. Belov, R. Kunetsky, R. Medda, A. Schonle, A. Egner, C. Eggeling, M. Bossi, and S. W. Hell. 2007. Photochromic rhodamines provide nanoscopy with optical sectioning. Angew. Chem. Int. Ed. Engl. 46: 6266–6270.
22. Bossi, M., J. Folling, V. N. Belov, V. P. Boyarskiy, R. Medda, A. Egner, C. Eggeling, A. Schonle, and S. W. Hell. 2008. Multicolor far-field fluorescence nanoscopy through isolated detection of distinct molecular species. Nano Lett 8:2463–2468.
23. Patterson, G. H., and J. Lippincott-Schwartz. 2002. A photoactivatable GFP for selective photolabeling of proteins and cells. Science 297:1873–1877.
24. Habuchi, S., R. Ando, P. Dedecker, W. Verheijen, H. Mizuno, A. Miyawaki, and J. Hofkens. 2005. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. Proc. Natl. Acad. Sci. USA 102:9511–9516.
25. Ando, R., H. Hama, M. Yamamoto-Hino, H. Mizuno, and A. Miyawaki. 2002. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. Proc Natl Acad Sci USA 99: 12651–12656.
26. Wiedenmann, J., S. Ivanchenko, F. Oswald, F. Schmitt, C. Rocker, A. Salih, K.-D. Spindler, and G. U. Nienhaus. 2004. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc. Natl. Acad. Sci. USA 101:15905–15910.
27. Gurskaya, N. G., V. V. Verkhusha, A. S. Shcheglov, D. B. Staroverov, T. V. Chepurnykh, A. F. Fradkov, S. Lukyanov, and K. A. Lukyanov. 2006. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat. Biotechnol. 24:461–465.

28. Chudakov, D. M., V. V. Verkhusha, D. B. Staroverov, E. A. Souslova, S. Lukyanov, and K. A. Lukyanov. 2004. Photoswitchable cyan fluorescent protein for protein tracking. Nat. Biotechnol. 22:1435–1439.
29. Stiel, A. C., S. Trowitzsch, G. Weber, M. Andresen, C. Eggeling, S. W. Hell, S. Jakobs, and M. C. Wahl. 2007. 1.8 A bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. Biochem. J. 402:35–42.
30. Tsutsui, H., S. Karasawa, H. Shimizu, N. Nukina, and A. Miyawaki. 2005. Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. EMBO Rep. 6:233–238.
31. Huang, B., W. Wang, M. Bates, and X. Zhuang. 2008. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. Science 319:810–813.
32. Juette, M. F., T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf. 2008. Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. Nat. Methods 5:527–529.
33. Kao, H. P., and A. S. Verkman. 1994. Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. Biophys. J. 67:1291–1300.
34. Holtzer, L., T. Meckel, and T. Schmidt. 2007. Nanometric three-dimensional tracking of individual quantum dots in cells. Appl. Phys. Lett. 90:053902.
35. Huang, B., S. A. Jones, B. Brandenburg, and X. Zhuang 2008. Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. Nat. Methods 5:1047–1052.
36. Shroff, H., C. G. Galbraith, J. A. Galbraith, and E. Betzig. 2008. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nat. Methods 5:417–423.
37. Hess, S. T., T. J. Gould, M. V. Gudheti, S. A. Maas, K. D. Mills, and J. Zimmerberg. 2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. Proc. Natl. Acad. Sci. USA 104:17370–17375.
38. Lord, S. J., N. R. Conley, H. L. Lee, R. Samuel, N. Liu, R. J. Twieg, and W. E. Moerner. 2008. A photoactivatable push-pull fluorophore for single-molecule imaging in live cells. J. Am. Chem. Soc. 130:9204–9205.
39. Giepmans, B. N. G., S. R. Adams, M. H. Ellisman, and R. Y. Tsien. 2006. The fluorescent toolbox for assessing protein location and function. Science 312:217–224.
40. Chen, I., and A. Y. Ting. 2005. Site-specific labeling of proteins with small molecules in live cells. Curr. Opin. Biotech. 16:35–40.
41. Prescher, J. A., and C. R. Bertozzi. 2005. Chemistry in living systems. Nat. Chem. Biol. 1:13–21.
42. Lin, M. Z., and L. Wang. 2008. Selective labeling of proteins with chemical probes in living cells. Physiology (Bethesda, MD) 23:131–141.
43. O'Hare, H. M., K. Johnsson, and A. Gautier. 2007. Chemical probes shed light on protein function. Curr. Opin. Struct. Biol. 17:488–494.
44. Griffin, B. A., S. R. Adams, and R. Y. Tsien. 1998. Specific covalent labeling of recombinant protein molecules inside live cells. Science 281:269–272.
45. Guignet, E. G., R. Hovius, and H. Vogel. 2004. Reversible site-selective labeling of membrane proteins in live cells. Nat. Biotechnol. 22:440–444.
46. Lin, C. W., and A. Y. Ting. 2006. Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins *in vitro* and on the surface of living cells. J. Am. Chem. Soc. 128:4542–4543.
47. Chen, I., M. Howarth, W. Y. Lin, and A. Y. Ting. 2005. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. Nat. Methods 2:99–104.
48. Fernandez-Suarez, M., H. Baruah, L. Martinez-Hernandez, K. T. Xie, J. M. Baskin, C. R. Bertozzi, and A. Y. Ting. 2007. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. Nat. Biotechnol. 25:1483–1487.
49. Tanaka, T., T. Yamamoto, S. Tsukiji, and T. Nagamune. 2008. Site-specific protein modification on living cells catalyzed by Sortase. ChemBioChem 9:802–807.
50. Zhou, Z., P. Cironi, A. J. Lin, Y. Xu, S. Hrvatin, D. E. Golan, P. A. Silver, C. T. Walsh, and J. Yin. 2007. Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. ACS Chem. Biol. 2:337–346.
51. Ramos-Vara, J. A. 2005. Technical aspects of immunohistochemistry. Vet. Pathol. 42:405–426.
52. Skepper, J. N. 2000. Immunocytochemical strategies for electron microscopy: choice or compromise. J. Microsc. 199:1–36.
53. Glauert, A. M., and P. R. Lewis. 1998. *Biological specimen preparation for transmission electron microscopy*. Princeton University Press, Princeton, NJ.
54. McIntosh, J. M., ed. 2007. *Cellular electron microscopy* (Methods in Cell Biology, Vol. 79). Academic Press, San Diego, CA.

55. Naviaux, R. K., E. Costanzi, M. Haas, and I. M. Verma. 1996. The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. J. Virol. 70:5701–5705.
56. Pear, W. S., G. P. Nolan, M. L. Scott, and D. Baltimore. 1993. Production of high-titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. USA 90:8392–8396.
57. Buchschacher, G. L., Jr., and F. Wong-Staal. 2000. Development of lentiviral vectors for gene therapy for human diseases. Blood 95:2499–2504.
58. Andresen, M., A. C. Stiel, J. Folling, D. Wenzel, A. Schonle, A. Egner, C. Eggeling, S. W. Hell, and S. Jakobs. 2008. Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. Nat. Biotechnol. 26:1035–1040.
59. Stetson, P. B. 1987. DAOPHOT: a computer program for crowded-field stellar photometry. Publ. Astron. Soc. Pacific 99:191–222.
60. Egner, A., and S. W. Hell. 2006. *Handbook of biological confocal microscopy*, Chapter 20. Springer, New York.
61. Hohng, S., C. Joo, and T. Ha. 2004. Single-molecule three-color FRET. Biophys. J. 87:1328–1337.
62. Cui, B. X., C. B. Wu, L. Chen, A. Ramirez, E. L. Bearer, W. P. Li, W. C. Mobley, and S. Chu. 2007. One at a time, live tracking of NGF axonal transport using quantum dots Proc. Natl. Acad. Sci. USA 104:13666–13671.

5

Single-Molecule FRET: Methods and Biological Applications

Ling Chin Hwang, Johannes Hohlbein, Seamus J. Holden, and Achillefs N. Kapanidis

Abstract Since the first single-molecule fluorescence resonance energy transfer (FRET) measurement in 1996, the technique has contributed substantially to our understanding of biological molecules and processes by probing the structure and dynamics of nucleic acids, protein molecules, and their complexes with other molecules. This review discusses basic concepts and current developments in single-molecule FRET methodology, as well as examples of applications to systems such as nucleic acid machines and molecular motors.

5.1. Introduction

Probing molecular structure, dynamics, and interactions at the 1- to 10-nm scale is central to our understanding of the molecules responsible for the myriad of fascinating processes occurring each minute in living cells. The special interest in this length scale arises from the simple fact that most of the proteins, nucleic acids, and lipids that form the structures, sensors, gates, and machines in the cell have nanometre-scale dimensions. For example, double-stranded DNA has a diameter of 2 nm; haemoglobin has a diameter of ∼6 nm; and the lipid bilayer of a cell membrane is ∼5 nm thick. Ever since it was realized that diffraction poses a fundamental limit in the direct visualisation of biomolecules using conventional optical microscopy, researchers have recruited various ingenious ways to break the diffraction barrier and access the all-important nanometre scale. A celebrated member in the category of such methods is based on the phenomenon of fluorescence resonance energy transfer (FRET), also known as Förster resonance energy transfer in honor of Theodor Förster, who in 1948 developed the main FRET theory. Since its first demonstration as a molecular ruler (Stryer and Haugland 1967), FRET has been used extensively at the levels of both molecular ensembles and single molecules and has been an instructive method in both *in vitro* and *in vivo* settings. FRET is not only a tool for basic science studies, but also has sparked several important applications. For example, clever designs in fluorophore combinations for FRET enabled faster and cheaper sequencing of the human genome. In addition, several bioassays used in the biotechnology industry are based on the concept of FRET. In this review, we provide a

L. C. Hwang, J. Hohlbein, S. J. Holden, and A. N. Kapanidis • Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, OX1 3PU, UK

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_5,

short introduction on FRET and discuss fundamentals of single-molecule FRET along with selected applications on biological and bio-inspired systems. For further reading, readers are also encouraged to consult other excellent reviews on ensemble FRET (Clegg 1992; Selvin 1995; Selvin 2000) and single-molecule FRET (Ha 2001; Ha 2004; Schuler and Eaton 2008; Weiss 1999; Weiss 2000).

5.2. FRET Fundamentals and Ensemble FRET

FRET is a distance-dependent, non-radiative energy transfer process that occurs between a fluorophore and a chromophore as a consequence of long-range dipole–dipole coupling (Förster 1948); in most cases, the chromophore is also a fluorophore, and the probes are introduced in a single biomolecule or a complex of biomolecules. After excitation of the high-energy fluorophore (donor, D; Figure 5.1A,C) via an external light source, the energy can be transferred to a low-energy fluorophore (acceptor, A) as long as two preconditions are fulfilled: first, both fluorophores are in close proximity (<10 nm); and second, the emission

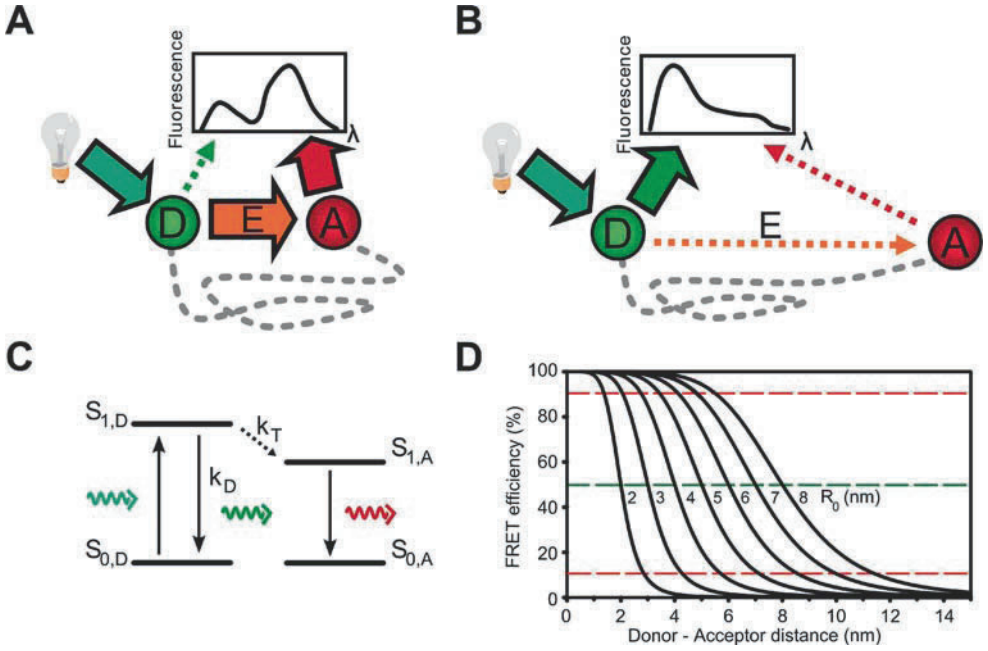


Figure 5.1. Principles of fluorescence resonance energy transfer (FRET). **A.** A light source directly excites the donor dye molecule (*green*). As long as the acceptor dye (*red*) is in close proximity, the energy can be transferred non-radiatively from the donor to the acceptor, resulting in increased emission intensity. **B.** If the distance between the donor and the acceptor is increased, the fraction of transferred energy decreases, resulting in increased emission intensity in the part of the spectrum corresponding to the donor emission wavelength. **C.** Simplified Jablonski diagram showing the transition pathways occurring in the energy transfer process. After excitation of the donor dye into its first excited singlet state, two processes can occur: the donor system can either relax to its ground state by emitting a green photon with a rate constant k_D, or the energy can be transferred to the acceptor molecule with a rate constant of k_T. In the latter case, the acceptor relaxes by emitting a red photon. Other electronic transitions, such as bleaching and intersystem crossing, are neglected. **D.** The FRET efficiency strongly depends on the Förster radius R_0 and is plotted as a function of the donor-to-acceptor distance and R_0.

spectrum of the donor overlaps with the absorption spectrum of the acceptor. As a consequence of energy transfer, the emission intensity of the donor decreases ("donor quenching") and the emission intensity of the acceptor increases ("acceptor sensitization").

The FRET process is often represented by a simplified Jablonski diagram (named after Polish physicist Aleksander Jabłoński) (Figure 5.1C), which is a free-energy state diagram that illustrates the electronic states of the fluorophores and the transitions between them. After excitation of the donor ($S_{0,D} \rightarrow S_{1,D}$), the donor can relax to its ground state $S_{0,D}$ either by emitting a photon or by losing energy as heat. In the presence of an acceptor in close proximity, the donor can also transfer its energy non-radiatively to the first excited singlet state $S_{1,A}$ of the acceptor. The FRET efficiency E between the two excited states $S_{1,D} \rightarrow S_{1,A}$ (Lakowicz 2006) is given by Eq. 5.1, where k_D the fluorescence emission rate constant of the donor in the absence of the acceptor and k_T, the energy transfer rate constant. According to Förster, E is inversely proportional to the sixth power of the distance R between the two fluorophores,

$$E = \frac{k_T}{k_T + k_D} = \frac{1}{1 + (R/R_0)^6} \tag{5.1}$$

The Förster radius R_0, which describes the donor-to-acceptor distance where the FRET efficiency equals 50%, takes values between 2 and 7 nm (Selvin 2000) for common pairs of fluorophores and can be calculated using

$$R_0^6 = \frac{9000(\ln 10)\Phi_D \kappa^2}{128\pi^5 N n^4} \int_0^\infty f_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \tag{5.2}$$

where Φ_D is the donor quantum yield in the absence of the acceptor, κ^2 is the orientation factor, N is Avogadro's number, and n is the refractive index of the intervening solution. The overlap integral of the donor emission spectrum and the acceptor absorption spectrum is calculated as a function of the wavelength λ using $\varepsilon_A(\lambda)$ as the molar extinction coefficient of the acceptor and $f_D(s)$ as the normalized emission spectrum of the donor, $\left(\int f_D(\lambda)d\lambda = 1\right)$. The orientation factor κ^2 describes the orientation of the emission dipole of the donor with respect to the absorption dipole of the acceptor and is often a matter of debate (Dale et al. 1979; Haas et al. 1978; Wu and Brand 1992). The factor κ^2 can take values between 0 (for perpendicular dipoles) and 4 (for collinear dipoles). In the limiting case of unrestricted rotational freedom of the dyes, averaging about all possible relative orientations yields $\kappa^2 = 2/3$ (Dale et al. 1979). Rotational freedom can be assessed by measuring fluorescence anisotropy, the change in polarisation between excitation and emission light. For polarised excitation, anisotropy is defined as:

$$r = \frac{I_\perp - I_\parallel}{I_\perp + 2I_\parallel} \tag{5.3}$$

where $I_\parallel$ is the emission intensity parallel to excitation, and $I_\perp$ is the emission intensity perpendicular to excitation. For cases of low anisotropy, the rotational freedom is high and the corresponding assumption of $\kappa^2 = 2/3$ is likely to be valid. For the less common case of high anisotropy, the rotational freedom of at least one fluorophore is likely to be restricted and therefore the uncertainty on R_0 and any associated distances will increase; the uncertainty increases further if both fluorophores have restricted rotational freedom.

The FRET efficiency can be determined in many ways (Clegg 1992; Jares-Erijman and Jovin 2003). Often, the FRET efficiency is measured using the decrease in the fluorescence

emission intensity of the donor or the "FRET-sensitized" emission intensity of the acceptor. Moreover, the equation $\tau^{-1} = k_D + k_T = \tau_D^{-1} + k_T$, which relates the rate constants in Eq. 5.1, to the fluorescence lifetime (τ is the donor fluorescence lifetime in the presence of the acceptor, and τ_D is the donor fluorescence lifetime in the absence of the acceptor) provides another means for determining the FRET efficiency according to

$$E = 1 - \frac{\tau}{\tau_D} \tag{5.4}$$

The $1/R^6$ dependence of the energy transfer on the donor-to-acceptor distance was experimentally verified in classic experiments by Haugland and Stryer in 1967 in which oligomers of poly-L-proline were used as spacers between the donor and the acceptor (Stryer and Haugland 1967); additional experiments using double-stranded DNA as a standard provided further validation (Clegg et al. 1993). Such studies established FRET as a type of "spectroscopic ruler", in that its use offers the intriguing possibility of determining the distance between the fluorophores by measuring the FRET efficiency between fluorophores and calculating the distance R_0 for the fluorophore pair used. During the last 30 years, FRET has become a widely used method on the ensemble level, with many applications on biomolecular structure and dynamics, especially with regards to proteins, nucleic acids, and protein–nucleic acid complexes. FRET between different variants of the green-fluorescence protein (GFP) has also been used extensively for testing for the presence of specific protein–protein interactions in living cells (Lippincott-Schwartz et al. 2001; Miyawaki and Tsien 2000; Zimmer 2002). Ensemble FRET has been extensively reviewed in many excellent reviews covering a broad range of applications (Clegg 1992; Clegg 1995; Jares-Erijman and Jovin 2003; Selvin 1995; Selvin 2000; Wu and Brand 1994). An extensive overview of different fluorophores for FRET can also be found in Sapsford et al. 2006.

Ensemble FRET measurements can provide much insight into the molecular world of biology. However, since ensemble methods report on the mean properties of populations of billions of molecules (e.g., a 16-μL solution of 10 nM DNA contains 1 billion DNA molecules), sample heterogeneity can skew the mean properties to a degree that complicates or even thwarts the interpretation of experiments. This is especially true in cases of large, unstable, or dynamic biomolecules such as protein–nucleic acid complexes or unfolded proteins. For instance, an ensemble of proteins in solution is composed of proteins in either their native state or a denatured state. In the latter case, the mean end-to-end distance of the N-terminus and the C-terminus increases. If both ends are labeled with a donor and an acceptor, respectively, it is clear that an ensemble FRET measurement will reveal only an averaged FRET efficiency with limited access to the different subpopulations (Haas et al. 1975). Ensemble methods are also limited by their inability to monitor stochastic (and, thus, unsynchronizable) dynamic motions, especially under equilibrium conditions. Such motions produce "dynamic heterogeneity", defined as the presence of molecular conformations in which a single molecule or molecular complex dwells for a measurable time during the timescale of observation.

5.3. Single-Molecule FRET Methods Based on Single-Laser Excitation

Some of the shortcomings of ensemble FRET were overcome by single-molecule FRET, which is essentially the observation of FRET at the level of single molecules (Figure 5.2). To achieve this high sensitivity, one needs to use dilute samples in which

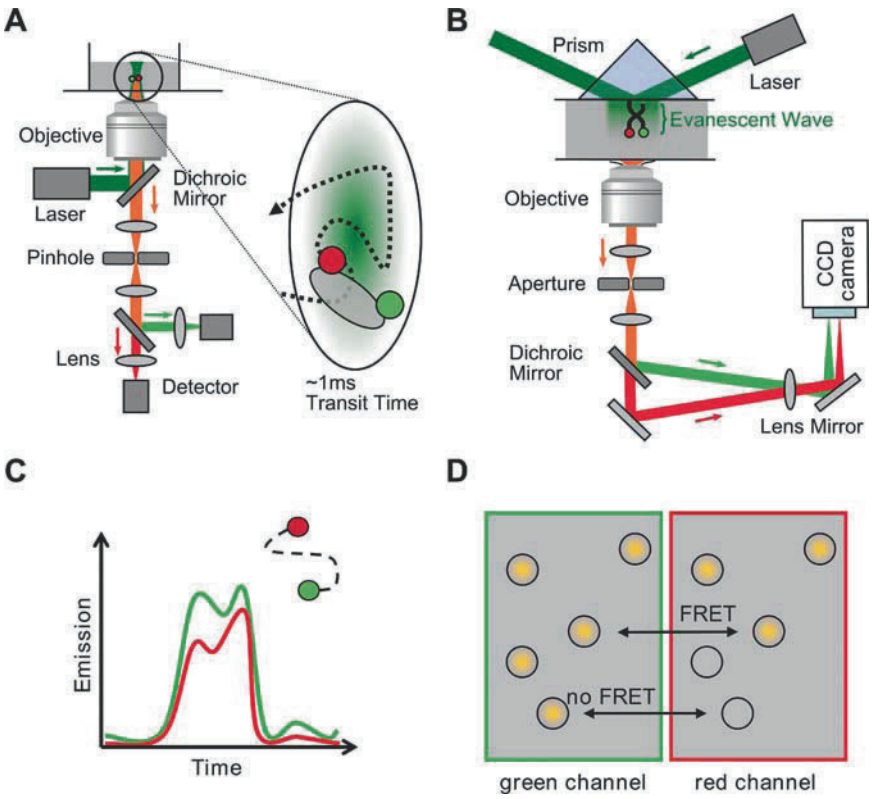


Figure 5.2. Instrumentation for single-molecule fluorescence resonance energy transfer (FRET). **A, B.** Simplified confocal microscopy setup for FRET detection of single molecules diffusing in solution and a setup for total internal reflection fluorescence (prism type) for imaging single molecules immobilised on a cover slide surface. Whereas in the confocal mode the avalanche photodiodes (APDs) act as point detectors of a diffraction-limited spot in the solution, the charge-coupled device (CCD) images color-separated molecules directly. **C.** A typical time trace obtainable for confocal detection in solution. If a molecule enters the focus with a donor and an acceptor attached to it, the detectable intensity in the green and the red detection channels increases. **D.** Immobilized molecules on a surface. If the molecule of interest is only labeled with a donor, no fluorescence will be visible in the red detection channel of the CCD camera.

individual molecules can be resolved; a small excitation/detection volume to reduce the contribution of the background; and bright and photostable fluorophores to achieve a high signal-to-noise ratio. All these features were included in the first successful detection of single-molecule FRET, reported in a pioneering paper by Ha et al. (1996). Near-field scanning optical microscopy (NSOM; a form of scanning probe microscopy that employs evanescent fields emerging from metal-coated fiber tips to probe fluorescence from immobilised single molecules) (Betzig and Chichester 1993; Betzig and Trautman 1992) was used to create a tiny excitation volume to resolve single-labeled DNA molecules dried on the surface of a microscope coverslip. The FRET efficiency was measured between a single tetramethylrhodamine fluorophore (TMR; acting as donor) and a single Texas red fluorophore (acting as an acceptor) attached to either end of a short, double-stranded DNA (dsDNA) fragment. The fluorescence was recorded either using two avalanche photodiodes (sensitive point detectors) dedicated to the donor-emission and the acceptor-emission channels or by a charge-coupled device (CCD)

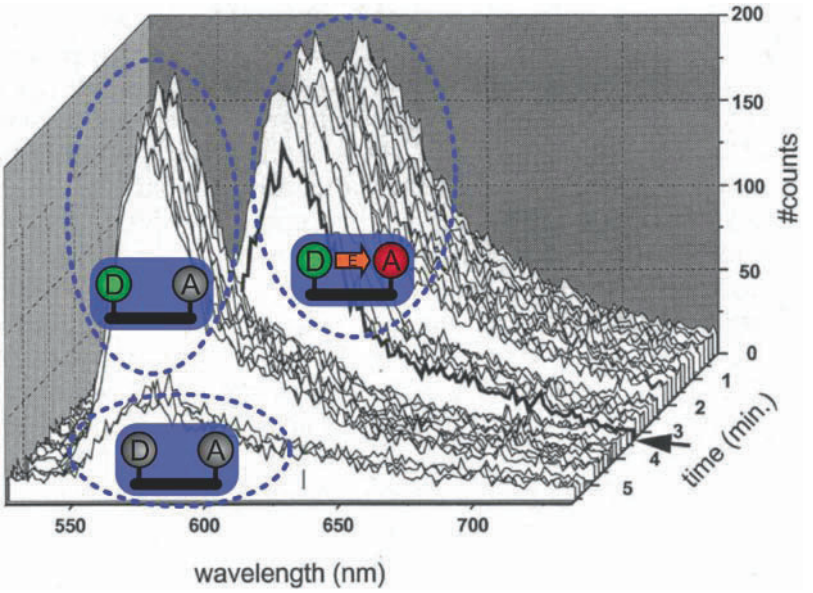


Figure 5.3. Time evolution of emissions during the first single-molecule fluorescence resonance energy transfer (FRET) observation. During the first minutes, the energy of the excited donor is transferred to the acceptor dye. After bleaching of the acceptor (around 3 min of detection), the donor intensity increases because no energy can be transferred to the acceptor. After 5 min the donor bleaches and the intensity is reduced to background levels (due to autofluorescence from the near-field scanning optical microscope tip and possibly fluorescence from distant molecules). See text for details. [Adapted from Ha et al. (1996). Copyright 1996 National Academy of Sciences.]

camera that recorded the emission spectra of individual molecules (Figure 5.3). The detection of single molecules was verified by single-bleaching steps of the fluorophores within the recorded emission spectra. If single molecules are present, the changes in the emission spectra are abrupt and represent either the "ON" or the "OFF" state of the donor and the acceptor, respectively. If, for example, the acceptor is photobleached and thereby becomes fluorescently inactive, the donor emission intensity increases and the acceptor emission vanishes. If the donor bleaches before the acceptor, both emission intensities decrease to the background level. Both cases allow the calculation of FRET efficiencies by using the integrated fluorescence intensities of the donor and acceptor. However, despite the superior resolution of NSOM compared to wide-field confocal microscopy, the NSOM approach requires sophisticated sample scanning and accurate positioning relative to the probed molecules to ensure effective excitation by the evanescent field. Moreover, the metallic tip may influence the emission properties of the fluorophores (Ambrose et al. 1994). Finally, such a method can only probe the surface of biological samples (such as cell membranes).

During this first demonstration of single-molecule FRET (smFRET), and, in general, in most single-excitation smFRET experiments (Figure 5.2A–C), one collects photons emitted in the donor- and acceptor-emission channels upon donor excitation for a smFRET pair under observation; we define these photon counts as $f_{D_{exc}}^{D_{em}}$ and $f_{D_{exc}}^{A_{em}}$, where $f_{D_{ex}}^{D_{em}}$ is the photon count for D-excitation-based D-emission and $f_{D_{ex}}^{A_{em}}$ is the photon count for D-excitation-based A-emission. The extent of FRET can be evaluated using a convenient expression of FRET efficiency, E^* (which does not account for spectral cross-talk in the acceptor-emission channel):

$$E^* = \frac{f_{D_{ex}}^{A_{em}}}{f_{D_{ex}}^{A_{em}} + \gamma f_{D_{ex}}^{D_{em}}} \tag{5.5}$$

Here γ is a detection-correction factor that depends on the donor and acceptor quantum yields and the detection efficiencies of the donor and acceptor emission channels; typically $0.5 < \gamma < 2$. A simpler expression that can still observe relative FRET changes assumes $\gamma = 1$; in that case, one recovers a *proximity ratio*, E_{PR}:

$$E_{PR} = \frac{f_{D_{ex}}^{A_{em}}}{f_{D_{ex}}^{A_{em}} + f_{D_{ex}}^{D_{em}}} \tag{5.6}$$

Where accurate FRET values are required (e.g., for evaluating donor–acceptor distances within biomolecular complexes), it is necessary to measure cross-talk and detection-correction factors (see Section 5.5.1).

In 1999, Ha et al. presented the first studies using a confocal microscope to observe temporal fluctuations of FRET at single-molecule level as a result of protein conformational changes. This study initiated a move from near-field to far-field excitation, which is the current standard for smFRET. It also marked a transition from studies in air to studies in biological buffers, which is necessary for relating smFRET results to the work from ensemble biochemical and biophysical experiments. The conformational changes in this study were modulated by ligand binding, which altered the rotational dynamics and the distances between the conjugated dyes (Ha et al. 1999b). A similar confocal microscope was also used to track conformational changes of individual three-helix-junction RNA molecules and changes induced by either RNA-binding proteins or Mg^{2+} (Ha et al. 1999c).

The use of confocal microscopy for single-molecule detection (Nie et al. 1994) with its high signal-to-noise ratio paved the way for the detection of single-molecule FRET in solution (Deniz et al. 1999). A series of 40-bp-long dsDNA fragments was labeled with a donor at the 5′ end and an acceptor at different positions along the DNA strand. For performing the measurements, a final concentration of ~30 pM was used. Based on the assumption of an excitation/detection volume of a confocal microscope of <1 fL (approximated as a rotational ellipsoid with a short axis of ~400 nm and a long axis of ~2 μm in diameter), the probability of single-molecule occupancy is ~1%, and thereby (following Poissonian statistics) it is several magnitudes larger than the probability of having >1 molecule at the same time within the focus. The paper showed the successful separation of different species within one sample due to the different FRET efficiencies of each species. Moreover, it showed changes in the FRET histograms arise due to DNA cleavage by a DNA-restriction enzyme causing the denaturation of DNA. Besides using intensity ratios for calculating FRET efficiencies, more sophisticated methods of detection and analysis use an array of parameters obtainable from fluorescence detection, such as fluorescence lifetimes and fluorescence anisotropies. This method was developed by Rothwell et al. (2003) to reveal the heterogeneity of HIV-1 reverse transcriptase (see Section 5.7.2).

Using confocal microscopy for the detection of single-molecule FRET has the advantage that the time range for studying dynamics is only limited by the time resolution of the single-point detector (up to nanoseconds). Therefore, confocal microscopy is an attractive tool for sorting molecules and monitoring biomolecular dynamics at the microsecond timescale. However, as a point-detection method, confocal microscopy typically detects only a single molecule at a given time. For parallel detection of multiple molecules, two-dimensional

detectors (e.g., CCD cameras) are necessary, which require the use of wide-field excitation. Two wide-field microscopy methods are commonly used: epifluorescence microscopy and total internal reflection fluorescence (TIRF) microscopy (Moerner and Fromm 2003). In epifluorescence microscopy, the laser excitation light is focused to the back- focal plane of the microscope objective, resulting in an expanded illuminated area that is collinated (Funatsu et al. 1995). In TIRF microscopy (Axelrod 2001), a collinated ray of light in a high–refractive index medium (e.g., glass) hits the interface with a lower–refractive index medium (e.g., water) at an angle larger than the critical angle for total reflection generating an evanescent wave in the second medium. The intensity distribution of the evanescent wave decays exponentially within a few hundred nanometres. For single-molecule detection, two TIRF excitation schemes can be used (Moerner and Fromm 2003). In the objective-type TIRF, the laser light is focused into an extreme edge of the back focal plane of the microscope objective. The prism-type TIRF uses an objective-independent excitation path and consists of a prism optically coupled to a cover glass (Figure 5.2B). In contrast to the epifluorescence wide-field illumination, TIRF excitation has the advantage of a reduced axial excitation volume, leading to lower background, but at the expense of limited time resolution as compared to point detection. The time resolution of ultra-sensitive electron-multiplying CCD cameras is limited to the CCD readout time, which is on the order of few milliseconds.

The first experimental realisation of TIRF-based single-molecule FRET examined the catalysis and folding of individual *Tetrahymena thermophila* ribozyme molecules (Zhuang et al. 2000) (see also Section 5.7.1). The ribozyme was labeled at both ends with a donor and an acceptor, surface-immobilized, and studied using either confocal scanning microscopy or TIRF microscopy. The fluctuations observed between two FRET states demonstrated the folding and unfolding of the ribozyme tertiary structure. Dwell-time analysis of both states produced equilibrium constants that were consistent with ensemble measurements (Zhuang et al. 2000). The studies revealed a rarely populated docked state (with a "fast" folding rate of 1 sec^{-1}), which was not observed using ensemble methods; this rate was discovered on top of the known longer rate of 0.016 sec^{-1} and a misfolded rate of 0.036 hr^{-1}, suggesting an additional folding pathway across a highly rugged energy landscape. Whereas the fast docking kinetics was measured using scanning confocal microscopy, the TIRF detection scheme was used to study the cleavage and overall folding kinetics for several hundred molecules simultaneously. Apart from its biological significance, this groundbreaking work included several aspects that are now the hallmarks of a successful use of a single-molecule technique to address a biological system and question. It was shown convincingly that the attachment of a donor and an acceptor to the large ribozyme does not alter the reaction kinetics of the ribozyme. Moreover, the surface immobilisation of the ribozyme did not affect the single-molecule reaction kinetics when compared to an ensemble kinetic. Finally, the single-molecule results showed directly the presence of distinct states and their dwell time; such complex analysis would simply have been impossible with ensemble measurements.

5.4. Single-Molecule FRET Methods Based on Alternating Laser Excitation

During the last few years, single-molecule FRET methods have been extended to allow direct probing of the FRET acceptor alternating laser excitation (ALEX) (Kapanidis et al. 2004) directly reports on the presence and state of both donor and acceptor fluorophores, by alternately exciting the sample at donor (*D*) and acceptor (*A*) excitation wavelengths.

The donor excitation provides the standard smFRET information, but it is the direct excitation of the acceptor that allows the relative fluorophore stoichiometry to be determined, virtual molecular sorting. ALEX (Figure 5.4D–E) provides smFRET with an additional photon count, for acceptor emission upon *direct* acceptor excitation the fluorescence stoichiometry ratio *S* (Kapanidis et al. 2004; Lee et al. 2005).

$$S = \frac{f_{D_{ex}}}{f_{D_{ex}} + f_{A_{ex}}} \tag{5.7}$$

The additional information is summarized in a two-dimensional histogram (Figure 5.4F). Sorting can remove artifacts that complicate FRET (such as the presence of states with inactive FRET donor or acceptor and the presence of complex fluorophore stoichiometries) while introducing new observables, such as the observation of an acceptor-only population, which is helpful for evaluating biomolecular interactions.

There are three complementary implementations of ALEX that among them cover a vast range of probe timescales, extending from nanoseconds to hours. Experiments on systems at chemical equilibrium (or systems undergoing kinetic changes occurring on the timescale of a few minutes) can be carried out using solution-based measurements and laser modulation at the microsecond timescale. This approach provides static snapshots of the FRET state of molecules as they diffuse through a femtolitre-size focal volume. This is the most straightforward implementation of the ALEX technique, and it has found numerous applications for measurements of the chemical kinetics of biomolecules (Kapanidis et al. 2005; Kapanidis et al. 2006; Lee et al. 2007).

For nonequilibrium experiments and minute-scale simultaneous observation of dynamics of multiple molecules, one can use the method of millisecond-ALEX (msALEX) (Margeat et al. 2006), which combines TIRF microscopy with laser modulation at the millisecond timescale. msALEX uses TIRF-illumination of surface-immobilised molecules combined with an alternating-laser scheme to carry out subpopulation sorting on surface-immobilised molecules, with the primary advantage over solution measurements of simultaneous observation of multiple smFRET molecules for extended periods of time (seconds to several minutes) (Margeat et al. 2006). One of the main advantages of msALEX over traditional surface smFRET is that by monitoring changes in the stoichiometry parameter *S* for time traces, it is possible to deconvolve fluctuations in FRET due to distance variations from fluctuations in fluorophore photophysics (such as acceptor blinking) (Margeat et al. 2006; Sabanayagam et al. 2004), a process that has been observed with many popular single-molecule fluorescence probes such as Cy5, tetramethylrhodamine, and Alexa647; blinking significantly increases the uncertainty on single-excitation smFRET measurements. Time traces of *S* can also report on the kinetics of assembly/disassembly reactions.

To probe fast dynamics and recover fluorescence-lifetime-based observables, one can use the method of nanosecond-ALEX or pulsed interleaved excitation (nsALEX/PIE) (Laurence et al. 2005; Müller et al. 2005), which uses alternation of pulsed laser sources at the nanosecond timescale. The time for a diffusing fluorophore to emit fluorescence after the excitation pulse is recorded and used to carry out fluorescence lifetime analysis in addition to ALEX-based analysis. This allows analysis of fast smFRET distance fluctuations, high-time-resolution fluorescence correlation spectroscopy (FCS), and time-resolved polarisation anisotropy measurements, which may be useful for increasing the accuracy of distance measurements extracted from FRET data (Laurence et al. 2005).

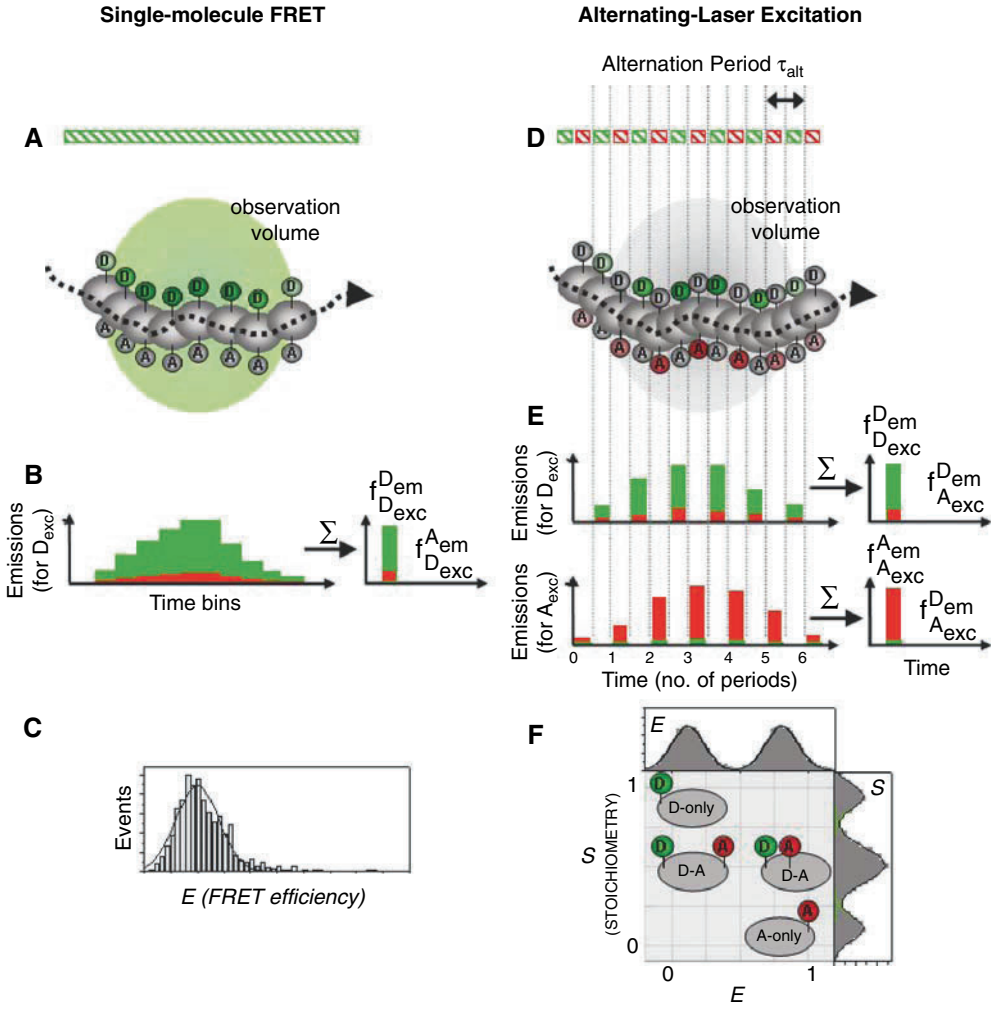


Figure 5.4. Concept of alternating-laser excitation (ALEX) and comparison with single-excitation, single-molecule fluorescence resonance energy transfer (FRET). **A–C.** Single-molecule FRET using single-laser excitation; diffusing-molecule example. A fluorescent molecule transverses a focused green-laser beam and emits photons at the donor- and acceptor-emission wavelengths. The photon counts at these two wavelengths are used to generate one-dimensional histograms of FRET efficiency, *E*. **D–F.** Single-molecule FRET using alternating-laser excitation. A fluorescent molecule transverses an observation volume illuminated in an alternating fashion using focused green- and red-laser beams. Using the photons emitted in the donor- and acceptor-emission wavelengths for each laser excitation, one can generate a two-dimensional histograms of FRET efficiency *E* and relative fluorophore stoichiometry *S*, enabling molecular sorting (see text for details).

5.5. Quantitative Single-Molecule FRET

5.5.1. Measuring Accurate FRET

To obtain *accurate FRET* values rather than proximity ratios, one can measure the contribution of excitation and emission cross-talk terms, the detection efficiency of each detection channel, and the quantum yield of the donor and acceptor fluorophores. The accurate E value is defined as

$$E = \frac{f^{FRET}}{\gamma f_{D_{ex}}^{D_{em}} + f^{FRET}} \tag{5.8}$$

where

$$f^{FRET} = f_{D_{ex}}^{A_{em}} - Lk - Dir \tag{5.9}$$

The terms *Lk* (donor leakage in the acceptor-detection channel) and *Dir* (acceptor emission in the acceptor-detection channel due to direct excitation by the donor-excitation laser) describe cross-talk contributions in the $f_{D_{ex}}^{A_{em}}$ signal; such cross-talk terms need to be subtracted from the $f_{D_{ex}}^{A_{em}}$ signal to recover the photon count associated only with FRET [Eq. (5.6)]. The Y term is the detection – correction factor from Eq. (5.5).

For solution-based measurements, these terms are obtained using μsALEX. The factor γ is directly related to the gradient on the *E*–*S* histogram of donor-acceptor labeled molecules of varying *E*, and *Lk* and *Dir* are obtained from the donor-only and the acceptor-only subpopulations, respectively. [for details see Lee et al. (2005)].

In the case of immobilised molecules, one follows the methodology of Ha et al. (1999a) or Edel et al. (2007), extracting γ and other correction factors by direct measurement of the change in emitted donor and acceptor intensity upon acceptor photobleaching, determined for individual molecules where the acceptor photobleaches before the donor. This can then be applied for all molecules in the experiment, allowing accurate E values to be obtained.

5.5.2. Obtaining Distances from Single-Molecule FRET Data

Once accurate FRET is obtained, Eq. (5.1) can be used to obtain interprobe distances. The major uncertainty on conversion of E to distances comes from uncertainty in the relative orientation of the smFRET pair, described by the orientation factor κ^2 (see Section 5.2). If κ^2 changes for a given sample (e.g., due to a change in rotational freedom of one of the fluorophores), R_0 will also change, and the distance value obtained will therefore be inaccurate (Dale et al. 1979). This can be addressed by measuring the fluorescence anisotropy (see Section 5.2), which provides a good estimate of the rotational freedom of the fluorophores. Ensemble anisotropy measurements (see Section 5.2) measure the anisotropy averaged over all subpopulations; for many applications, this will be an entirely sufficient check on the rotational freedom. However, ensemble measurements do not entirely address the possibility that rotational freedom is restricted for only some of the sample subpopulations, although in many cases this will still result in a noticeably high ensemble value, which will indicate that rotation is likely to be restricted to some extent. A very useful application of nsALEX/PIE is the ability to directly measure the rotational freedom of both fluorophores at a subpopulation level

(Laurence et al. 2005), thereby better estimating κ^2 and reducing the uncertainty on distance measurements obtained with smFRET.

5.5.3. Triangulation Methods

A promising application of smFRET methods is the use of multiple surface-based FRET measurements to accurately triangulate the position of part of a biomolecule in three dimensions (Andrecka et al. 2008; Rasnik et al. 2004). Multiple FRET-based distance measurements are made between an element with a known crystal structure, such as DNA helicase (Rasnik et al. 2004) or RNA polymerase (RNAP) (Andrecka et al. 2008), and a single point of unknown structure or dynamics (DNA for DNA helicase, mRNA for RNAP). The combination of multiple measurements allows the three-dimensional position of the unknown element to be determined relative to the known structure by simple geometry. In the case of DNA helicase, this provided important information on its DNA-binding orientation; in the case of RNAP, this allowed tracking of the exiting mRNA during transcription elongation. The use of more measurements than the minimum three required for triangulation allows self-consistency checks to be carried out on the measurements, the uncertainty on the measurements, particularly with respect to the orientation factor κ^2 (Rasnik et al. 2004).

5.6. Current Developments in Single-Molecule FRET

5.6.1. Multiple FRET Pair Methods

Single-molecule FRET typically consists of a single two-color donor–acceptor pair. However, a single pair will report only a single distance at a time. To probe several distances simultaneously, adding a third fluorophore is equivalent to performing three two-color experiments. This labeling scheme detects molecular interactions of up to three molecules and adds a third dimension for measuring conformational changes and dynamics of complex biomolecules.

The laboratory of Taekjip Ha was the first to extend three-color FRET to the single-molecule level (Hohng et al. 2004); this was achieved on DNA four-way junctions (see also Section 5.7.1) labeled with a donor (Cy3) and two acceptors (Cy5 and Cy5.5) on three arms of the junction. Three-color smFRET measurements were performed on surface-immobilised junctions using a single 532-nm laser for Cy3 excitation. Three avalanche photodiodes recorded the fluorescence intensities of single DNA molecule in a focused laser spot. By observing correlated intensity changes among three fluorophores, the group was able to probe the well-synchronized movements of two of the acceptor arms to and from the third donor arm and to show that nonparallel conformers of the junctions are dominant. Three-color FRET using a single excitation laser has also been applied to freely diffusing single molecules to simultaneously monitor mixtures of triply labeled (Alexa488, TMR, and Cy5) DNA duplex at multiple distances (Clamme and Deniz 2005).

Use of single-laser donor excitation can pose problems with distinguishing different FRET pathways among the three fluorophores. Moreover, multicolor FRET involves the selection of fluorophores with overlapping donor emission and acceptor excitation spectra, depending on the interprobe distances probed and the R_0 for the different FRET pairs. If

there is large spectral overlap between the dye pairs (such as acceptors, Cy5 and Cy5.5), careful selection or custom-made dichroic mirrors and filters are crucial for separating their signal. Because there is significant cross-talk into the acceptor channels, control experiments have to be conducted to determine their cross-talk and correct for them. To address such challenges and monitor multiple accurate distances without previous information about the molecular structure, Lee et al. (2007) introduced alternating-laser excitation with three-color FRET (3c-ALEX) of diffusing molecules. Using alternating-laser excitation for each of the fluorophores on a triply labeled DNA duplex (Alexa488, TMR, Alexa647), they were able to sort molecules of a single sample measurement into three-dimensional stoichiometry and FRET histograms (Figure 5.5). The stoichiometry histogram identifies the molecules according to their fluorophore stoichiometry, that is, singly, doubly, or triply labeled molecules, independent of FRET efficiency. The FRET histogram depends on the three intermolecular distances of each dye pair, improving the resolution of conformational heterogeneity compared to a one-dimensional FRET histogram. The authors applied 3c-ALEX to study the translocation of *Escherichia coli* RNAP on DNA. They showed that with a single system consisting of a singly labeled RNAP and a DNA doubly labeled at both ends surrounding the start site, there is a concomitant change in the FRET efficiency of the fluorophore pairs arising from the downstream translocation of RNAP on DNA. The method of 3c-ALEX was also used recently to study the structure and folding of a triply labeled 8-17 deoxyribozyme in solution (Lee et al. 2007); the group showed that the deoxyribozyme folds to form a pyramidal structure upon adding Mg^{2+} by monitoring three-color FRET changes with increasing concentrations of Mg^{2+}.

A major challenge for multicolor setups is to have a high signal-to-noise ratio in all detection channels, and this has been limited by tri-band dichroic mirrors with limited transmission wavelengths and broad reflective regions. To improve the signal-to-noise ratio of multicolor ALEX setups, Ross et al. (2007) used a programmable acousto-optical beamsplitter with user-defined wavelengths to achieve the same detection efficiency as single-color setups. Using DNA three-way junctions hybridized on surfaces, the authors resolved seven subpopulations with different stoichiometry, quantified FRET efficiencies in the presence of competing FRET pathways, and observed correlated molecular dynamics.

5.6.2. Combinations of Single-Molecule FRET with Other Single-Molecule Methods

Combining smFRET with methods allowing the manipulation of single molecules by applying external forces opens a new field of exciting applications. A range of such applications was proposed in 1999 by Shimon Weiss in a visionary article (Weiss 1999); among other combinations, methods that combine FRET with patch-clamp, atomic force microscopy (AFM), and optical/magnetic tweezers were suggested. A number of recent reviews also discussed the prospects of combination methods and possible applications (Deniz et al. 2008; Greenleaf et al. 2007; Walter et al. 2008).

The combination of FRET and electrical recording of single-ion channels was shown in 2003 by Borisenko et al. (2003 and Harms et al. (2003); both groups studied the heterodimer formation from gramicidin monomers in a lipid bilayer. Whereas a FRET signal can give information about the distance between differently labeled monomers, the electrical recording allows validation of ion channel formation at the same time.

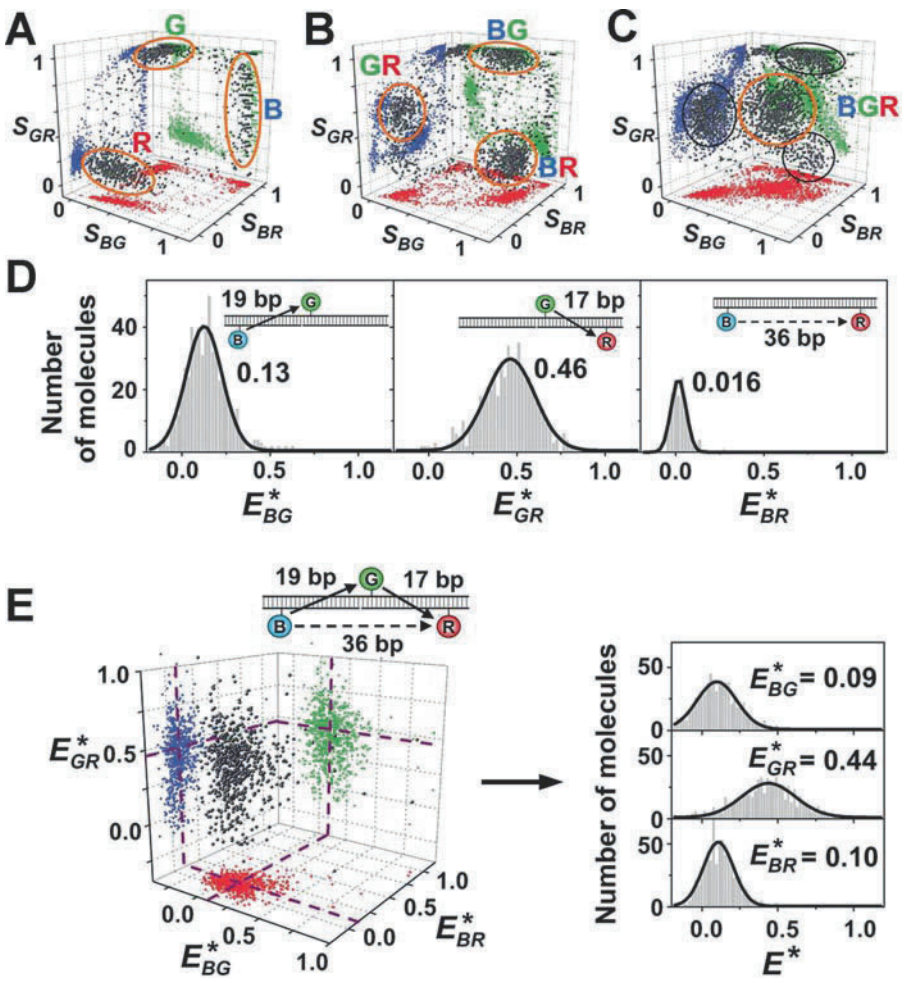


Figure 5.5. Molecule sorting based on probe stoichiometry *S*. **A.** Three-dimensional *S* histograms for a 1:1:1 mixture of B-only, G-only, and R-only. **B.** Three-dimensional *S* histograms for a 1:1:1 mixture of three doubly labeled species (B-G, G-R, and B-R). **C.** Three-dimensional *S* histograms of triply labeled species (B-G-R). The cluster at the center of the histogram (*orange ovals*) corresponds to B-G-R. Black ovals surround clusters corresponding to the doubly labeled species. **D, E.** Measuring distances within triply labeled and doubly labeled species in the same solution. The reported E^* values represent the means of the fitted Gaussian distributions. Panel D shows the one-dimensional E^* histograms for B-G, G-R, and B-R species, selected from the black circles in panel C. Panel E shows the three-dimensional E^* histograms for the triply labeled species B-G-R, selected from the orange circles in panel C, and the one-dimensional E^* histograms obtained after collapsing the three-dimensional histogram on each of the E^* axes.

An instrument capable of combining FRET with optical trapping was suggested in 2003 by Lang et al. (2003) and was demonstrated experimentally in 2007 by two groups (Hohng et al. 2007; Tarsa et al. 2007). Tarsa et al. (2007) described the use of optical traps mechanically to unzip DNA hairpins by simultaneously monitoring the conformational states of the DNA using FRET, whereas Hohng et al. (2007) extended the method by watching DNA four-way junctions (see also Section 5.7.1), allowing the detection of nanometre-scale motion at sub-pN forces.

Whereas optical traps allow only the translational movement of a loaded bead, magnetic traps can be used to rotate the paramagnetic bead. In addition, using magnetic tweezers prevents photophysical effects that can arise from optical traps. The combination of FRET and magnetic tweezers (Shroff et al. 2005) used a single-stranded DNA oligomer to measure the forces internal to a small DNA loop, showing the possibility of tuning the sensor's force response by using the distances between the dyes.

Single-Molecule FRET in Living Cells

Despite the potential of single-molecule FRET as a noninvasive technique for *in vivo* studies, there have been few *in vivo* single-molecule FRET reports, mainly due to a series of experimental challenges. First, the background autofluorescence from the cell interferes with the detection of fluorescence emission from labeled biomolecules. Second, due to size and labeling requirements, typical fluorophores used in living cells (organic fluorophores and fluorescent proteins) are of moderate brightness and photostability, leading to low signal-to-noise ratio and short photobleaching lifetimes, respectively. Premature bleaching of the fluorophores may lead to misinterpretation of the molecular interactions and dynamics. Semiconductor quantum dots (Michalet et al. 2005)—novel probes displaying tunable emission wavelengths, superior brightness, and longer photobleaching rates—may provide an alternative for *in vivo* labeling; however, quantum dot internalisation and specific targeting to sites or proteins of interest is not trivial. Third, fluorophore labeling methods in living cells can be rather complex and often not as specific (Kapanidis and Weiss 2002).

In vivo applications of smFRET have been limited to cell membrane proteins because these proteins are accessible for fluorescent labeling and molecular interactions. Moreover, fluorescent molecules on membranes can be visualised using TIRF microscopy (see Section 5.2),which does not suffer as much as epifluorescence from out-of-focus background and concomitant decrease in the signal-to-background ratio. Along these lines, one of the first examples of single-molecule FRET examined epidermal growth factor receptor (EGFR) signaling in living cells using objective-type TIRF microscopy equipped with dual-view optics (Sako et al. 2000). The study tracked single molecules of Cy3-labeled epidermal growth factor (EGF) in the plasma membrane of A431 carcinoma cells and observed that EGFR dimers were preformed before the binding of the second EGF molecule; this observation was confirmed by adding a mixture of EGF–Cy3 and EGF–Cy5 and measuring their FRET signal as they colocalized on the membrane (Figure 5.6). The study also used a labeled monoclonal antibody specific to the phosphorylated EGFR (Cy3–mAb74) and Cy3–Cy5 coincidence analysis to show that EGFR becomes phosphorylated after dimerisation; this conclusion was reached by observing the colocalization of Cy3–mAb74 with spots of Cy5–EGF with double the fluorescence intensity. Most recently, smFRET and ensemble FRET based on fluorescence lifetime imaging (FLIM-FRET) were used to determine the three-dimensional architecture of high-affinity and low-affinity EGFR–EGF complexes in cells (Webb et al. 2008). By measuring the inter-EGF distances within discrete EGF pairs and the vertical distance from EGF to the plasma membrane, structural and orientation differences in the high- and low-affinity EGFR complexes have been revealed. These findings could explain how EGFR oligomers achieve multiple levels of signaling in normal and diseased cells. smFRET has also been used to detect the binding of BodipyTR-GTP to a G protein Ras-YFP in living cells (Murakoshi et al. 2004), a demonstration that smFRET could be useful for studying the dynamics and activation of G proteins in living cells.

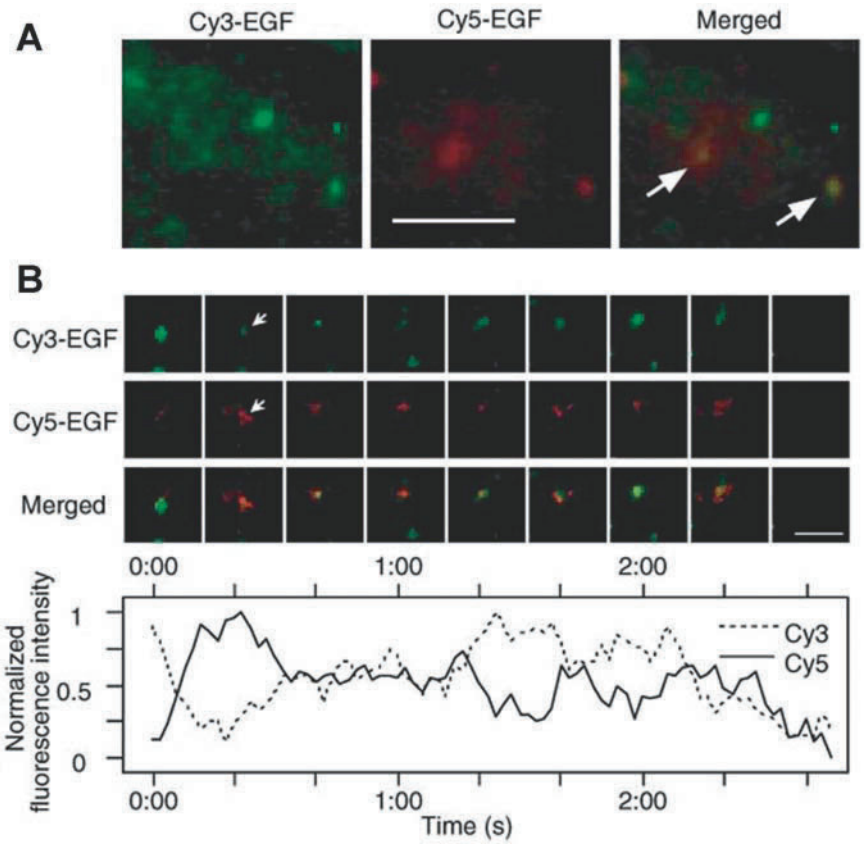


Figure 5.6. Single-molecule fluorescence resonance energy transfer (FRET) between Cy3–epidermal growth factor (EGF) and Cy5–EGF on the surface of living cells. **A.** A mixture of Cy3–EGF and Cy5–EGF was added to A431 cells and imaged with total internal reflection fluorescence microscopy and dual-view optics. Fluorescent spots shown in the Cy5 channel appeared as a result of FRET from Cy3–EGF to Cy5–EGF (*arrows*), indicating interaction between single molecules. **B. Top:** Fluorescence intensity over time of a single spot (*arrow*) was measured in both Cy3 and Cy5 channels. The diffusion of the spot was slow, probably because of slow Brownian motion of EGF receptor (EGFR) bound to the cytoskeleton. **Bottom:** Anticorrelation between Cy3 and Cy5 fluorescence intensities indicated FRET occurring due to the binding of Cy3– and Cy5– EGF on the same EGFR dimer. [Adapted from Sako et al. 2000. Copyright Nature Publishing Group.]

5.7. Applications of Single-Molecule FRET to Biomolecular Systems

Since the first single-molecule FRET experiment measured on doubly labeled DNA molecules immobilized on a glass surface (Ha et al. 1996), smFRET has seen an escalation in applications on biomolecular systems within the last decade. Based on the principle of acting as a nanoscale ruler at 2–10 nm, it can map out molecular distances in three dimensions; monitor conformational changes in real time, and catch molecular motors in action. In this section, we discuss several biomolecular systems to which smFRET has been applied. Due to space limitations, this survey cannot be comprehensive; we discuss mainly applications of single-molecule FRET to the structure, dynamics, and mechanism of nucleic acids, protein–nucleic acid complexes, and selected proteins (molecular motors and protein-folding model proteins). We should mention, however, that smFRET has been used to make contributions to

the understanding of many mechanisms and molecular devices, such as translation (Blanchard et al. 2004a, b), DNA processing (Smiley et al. 2007), homologous recombination (Joo et al. 2006), transcription regulation (Morgan et al. 2005; Schluesche et al. 2007), nucleosome dynamics (Koopmans et al. 2007), membrane fusion (Margittai et al. 2003), and energy transfer in photonic wires (Heilemann et al. 2004).

5.7.1. Applications to Nucleic Acids

Holliday Junctions

One of the first systems extensively studied using single-molecule FRET is the Holliday junction, a four-way DNA junction that constitutes a central intermediate of recombination and DNA replication. The Holliday junction is initially formed after strand exchange between two homologous sequences of two DNA strands; it subsequently undergoes branch migration, a process in which the initial heteroduplex DNA is extended by movement of the junction along DNA. Prior to 2003, it was known that in the absence of Mg^{2+}, the Holliday junction adopts an open structure in which the four helices point to the corners of a square. It was also known that in the presence of Mg^{2+}, the junction folds into a more compact stacked-X structure, in which two neighbouring helices undergo coaxial stacking, leading to two possible stacked structures. However, only one of the stacked structures had been observed using X-ray crystallography; moreover, there was no information about possible transitions between the stacked conformers and no clear understanding of branch migration and its coupling to the transitions between stacked and open conformations of the Holliday junction. A major obstacle for such studies was the inability to synchronise such stochastic conformational changes; single-molecule FRET offered an elegant and direct way to address this limitation and shed light on the thermodynamics and kinetics of the process by observing dynamic transitions between stacked conformers, as well as branch migration.

The first report on the dynamics of Holliday junctions used “immobile” Holliday junctions (junctions with nonhomologous sequences that prevent branch migration) labeled at the ends of two of the helices with a donor and acceptor. Use of confocal microscopy on individual surface-immobilised Holliday junctions led to direct observations of switching between low- and high-FRET states that correspond to the stacked conformers. Analysis of the dwell times at the two states provided the rate constants of the transitions, which were Mg^{2+} concentration dependent, as expected (Hohng et al. 2004; McKinney et al. 2003).

The initial studies on immobile junctions did not report on branch migration. Towards such a goal, McKinney et al. (2005) designed DNA sequences that allow a single–base pair step of branch migration; again, the labeling strategy led to distinct FRET efficiencies for the two stacked conformers. It is intriguing that FRET time traces showed that the junction interconverted between FRET states at two distinct rates (fast and slow), with the rate dependent on the sequence of the central part of the junction; the switch from one rate to the other was interpreted as a branch migration step (since the step changes the central sequence context and therefore the interconversion rate). As previously, the transition rate depended on Mg^{2+} concentration and on the G-C versus A-T content of the central sequence, with the latter having fewer transitions between stacked conformers before branch migration. One surprising finding was that different branch points had more than a 30-fold variation in their lifetimes, which led to the proposal of a highly sequence-dependent energy landscape for spontaneous branch migration. An independent smFRET study on mobile Holliday junctions also studied real-time branch migration over a 5-bp homology region, establishing that branch migration

proceeds in a stepwise pattern and that each step could represent more than a single base pair, with Mg^{2+} modulating the rate of branch migration (Karymov et al. 2005; Karymov et al. 2008a, b). This swift progress in the understanding of the mechanism of branch migration will certainly be followed by extensive analysis of the effect of proteins that recognise and process this important DNA intermediate.

RNA Folding

smFRET was also used extensively in studies of RNA folding. One form of single-stranded RNA folds to form secondary structures with catalytic activity called ribozymes or RNA enzymes. Ribozymes catalyze numerous cellular processes, such as translation and RNA splicing, and are used in biotechnology as therapeutic agents and in functional genomics. Their catalytic activity of performing site-specific cleavage and ligation reactions depends on their folded structures and their interactions. Various biochemical and ensemble techniques have discovered that ribozymes undergo multiple folding pathways and intermediate states to achieve the native structure. However, the characterization of many transient intermediate states was not accessible by ensemble methods that probed only well-populated, accumulative intermediate states. Furthermore, large multidomain ribozymes have complex folding pathways and interactions that make the characterization of each pathway difficult with ensemble methods. smFRET offers additional information that ensemble methods lack by allowing the observation of structural dynamics of transient intermediate states and subpopulations in a heterogeneous system. The first smFRET study on the dynamics of RNA folding used a large multidomain ribozyme (Zhuang et al. 2000) and was discussed in Section 5.3.

Fast conformational kinetics of a two-way junction hairpin ribozyme cleavage pathway was also determined for the first time with smFRET (Zhuang et al. 2002). Based on the use of labels at both ends of the helices, the FRET time traces revealed the folding and unfolding transitions in real time, and their kinetics was determined to have one folding rate of 0.008 sec^{-1} and four other unfolding rates. In addition, the time traces showed a "memory effect" by which similar dwell times were repeated in the folded state and rarely switched between different folded states after unfolding, suggesting four different folded and unfolded states, in which three of them were too few populated and too fast to be detected with ensemble methods. The four-way junction hairpin ribozyme was found instead to have a folding rate that was three to four orders of magnitude faster, with both heterogeneous folding and unfolding kinetics dependent on Mg^{2+} concentration (Tan et al. 2003). FRET time traces provide the history of a given state and allowed the finding of a previously unknown intermediate with a proximal conformation that was obligatory for the folding of the hairpin ribozyme. This intermediate works by folding the loops in close proximity, increasing the probability of interaction and hence speeding up the folding rate. smFRET has demonstrated the robustness of obtaining highly heterogeneous structural and kinetic information of complex RNA folding pathways and rugged energy landscapes, which had been largely an unexplored field.

DNA Nanomachines

DNA nanomachines are examples of DNA nanotechnology (Bath and Turberfield 2007; Rothemund 2006; Seeman 2003; Seeman 2005), a branch of synthetic biology that uses the unique properties of DNA (ease and programmability of DNA hybridization, rigid and

well-understood structure at the 10- to 50-nm scale, facile synthesis and manipulation) to assemble artificial molecular machines that can generate motion and accomplish a specific task as a response to a specific signal. The nanometer measurement scale of FRET is particularly suitable to study the conformational changes of such machinery.

Recently, a novel DNA nanomachine has also been studied with smFRET using microsecond ALEX (Section 5.4). Specifically, Goodman et al. used the specificity of DNA hybridization to prepare self-assembling DNA tetrahedra (Goodman et al. 2004; Goodman et al. 2005) that can reversibly change shape in response to specific signals (Goodman et al. 2008). In the reconfigurable design, one of the edges of the tetrahedron contains a hairpin loop. In the absence of "fuel" the loop remains closed, producing a 10-bp-long edge. When the "fuel" (a single-stranded DNA strand complementary to the hairpin loop) is added, it hybridises to both halves of the hairpin, producing a substantially extended edge (30-bp long). This cycle is reversible, as shown using polyacrylamidegel electrophoresis (PAGE) and ensemble-FRET measurements; the addition of an "antifuel" hairpin complementary to the fuel strand binds and displaces the fuel from the tetrahedra and returns them to the closed state. This cycle can be repeated several times. The FRET measurements were carried out by labeling each end of the hairpin edges with donor and acceptor fluorophores so the closed state would give high FRET and the open state would give low FRET. However, this did not exclude the presence of either static heterogeneity, where the tetrahedra could occupy a range of static partially open states; or dynamic heterogeneity, where the tetrahedra could undergo transient opening. To address this issue, μsALEX was used to measure FRET for the initial conversion to closed and then open tetrahedra. Each state showed a single homogeneous population at expected FRET values, showing that dynamic interconversion was not present on any significant scale.

Müller et al. (2006) used nsALEX/PIE to characterise DNA tweezers (Section 5.4), one of the first DNA machines (Yurke et al. 2000). The principle of DNA tweezers is as follows: Initially the tweezers are in the open, low-FRET state; upon addition of a closing stand ("fuel") that hybridises with the tweezers, the tweezers convert to the closed, high-FRET state. The addition of fuel may produce several DNA nanostructures other than the desired DNA tweezers; such nanostructures will have different FRET efficiencies and stoichiometries difficult to study using ensemble methods because their FRET efficiencies are expected to average out at the ensemble level. However, it was elegantly shown in a single diffusion-based nsALEX/PIE measurement that whereas open tweezers exist in a single conformation, the addition of fuel generated three conformations corresponding to various states of "closed" tweezers. These closed states were tentatively assigned to some of the states seen in the FRET histogram generated by nsALEX/PIE.

These experiments illustrate well how DNA nanotechnology and smFRET methods are natural partners because both are essentially concerned with the properties of single molecules, smFRET can be used to directly report on the structure and function of DNA nanostructures, and this information can then be used to inform the development and design of new nanostructures.

5.7.2. Applications to Nucleic Acid Machines

DNA and RNA Helicases

smFRET methods have proven to be a powerful tool for probing the structure and mechanism of helicases in ways not possible with ensemble methods. Helicases are motor

proteins that translocate and unwind double-stranded nucleic acids into their complementary single strands using energy derived from adenosine triphosphate (ATP) or nucleotide hydrolysis. These enzymes are crucial for many cellular processes, such as DNA replication, transcription, repair, and recombination. Whereas the knowledge of helicase mechanistic properties has generally been determined by pre–steady-state kinetic ensemble studies of single turnovers, they only provide overall kinetic information of the unwinding and translocation of DNA, two core mechanisms of the helicase. smFRET can provide high base pair resolution ($\leq$10 bp), resolving the problem of low helicase processivity *in vitro*, and can probe detailed mechanistic processes not detectable by ensemble methods, such as pausing, rewinding, and repetitive shuttling.

The earliest smFRET work on helicase was the investigation of unwinding of double-stranded DNA by *E. coli* Rep helicase (Ha et al. 2002). Based on the use of a donor and acceptor–labeled DNA duplex or biotinylated-Rep immobilized on a surface and imaged using TIRF microscopy, FRET fluctuations showed the translocation of Rep monomer along single-stranded DNA (ssDNA) using ATP hydrolysis, stopping at the junction of ssDNA and dsDNA. At increasing Rep concentrations, DNA unwinding was initiated only with the binding of additional Rep to form a functional oligomer, which was indicative of the rate-limiting step. This result led to the explanation of limited unwinding processivity *in vitro*, in which frequent stalling of the Rep helicase was observed from a partially dissociated Rep oligomer. The stalled state could then proceed in two ways: rewinding of dsDNA upon complete dissociation of Rep or reinitiation of unwinding upon formation of functional Rep oligomer. This revealed new mechanistic information about helicase and raised the questions about the connection of the reactions to the structure or conformations at each reaction state. Crystal structures of Rep helicase showed two different conformations, open and closed, implying the structural flexibility of the helicase. The binding orientation of Rep to DNA was determined, using smFRET, and from powerful triangulation methods of eight different donor-labeling sites and an acceptor labeled at the ssDNA/dsDNA junction, the Rep monomer was found mainly to be in the closed conformation when bound to the DNA junction, a conformation representing an inactive state of the Rep helicase (Rasnik et al. 2004). Single-molecule FRET was also used to identify an unexpected translocation mode for Rep helicase: its repetitive shuttling motion on DNA (Myong et al. 2005). The FRET increased as a donor-labeled Rep at the 3′ end of ssDNA translocated towards the acceptor at the 5′ end (Figure 5.7). As the Rep monomer stopped translocation when it encountered the ssDNA/dsDNA junction, it snapped back to its initial position and repeated this shuttling, producing a saw-toothed pattern in the FRET time trace. The functionality of the repetitive shuttling was proposed to keep the ssDNA clear of unwanted proteins such as toxic recombination intermediates.

Helicases can be characterized by their step sizes upon translocation or unwinding of DNA or RNA. Single-molecule FRET was used to study the RNA-unwinding mechanism by nonstructural protein 3 (NS3; an essential helicase for replication of hepatitis C virus) and showed that NS3 unwinds DNA (used instead of RNA due to ease of synthesis) at discrete steps of 3 bp. In the dwell-time analysis of each FRET state, the unwinding proceeded in a proposed spring-loaded mechanism where domains 1 and 2 of NS3 translocated three hidden steps consisting of one ATP hydrolysis each while domain 3 remained fixed. This accumulated tension on the NS3–DNA complex, which was released by unwinding abruptly in a 3-bp step (Myong et al. 2007).

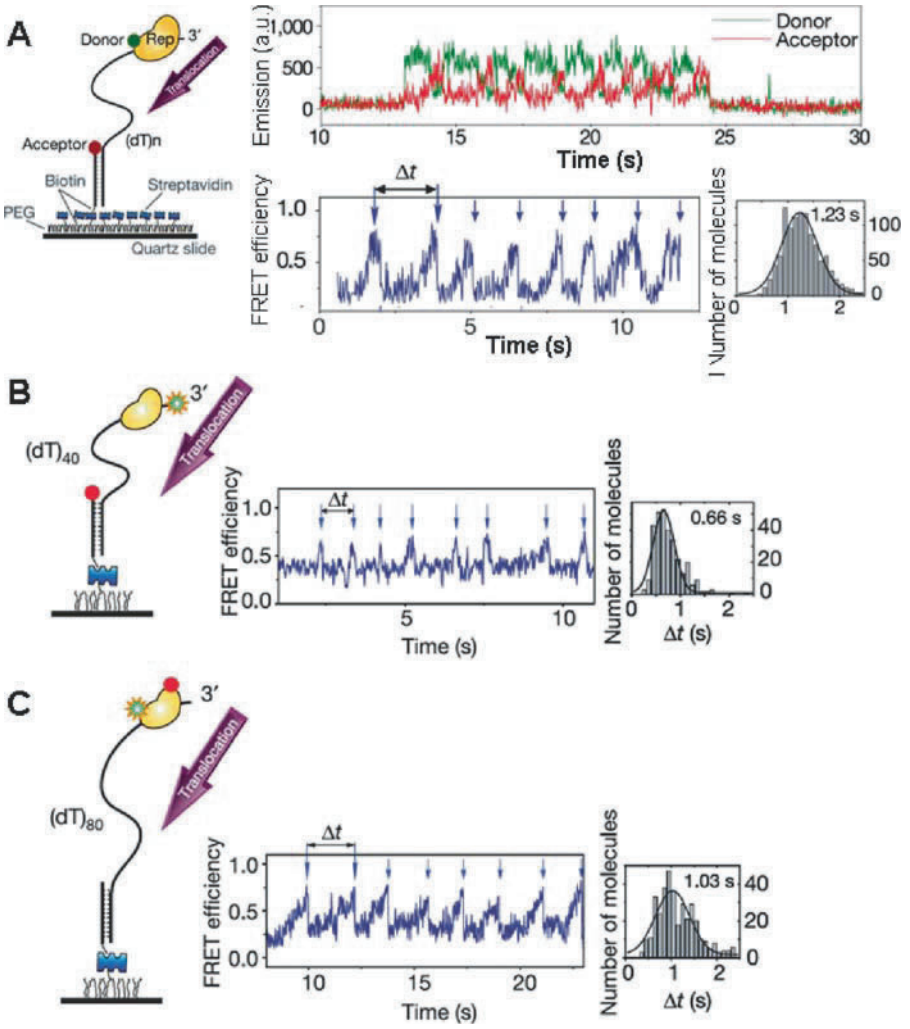


Figure 5.7. Repetitive shuttling of Rep Helicase on single-stranded DNA (ssDNA). Rep helicase was donor-labeled and applied to acceptor-labeled ss/double-stranded (ds) DNA molecules immobilized on a PEGylated surface. Fluorescence intensity time traces displayed anticorrelation between green and red intensity and the fluorescence resonance energy transfer (FRET) trace, pointing to Rep translocation towards the acceptor (gradual increase in FRET), followed by an unexpected snapback (sharp drop in FRET) as it hits a blockade at the junction of ss/dsDNA. **B.** Mechanism of repetitive shuttling. An unlabeled Rep was added to DNA labeled at both ends. The FRET trace showed regular spikes of high FRET, demonstrating the forming of a temporary ssDNA loop as Rep is simultaneously bound to the 3′ end and the ss/dsDNA junction. **C.** Conformation of Rep. Rep helicase was labeled with both a donor and an acceptor. The FRET trace showed a gradual increase as Rep gradually adopts an open conformation as it translocates towards the junction, then closes abruptly as it snaps back. PEG, poly(ethylene glycol). [Adapted from Myong et al. 2005. Copyright 2005 Nature Publishing Group.]

DNA Polymerase

Since the completion of the human genome sequence using the Sanger sequencing method and the subsequent launch of many other genome projects, there has been a race to develop faster, cheaper, and more sensitive sequencing methods. The specificity and reproducibility of DNA polymerase (DNAP) enzyme can be exploited to image sequence information on a DNA template during the synthesis of the complementary strand, as shown by several recent examples (Guo et al. 2008; Harris et al. 2008). The first report that demonstrated DNA sequencing by smFRET performed cyclic synthesis by DNAP using TIRF microscopy (Braslavsky et al. 2003). The sequencing assay involved a surface-immobilised DNA template with an annealed donor-labeled primer and the flowing of DNAP and labeled nucleotide triphosphates while the DNA template positions monitored for FRET signal. The reagent exchange for smFRET sequencing proceeded as follows: The position of the donor-labeled DNA primer was recorded and the donor was bleached before the first round of incorporation. This round consisted of donor-labeled nucleotides deoxyuridine triphosphate (dUTP)–Cy3 or deoxycytidine triphosphate (dCTP)–Cy3 with DNAP, which was flowed in and checked for donor signal. Unlabeled deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), and DNAP were then flowed in to extend the primer until the next A or G. If there was no donor signal detected for the first incorporation, this process was repeated. Otherwise, the second round of incorporation used acceptor-labeled nucleotides dUTP–Cy5 and dCTPxCy5 and checked for FRET signal, followed by bleaching the acceptor before flowing unlabeled dATP, dGTP, and RNAP. Green and red lasers were alternated to excite the donor, check for FRET, and bleach the acceptor. This whole cycle was repeated several times. This method of sequencing DNA allowed the determination of sequence fingerprints up to 5 bp in length with 97% confidence level. However, this FRET assay was limited by the dynamic range of FRET at 5 nm, thus restricting the read length equivalence to $\sim$15 bases. Instead, labeling DNAP with a donor and monitoring the FRET signal as it incorporates acceptor-labeled nucleotides may overcome this constraint. Nevertheless, Braslavsky et al.'s pioneering work on using DNAP and FRET between labeled nucleotides has been instrumental for the development of subsequent generations for DNA sequencing.

RNA Polymerase

RNAP is the multifunctional machine at the heart of transcription, one of the most important biological processes in gene expression and regulation. RNAP works by recognizing specific sites on the DNA, unwinding the DNA helix around the start site, and transcribing the DNA by synthesizing a complementary RNA strand. There have been numerous studies on the structure and function of RNAP and the mechanism of transcription at initiation, elongation, and termination. Recent papers have reviewed extensively the studies of RNAP transcription with single-molecule methods (Bai et al. 2006; Herbert et al. 2008).

One of the first questions addressed using single-molecule FRET involved the fate of the bacterial transcription initiation protein σ^{70}. Early observations of σ^{70} absence in RNAP–DNA elongation complexes led to proposals that σ^{70} release is required for the transition from transcription initiation to elongation and that the molecular machinery in transcription is essentially different between initiation and elongation. Kapanidis et al. (2005) used single-molecule FRET combined with ALEX to measure the extent of σ^{70} retention at various points along elongation. The ALEX-based σ^{70} release assay was based on observations of changes

in intermolecular distances and binding stoichiometries upon escape to elongation. The assay uses analytical sorting to distinguish between σ^{70} retention and σ^{70} release in elongation: For a leading-edge FRET assay (which uses a donor introduced at the leading edge of RNAP and an acceptor at the downstream end of DNA), forward translocation and formation of a σ^{70}-containing elongation complex converts a donor–acceptor species with low FRET to a donor–acceptor species with high FRET; on the other hand, formation of σ70-free elongation complex with release of σ^{70} converts a donor–acceptor species with low FRET to donor-only and acceptor-only species, which can be identified using the stoichiometry dimension of ALEX histograms. The results show clearly that σ^{70} release is not obligatory but stochastic, with 70%–90% of early elongation complexes and 50%–60% of mature elongation complexes retaining σ^{70}. The long half-life of σ^{70} in mature elongation complexes suggested that some complexes might retain σ^{70} throughout elongation, allowing multiple additional levels of gene regulation with σ^{70}.

The mechanism of initial transcription in which RNAP translocates relative to DNA had also been controversial. Based on ensemble biochemical results, there were three proposed models that attempted to describe the physical mechanism of the process: first, the model of transient excursions, which invoked transient cycles of forward/reverse translocation of RNAP; second, the model of RNAP inchworming, which invoked a flexible element in RNAP that translocates downstream during abortive RNA synthesis and retracts upon release of the abortive RNA; and third, the model of DNA scrunching, which invoked a flexible element in DNA that was pulled in by a fixed RNAP during translocation and extruded upon release of the abortive RNA (Figure 5.8A). The three models were tested by monitoring the open complex and initial transcribing complexes labeled at different positions on RNAP and on promoter DNA (Kapanidis et al. 2006). Based on the changes in FRET efficiencies, it was shown that the only distances changing during abortive transcription were distances between DNA downstream of the transcription bubble and the RNAP leading edge, as well as between

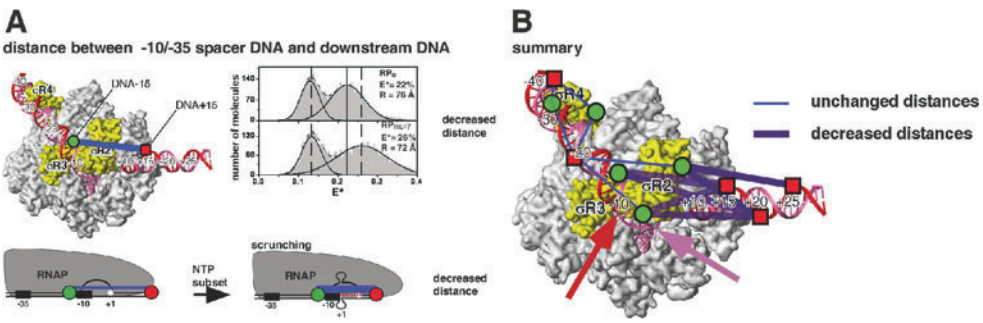


Figure 5.8. Initial transcription involves DNA scrunching. **A.** Experiment documenting the contraction of DNA between positions –15 (Cy3b) and +15 (Alexa647). Subpanels show E^* histograms of open complex (RP_o) and initial transcribing complexes with up to 7-nt RNA ($RP_{itc,\leq 7}$). The histograms comprise free DNA (lower-E^* species) and the RNA polymerase (RNAP)–DNA complexes (higher-E^* species attributable to RNAP-induced DNA bending). Free DNA is present in all experiments. The increase in the fluorescence resonance energy transfer (FRET) efficiency in $RP_{itc,\leq 7}$ (compared to RP_o) point to DNA compaction during abortive initiation, consistent with the predictions of the DNA-scrunching model (see text for details). **B.** Structural model of RP_o showing all donor–acceptor distances monitored. Thin blue lines represent distances that remain unchanged on transition from RP_o to $RP_{itc,\leq 7}$. Thick blue lines represent distances that decrease on transition from RP_o to $RP_{itc,\leq 7}$. Red and pink arrows show the proposed positions at which scrunched template (–9 to –10) and nontemplate strand DNA (–5 to −6), respectively, emerge from RNAP.

DNA downstream of the transcription bubble and DNA upstream of the transcription bubble (Figure 5.8B). This validated the scrunching mechanism and supported a model in which the stress accumulated from the DNA unwinding and compaction drives the breaking of bonds between RNAP and DNA into promoter escape and elongation. Conformational heterogeneity of RNAP during elongation (Coban et al. 2006) and the scrunching and rotation of T7 RNAP during initiation has also been observed with smFRET (Sharma et al. 2008).

Reverse Transcriptase

Reverse transcriptase (RT), in particular the human immunodeficiency virus type 1 (HIV-1) RT, is a nucleic acid polymerase that uses RNA as a template to synthesize double-stranded DNA for integration into the host genome. RT is a heterodimer consisting of a p51 and p66 subunit, with the latter containing RNA- and DNA-dependent DNA polymerase activities and an RNase H domain. Crystallographic studies of the RT–nucleic acid complexes showed a single primer/template binding mode, but recent ensemble kinetic studies suggested a heterogeneous mixture of several binding modes. To determine the existence of several species and obtain functional and structural information of each, the group of Claus Seidel developed a multiparameter fluorescence detection technique (MFD) that measures FRET based on the reduction of donor lifetimes (Rothwell et al. 2003). The group investigated donor-labeled RT at the p66 domain and acceptor-labeled primer to reveal three distinct states of RT–nucleic acid complexes. One of the states, not seen before by X-ray crystallography, did not incorporate nucleotides and was structurally different from the other two. The other two states were similar to the crystal structure but differed only by 5Å in the nucleic acid position from each other. The two states undergo conformational changes upon binding a nucleotide, allowing the researchers to assign the structures as productive stages in DNA polymerization.

The mechanism of RT that coordinates the DNA polymerase and RNase H activities is unknown. Although RNase H cleavage analysis has shown different interaction modes with substrates, crystal structures have only shown one enzyme-binding orientation. To address this question and study RT binding conformations, an smFRET assay with msALEX was used (Abbondanzieri et al. 2008). Surface-immobilised DNA or RNA substrate was acceptor-labeled at the 5′ or 3′ end and immersed in a solution containing free RT molecules, donor-labeled either at the RNase H domain or at the fingers subdomain (at the DNA polymerase domain) on the p66 subunit. The distinct FRET efficiencies showed that RT binds DNA and RNA primers in opposite directions, hence defining its enzymatic activity. For the binding of RT to duplexes containing polypurine RNA primers for plus-strand DNA synthesis, RT was found to dynamically switch between binding orientations. Cognate nucleotides and nonnucleoside RT inhibitors, a class of anti-HIV drugs, were found to have opposing effects on switching rates. The dynamic conformational changes of RT binding to substrates allowed the enzyme to explore multiple binding orientations, thereby regulating enzymatic kinetics and replication efficacy. This provided new insights into the dynamic structure and function of RT that had not been shown before with ensemble methods or crystal structures and could help in the development of pharmacological agents.

5.7.3. Applications to Molecular Motors

Molecular motors are proteins that use the energy from nucleotide hydrolysis to transport cargoes in cells and exert mechanical tension. Motor proteins have been extensively studied by single-molecule methods (especially optical traps and single-molecule fluorescence)

that revealed the stepping characteristics of each motor, the energy requirements, the determinants, and conformational changes occurring during each stepping cycle. Single-molecule FRET has also contributed to the improved understanding of two of these motor proteins: linear motor kinesin and rotary machine F_0F_1-ATP synthase. In both cases, FRET pairs were incorporated in the molecular motor, either in different subunits or within a single subunit, and conformational changes were observed during the motor's function.

ATP Synthase

F_0F_1-ATP synthase is a multi-subunit molecular machine that uses the proton-motive force (a transmembrane gradient of proton concentration) to generate ATP from adenosine diphosphate (ADP) and phosphate in the membranes of mitochondria, chloroplasts, and bacteria. The machine can also work in reverse, consuming ATP molecules to translocate protons across membranes, and it consists of two coupled components: a membrane-bound hydrophobic part (F_0) responsible for proton translocation and a hydrophilic part (F_1) responsible for ATP hydrolysis and synthesis (Figure 5.9A). A salient feature of the motor is the presence of a multisubunit rotor (subunit composition of $\gamma\epsilon c_{10\text{-}12}$) that rotates 120° per step (with one step corresponding to the synthesis or consumption of a single ATP) with regards to a stator assembly (subunit composition of $\alpha_3\beta_3\delta ab_2$) and the rest of the protein, with each β subunit containing an ATP-binding site. Extensive biochemical, structural, and single-molecule studies have made this rotary motor one of the best-studied proteins in bioscience (Abrahams et al. 1994; Boyer et al. 2000) and inspired some of the most elegant and ground-breaking single-molecule experiments (Noji et al. 1997). Despite this impressive body of work, a full understanding of the mechanism of coupling of the two main functions of the protein is first beginning to emerge.

The first single-molecule FRET on the motor (Borsch et al. 2002) examined the intact F_0F_1 proteins reconstituted in liposomes, a system more complex that the F_1 part used for initial single-molecule fluorescence studies (Noji et al. 1997). Single-molecule FRET is well suited for studying the rotary motion because it is stochastic and cyclic and, as such, it cannot be synchronised and dissected at the ensemble level. The rotation of the γ subunit versus the rest of the protein was observed by placing a donor on the b subunit of F_0F_1 and an acceptor on the γ subunit (Figure 5.9A). Upon rotation of the γ subunit, a stepwise change in FRET states was expected due to the stepwise changes among the three different donor–acceptor distances corresponding to the three main conformations of the γ subunit against the rest of the protein (similar to the conformations in Figure 5.9B). The first experiments used a confocal setup with ∼1-ms temporal resolution to observe diffusing liposomes with single F_0F_1 molecules. Analysis of fluorescence intensity and FRET time traces led to observation of the predicted switching among three FRET states during the diffusion of one membrane-bound F_0F_1 molecule through an expanded confocal volume (Figure 5.9C). Of importance, the rate of FRET transitions due to presumed rotation matched the rate of ∼30 sec^{-1} obtained from enzymatic assays performed for this protein preparation. In further studies of F_0F_1 using a revised labeling scheme (Diez et al. 2004), an improved temporal resolution, and additional FRET information (arising from donor-lifetime analysis), it was shown that the γ subunit rotates in a stepwise fashion (as opposed to continuously) during ATP synthesis and in a direction opposite to the rotation observed during ATP hydrolysis (Figure 5.9D, E). Similar results were obtained when a fluorophore was introduced in the ϵ subunit, which also forms part of the rotor assembly (Zimmermann et al. 2006).

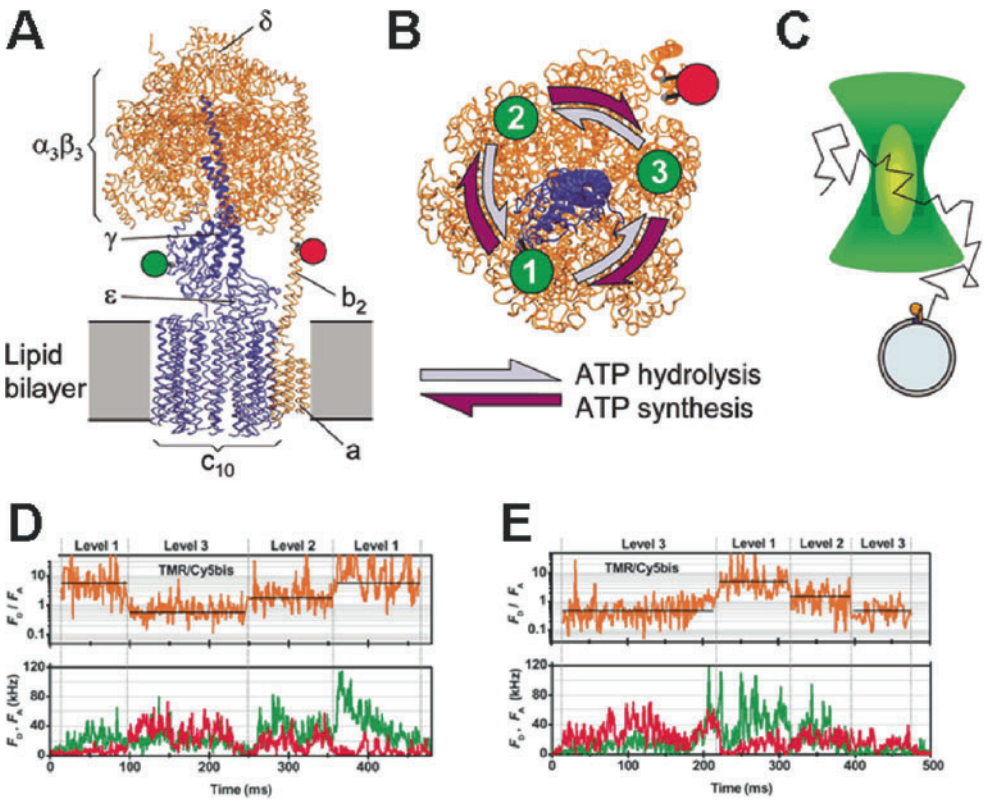


Figure 5.9. Single-molecule fluorescence resonance energy transfer (FRET) studies of adenosine triphosphate (ATP) hydrolysis and synthesis by ATP synthase. **A.** Side view of a model of bacterial F_1F_0-ATP synthase integrated in a lipid bilayer along with FRET probes. The rotor assembly is in blue, and the stator assembly is in orange. The donor (*green circle*) is attached to the γ subunit, and the acceptor (*red circle*) is attached to the b_2 subunit. **B.** Top view of a cross section of the model shown in panel A. Rotation of the stator generates three distinct FRET states in a cyclic fashion. ATP hydrolysis causes counterclockwise rotation, whereas ATP synthesis leads to clockwise rotation. **C.** Experimental format. Diffusing labeled ATP-synthase molecules in liposomes cross a confocal volume and generate long fluorescence bursts (see panels D and E). **D.** Photon bursts from ATP-synthase molecules during ATP hydrolysis show cycling among three FRET states (represented by three levels of the ratio F_D/F_A, which depends on FRET efficiency). **E.** Photon bursts from ATP-synthase molecules during ATP synthesis show cycling along three FRET states in a direction opposite to the one for ATP hydrolysis. [Adapted from Diez et al. (2004). Copyright 2004 Nature Publishing Group]

Kinesin

Kinesin, one of the smallest molecular motors in cells, is a dimeric protein that uses the energy of ATP to move cargoes towards the plus end of microtubules (Vale 2003). Kinesin has been the focus of extensive biochemical, structural, and biophysical studies, which have showed that kinesin moves in 8-nm steps by a hand-over-hand processive motion (Yildiz et al. 2004). The group of Ron Vale was the first to use FRET methods to look at conformational changes while kinesin reconfigures itself during its mechanochemical cycle, either using static measurements (e.g., Rice et al. 1999) or by probing kinesins in action on microtubules.

The first application of smFRET to the study of kinesin (Tomishige et al. 2006) examined conformations of the neck-linker region of kinesin, a 12–amino acid region crucial for the coordination of biochemical activities and conformational changes within each kinesin step (Figure 5.10A). From prior work, it was known that the neck region adopts a "docked" (forward-extending) state in the presence of ATP and an undocked state (backward-extending) in the presence of ADP or in the nucleotide-free state (Figure 5.10B). These observations led to a model in which ATP conversion to ADP in the leading head of the kinesin leads to docking of the neck-linker region and to a ∼16-nm movement of the trailing head towards the direction of motion. Previous attempts to test this model were complicated by ensemble averaging of conformations (e.g., both linkers were labeled in a kinesin homodimer) and by operating at high ATP concentrations, where transitions between linker conformers are too fast to be resolved directly. To identify the different conformations of the neck linker and to see dynamic changes in neck-linker conformations as kinesins move on axonemes, the group prepared kinesins in which only a single neck linker was labeled with a FRET pair. To achieve this challenging task, the group used a tour-de-force of site-directed mutagenesis, labeling, and functional assays to identify six labeling positions with kinesin that did not affect function.

Using a TIRF microscope and alternating-laser excitation, it was possible to identify kinesins with a donor and an acceptor and generate FRET histograms for several interprobe positions and nucleotide states. This experiment was a static one, providing equilibrium distributions of FRET efficiencies; the FRET analysis confirmed the expected docked conformation for the ATP-bound state (using a nonhydrolyzable analog of ATP) and the undocked conformation for the nucleotide-free state. The group went on to study transitions between conformations in real time by tracking doubly labeled kinesins as they moved on microtubules (Figure 5.10C); this impressive experiment combined tracking of a large-scale, "global" probing of the unidirectional movement of kinesin along with "local" probing of the intrasubunit conformational changes occurring during each step. Indeed, at low ATP concentrations (∼1 μM), the group succeeded in observing transitions between dwells in high- and low-FRET states that represent docked and undocked states, respectively, for the FRET pair used (Figure 5.10D).

This first application of smFRET to kinesin dynamics was followed by a study that addressed the question of the prevailing conformation of kinesin while it waits for the second ATP molecule to bind between steps (Mori et al. 2007). The group this time used two FRET pairs that examined intersubunit distances within kinesin heterodimers. An initial calibration using static distributions of kinesins showed that in the presence of an ATP analog, a two-head-bound state predominates, whereas in the presence of low ADP, states in which the heads come closer to each other (before dissociation) are populated; these states were interpreted as one-head-bound states, which was confirmed using a heterodimeric kinesin able to bind only through a single head.

The most interesting conclusions of this work came from studies on FRET-labeled kinesin molecules moving on microtubules. At saturating ATP conditions, kinesins occupy FRET states consistent with a two-head-bound state, pointing to the fact that the time spent in the one-head-bound state is very small. In contrast, at limiting ATP concentrations, the predominant FRET state is consistent with an one-head-bound state interdispersed with transitions to lower- and higher-FRET states, which are interpreted as transient two-head-bound states (mostly towards a forward direction). This detailed experimental treatment of such a complex system allowed formulation of an updated model for kinesin translocation at different levels of ATP and explained some inconsistencies existing in the field.

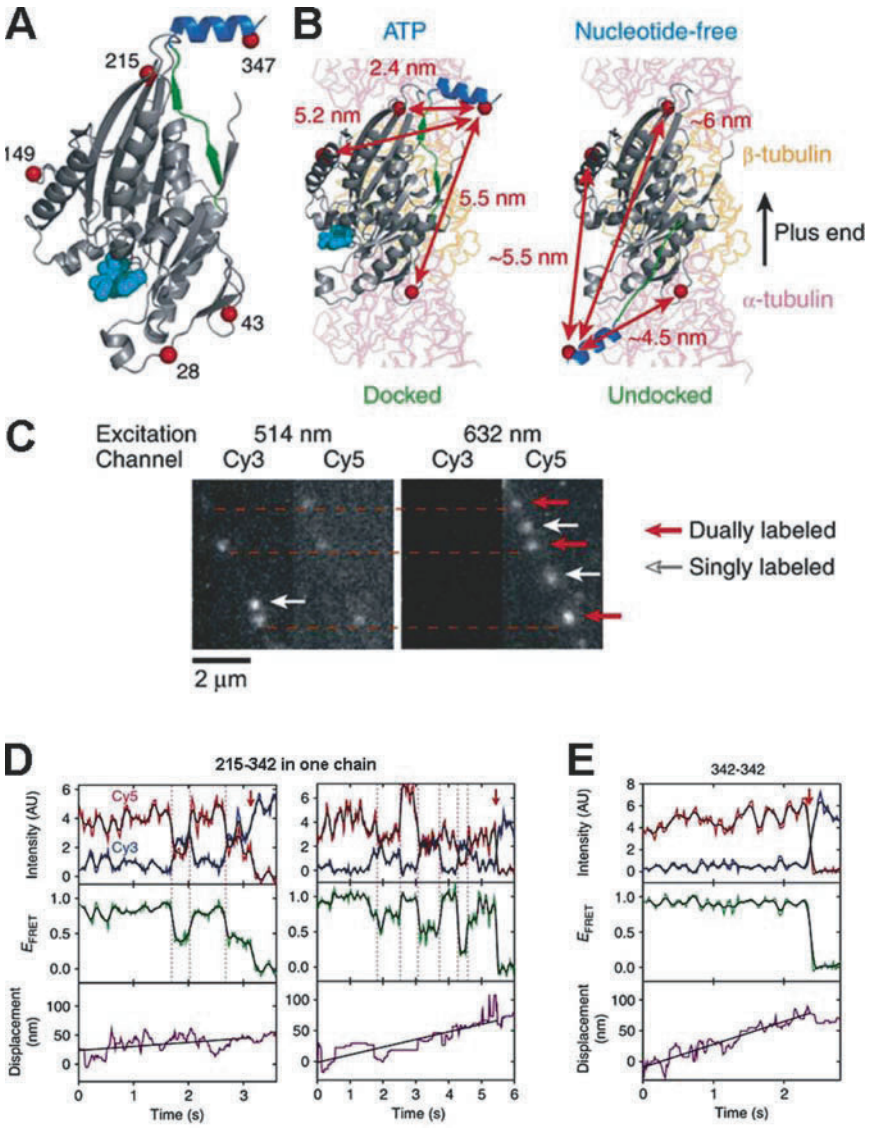


Figure 5.10. Single-molecule fluorescence resonance energy transfer (smFRET) measurements of kinesin dynamics. **A.** Model of kinesin showing labeling sites (*red spheres*) used in Tomishige et al. (2006). The neck linker (*green*) is in the docked conformation; adenosine diphosphate (ADP) is shown as a cyan space-fill. **B.** Intraprotein distances between labeling sites for the adenosine triphosphate (ATP)–bound, docked state of kinesin and for the nucleotide-free, undocked state of kinesin. **C.** Experimental format. Sequential excitation by 514- and 632-nm light and wide-field imaging identifies singly- and doubly- labeled kinesins on surface-immobilised microtubules. **D.** FRET fluctuations due to conformational changes in the neck linker during the motion of kinesins along microtubules at low ATP concentration. The donor and intensity time traces **(top)** are used to calculate FRET efficiencies **(middle)** during the motion of doubly labeled kinesins; kinesin motion is tracked by localising the fluorophores **(bottom)**. Abrupt and large changes in FRET are due to conformational changes in the neck linker. **E.** Negative control (heterodimeric kinesin 342-342) shows no significant conformational changes under conditions identical to ones used in panel D. [Adapted from Tomishige et al. 2006. Copyright 2006 Nature Publishing Group.]

5.7.4. Applications to Protein Folding and Dynamics

The multistep nature of protein folding and dynamics means that bulk measurements can only provide ensemble-averaged information on the process, whereas theory and simulations consider folding on the single-molecule or subpopulation level. By contrast, single-molecule methods allow direct observation of individual proteins under various experimental conditions for extended periods. As a result, smFRET has found numerous applications to the study of protein folding and dynamics smFRET approaches have been discussed in recent excellent reviews (Michalet et al. 2006; Schuler and Eaton 2008; Zhuang and Rief 2003). The general approach used for investigations of folding with smFRET has been to study labeled proteins in varying concentrations of denaturant, using the ability of smFRET to monitor individual molecular subpopulations to resolve the folded and unfolded states and thus obtain novel information on folding mechanisms and pathways.

The first smFRET study of protein folding by Deniz et al. (2000), used diffusion methods to directly confirm a two-state model of folding for the chymotrypsin inhibitor and extract free energy landscapes for the protein under varying concentrations of denaturant. This work was extended by a similar study (Schuler et al. 2002) on a doubly-labeled cold-shock protein (Csp). After careful comparison with a series of rigid standards, Schuler et al. recovered a limit in the reconfiguration time for the polypeptide chain in the unfolded state, an important parameter not available from ensemble measurements. The studies on Csp were further extended by measurements of the kinetics of folding in a microfabricated laminar-flow mixing device (Lipman et al. 2003).

Parallel studies have applied surface immobilisation approaches with great success, observing folding kinetics for extended periods of time. After early observations of protein folding on silanized surfaces that may have influenced the folding landscape to a certain degree (Talaga et al. 2000), Rhoades et al. (2003) used the interesting method of encapsulation of labeled proteins in surface-tethered 100-nm lipid vesicles to avoid any surface interactions and found evidence for multiple folding pathways for the folding of adenylate kinase. Kuzmenkina et al. (2005) used surface immobilisation to investigate the dynamics of ribonuclease HI under varying concentrations of denaturant, successfully obtaining a partial free energy landscape for the protein and providing strong evidence that the protein retains significant residual nonrandom structure even in the unfolded state. Recent work has also studied the mechanisms of chaperonin-mediated protein folding (Sharma et al. 2008) and provided strong evidence for the proposal that protein confinement by GroEL/GroES stretches apart strongly hydrophobic regions of a polypeptide and thus opens up folding pathways inaccessible to spontaneous hydrophobic collapse.

smFRET has also been used to study the fundamental mechanisms of enzymatic catalysis. Henzler-Wildman et al. (et al. 2007) studied the enzymatic reaction trajectory of adenylate kinase (Adk) using a number of methods, including smFRET both in solution and on the surface. Adk was fluorescently labeled and monitored in the ligand-bound and ligand-free states (Figure 5.11A). Solution smFRET measurements showed that a state similar to the closed state was significantly populated even in the absence of ligand (Figure 5.11A). Surface-immobilised measurements (Figure 5.11B–D) supported this finding and, by analysing the intensity correlation function (Figure 5.11G), provided estimated interconversion rates between the ligand free open and closed states that agreed well with nuclear magnetic resonance (NMR) data, which had indicated the presence of interconverting states but was unable to determine the amplitude of any dynamics. By combining NMR, molecular dynamics, and smFRET measurements, this research provided strong evidence that

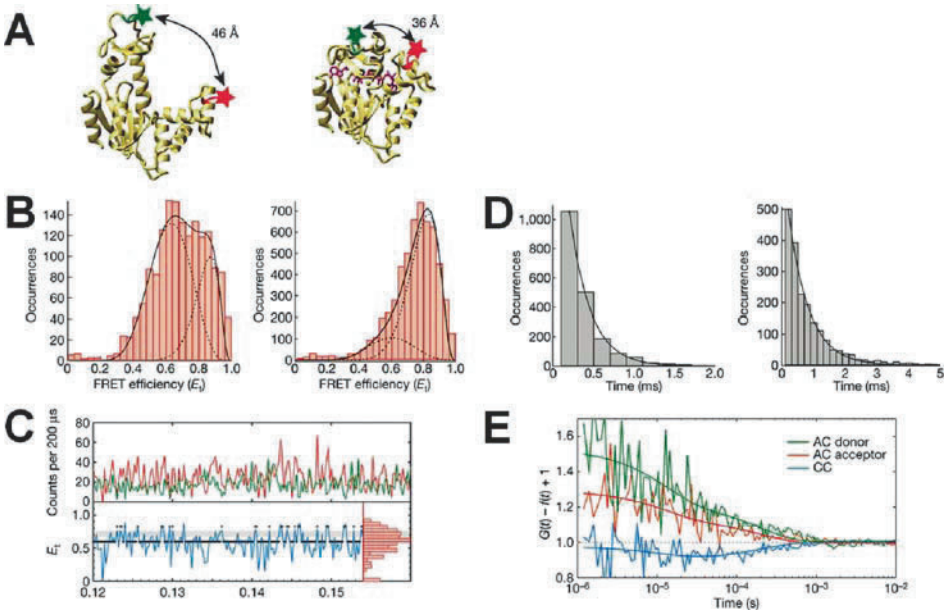


Figure 5.11. Using single-molecule fluorescence resonance energy transfer (smFRET) to study the enzymatic reaction trajectory of adenylate kinase (Adk). **A.** Adk was fluorescently labeled and monitored in the ligand-free and ligand-bound state. **B.** Histograms of FRET efficiencies of ligand-free **(left)** and ligand-bound Adk **(right)** measured from single-molecule diffusion experiments. The dashed lines represent Gaussian fits of the corresponding distance histograms back-transformed into FRET efficiency E_t, and the solid line is the sum of those distributions. The ligand-free state shows a state similar to the closed state, significantly populated even in the absence of ligand. **C.** Section of a single-molecule fluorescence time trace of ligand-free Adk tethered on a glass surface. Donor (*green*) and acceptor (*red*) intensities are shown together with the corresponding E_t (*blue*), including the E_t histogram over 25 time traces (*red, right*). Lifetimes of the open and closed state (*black lines*) were determined using a transition zone for E_t values where the Gaussian distributions of the open and closed states overlap (*grey*). **D.** Resulting lifetimes of 25 time traces were averaged, fitted exponentially, and corrected for missed events and triplet-state dynamics to yield $k_{\text{open}} = 6{,}500 \pm 500\ \text{sec}^{-1}$ **(left)** and $k_{\text{closed}} = 2{,}000 \pm 200\ \text{sec}^{-1}$ **(right)**, respectively. **E.** The intensity correlation function $G(t)$ for autocorrelation (AC) and cross-correlation (CC) analysis for six immobilized ligand-free Adk molecules was calculated to obtain the overall opening/closing rate. Anticorrelation of the CC function is observed in the $10^{-4}-10^{-3}$ sec timescale with a fitted overall opening/closing rate of 7,000 ± 2,000 sec^{-1}. [Adapted from Henzler-Wildman (2007). Copyright 2007 Nature Publishing Group.]

adenylate kinase samples states resembling the catalytically active closed conformation even in the absence of ligand binding, with implications for the theory of dynamics of enzymatic catalysis.

5.8. Conclusion and Future Prospects

The applications discussed show clearly that single-molecule FRET has affected several fields in basic and applied biosciences. However, there is also room for improvement that will help the implementation and interpretation of smFRET measurements. Moreover, as the research community becomes more familiar with the method, more challenging applications are being considered, raising the bar for single-molecule FRET. Some of the active

areas under development were outlined in Section 5.6. Moreover, advances in site-specific labeling technologies (Deniz et al. 2008; Kapanidis et al. 2001) will expand the number of FRET pairs and overcome distance constraints that can be incorporated into proteins to study protein structure, protein dynamics, and conformational changes. New buffer components that delay photobleaching (e.g., Vogelsang et al. 2008) and control fluorophore photophysics (e.g., Rasnik et al. 2006), as well as the development of reliable quantum-dot biolabels (Bruchez et al. 1998; Michalet et al. 2005) will lead to extended and uninterrupted recordings in a variety of *in vitro* experiments. Ways to achieve this in living cells will enable exciting smFRET experiments in a natural context. Finally, the appeal of the method for observing molecular interactions will increase if the working concentration can be extended to study fluorescent species in concentrations greater than nanomolar range. This may be achieved by combining smFRET with methods that confine the laser excitation volume (Foquet et al. 2004; Levene et al. 2003; Vogelsang et al. 2007) or by using nonfluorescent analytes. Advances in all these aspects, along with improvements in detector technologies and FRET data-analysis algorithms (McKinney et al. 2006), will keep single-molecule FRET in the limelight for years to come.

References

Abbondanzieri EA, Bokinsky G, Rausch JW, Zhang JX, Le Grice SFJ, Zhuang XW. 2008. Dynamic binding orientations direct activity of HIV reverse transcriptase. Nature 453(7192):184–9.

Abrahams JP, Leslie AG, Lutter R, Walker JE. 1994. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature 370(6491):621–8.

Ambrose WP, Goodwin PM, Martin JC, Keller RA. 1994. Alterations of single-molecule fluorescence lifetimes in near-field optical microscopy. Science 265(5170):364–7.

Andrecka J, Lewis R, Bruckner F, Lehmann E, Cramer P, Michaelis J. 2008. Single-molecule tracking of mRNA exiting from RNA polymerase II. Proc Natl Acad Sci USA 105(1):135–40.

Axelrod D. 2001. Total internal reflection fluorescence microscopy in cell biology. Traffic 2(11):764–74.

Bai L, Santangelo TJ, Wang MD. 2006. Single-molecule analysis of RNA polymerase transcription. Annu Rev Biophys Biomol Struct 35:343–60.

Bath J, Turberfield AJ. 2007. DNA nanomachines. Nat Nanotechnol 2(5):275–84.

Betzig E, Chichester RJ. 1993. Single molecules observed by near-field scanning optical microscopy. Science 262(5138):1422–5.

Betzig E, Trautman JK. 1992. Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit. Science 257(5067):189–95.

Blanchard SC, Gonzalez RL, Kim HD, Chu S, Puglisi JD. 2004a. tRNA selection and kinetic proofreading in translation. Nat Struct Mol Biol 11(10):1008–14.

Blanchard SC, Kim HD, Gonzalez RL Jr, Puglisi JD, Chu S. 2004b. tRNA dynamics on the ribosome during translation. Proc Natl Acad Sci USA 101(35):12893–8.

Borisenko V, Lougheed T, Hesse J, Fureder-Kitzmuller E, Fertig N, Behrends JC, Woolley GA, Schutz GJ. 2003. Simultaneous optical and electrical recording of single gramicidin channels. Biophys J 84(1):612–22.

Borsch M, Diez M, Zimmermann B, Reuter R, Graber P. 2002. Stepwise rotation of the gamma-subunit of EF(0)F(1)-ATP synthase observed by intramolecular single-molecule fluorescence resonance energy transfer. FEBS Lett 527(1–3):147–52.

Boyer LA, Shao X, Ebright RH, Peterson CL. 2000. Roles of the histone H2A-H2B dimers and the (H3-H4)(2) tetramer in nucleosome remodeling by the SWI-SNF complex. J Biol Chem 275(16):11545–52.

Braslavsky I, Hebert B, Kartalov E, Quake SR. 2003. Sequence information can be obtained from single DNA molecules. Proc Natl Acad Sci USA 100(7):3960–4.

Bruchez M, Jr., Moronne M, Gin P, Weiss S, Alivisatos AP. 1998. Semiconductor nanocrystals as fluorescent biological labels. Science 281(5385):2013–6.

Clamme JP, Deniz AA. 2005. Three-color single-molecule fluorescence resonance energy transfer. Chemphyschem 6(1):74–7.

Clegg RM. 1992. Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol 211:353–88.
Clegg RM. 1995. Fluorescence resonance energy transfer. Curr Opin Biotechnol 6(1):103–10.
Clegg RM, Murchie AI, Zechel A, Lilley DM. 1993. Observing the helical geometry of double-stranded DNA in solution by fluorescence resonance energy transfer. Proc Natl Acad Sci USA 90(7):2994–8.
Coban O, Lamb DC, Zaychikov E, Heumann H, Nienhaus GU. 2006. Conformational heterogeneity in RNA polymerase observed by single-pair FRET microscopy. Biophys J 90(12):4605–17.
Dale RE, Eisinger J, Blumberg WE. 1979. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. Biophys J 26(2):161–93.
Deniz AA, Dahan M, Grunwell JR, Ha T, Faulhaber AE, Chemla DS, Weiss S, Schultz PG. 1999. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. Proc Natl Acad Sci USA 96(7):3670–5.
Deniz AA, Laurence TA, Beligere GS, Dahan M, Martin AB, Chemla DS, Dawson PE, Schultz PG, Weiss S. 2000. Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. Proc Natl Acad Sci USA 97(10):5179–84.
Deniz AA, Mukhopadhyay S, Lemke EA. 2008. Single-molecule biophysics: at the interface of biology, physics and chemistry. J R Soc Interface 5(18):15–45.
Diez M, Zimmermann B, Borsch M, Konig M, Schweinberger E, Steigmiller S, Reuter R, Felekyan S, Kudryavtsev V, Seidel CA, et al. 2004. Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. Nat Struct Mol Biol 11(2):135–41.
Edel JB, Eid JS, Meller A. 2007. Accurate single molecule FRET efficiency determination for surface immobilized DNA using maximum likelihood calculated lifetimes. J Phys Chem B 111(11):2986–90.
Foquet M, Korlach J, Zipfel WR, Webb WW, Craighead HG. 2004. Focal volume confinement by submicrometer-sized fluidic channels. Anal Chem 76(6):1618–26.
Förster T. 1948. Intramolecular energy migration and fluorescence. Ann Physik 2:55–75.
Funatsu T, Harada Y, Tokunaga M, Saito K, Yanagida T. 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. Nature 374(6522):555–9.
Goodman RP, Berry RM, Turberfield AJ. 2004. The single-step synthesis of a DNA tetrahedron. Chem Commun (Cambridge) 12:1372–3.
Goodman RP, Heilemannt M, Dooset S, Erben CM, Kapanidis AN, Turberfield AJ. 2008. Reconfigurable, braced, three-dimensional DNA nanostructures. Nat Nanotechnol 3(2):93–96.
Goodman RP, Schaap IA, Tardin CF, Erben CM, Berry RM, Schmidt CF, Turberfield AJ. 2005. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. Science 310(5754):1661–5.
Greenleaf WJ, Woodside MT, Block SM. 2007. High-resolution, single-molecule measurements of biomolecular motion. Annu Rev Biophys Biomol Struct 36:171–90.
Guo J, Xu N, Li Z, Zhang S, Wu J, Kim DH, Sano Marma M, Meng Q, Cao H, Li X, et al. 2008. Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. Proc Nat Acad Sci USA 105(27):9145–50.
Ha T. 2001. Single-molecule fluorescence resonance energy transfer. Methods 25:78–86.
Ha T. 2004. Structural dynamics and processing of nucleic acids revealed by single-molecule spectroscopy. Biochemistry 43(14):4055–63.
Ha T, Enderle T, Ogletree DF, Chemla DS, Selvin PR, Weiss S. 1996. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proc Natl Acad Sci USA 93(13):6264–8.
Ha T, Rasnik I, Cheng W, Babcock HP, Gauss GH, Lohman TM, Chu S. 2002. Initiation and re-initiation of DNA unwinding by the *Escherichia coli* Rep helicase. Nature 419(6907):638–41.
Ha T, Ting AY, Liang J, Caldwell WB, Deniz AA, Chemla DS, Schultz PG, Weiss S. 1999a. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. Proc Natl Acad Sci USA 96(3):893–8.
Ha T, Ting AY, Liang J, Deniz AA, Chemla DS, Schultz PG, Weiss S. 1999b. Temporal fluctuations of fluorescence resonance energy transfer between two dyes conjugated to a single protein. Chem Phys 247:107–118.
Ha T, Zhuang X, Kim HD, Orr JW, Williamson JR, Chu S. 1999c. Ligand-induced conformational changes observed in single RNA molecules. Proc Natl Acad Sci USA 96(16):9077–82.
Haas E, Katchalski-Katzir E, Steinberg IZ. 1978. Effect of the orientation of donor and acceptor on the probability of energy transfer involving electronic transitions of mixed polarization. Biochemistry 17(23): 5064–70.
Haas E, Wilchek M, Katchalski-Katzir E, Steinberg IZ. 1975. Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer. Proc Natl Acad Sci USA 72(5):1807–11.

Harms GS, Orr G, Montal M, Thrall BD, Colson SD, Lu HP. 2003. Probing conformational changes of gramicidin ion channels by single-molecule patch-clamp fluorescence microscopy. Biophys J 85(3):1826–38.

Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcavitch JW, et al. 2008. Single-molecule DNA sequencing of a viral genome. Science 320(5872):106–9.

Heilemann M, Tinnefeld P, Sanchez Mosteiro G, Garcia Parajo M, Van Hulst NF, Sauer M. 2004. Multistep energy transfer in single molecular photonic wires. J Am Chem Soc 126(21):6514–5.

Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, et al. 2007. Intrinsic motions along an enzymatic reaction trajectory. Nature 450(7171):838–44.

Herbert KM, Greenleaf WJ, Block SM. 2008. Single-molecule studies of RNA polymerase: motoring along. Annu Rev Biochem 77:149–76.

Hohng S, Joo C, Ha T. 2004. Single-molecule three-color FRET. Biophys J 87(2):1328–37.

Hohng S, Zhou RB, Nahas MK, Yu J, Schulten K, Lilley DMJ, Ha TJ. 2007. Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the Holliday junction. Science 318(5848):279–83.

Jares-Erijman EA, Jovin TM. 2003. FRET imaging. Nat Biotechnol 21(11):1387–95.

Joo C, McKinney SA, Nakamura M, Rasnik I, Myong S, Ha T. 2006. Real-time observation of RecA filament dynamics with single monomer resolution. Cell 126(3):515–27.

Kapanidis AN, Ebright YW, Ebright RH. 2001. Site-specific incorporation of fluorescent probes into protein: hexahistidine-tag-mediated fluorescent labeling with (Ni(2+):nitrilotriacetic acid (n)-fluorochrome conjugates. J Am Chem Soc 123(48):12123–5.

Kapanidis AN, Lee NK, Laurence TA, Doose S, Margeat E, Weiss S. 2004. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. Proc Natl Acad Sci USA 101(24):8936–41.

Kapanidis AN, Margeat E, Ho SO, Kortkhonjia E, Weiss S, Ebright RH. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science 314(5802):1144–7.

Kapanidis AN, Margeat E, Laurence TA, Doose S, Ho SO, Mukhopadhyay J, Kortkhonjia E, Mekler V, Ebright RH, Weiss S. 2005. Retention of transcription initiation factor sigma(70) in transcription elongation: single-molecule analysis. Mol Cell 20(3):347–56.

Kapanidis AN, Weiss S. 2002. Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules. J Chem Phys 117(24):10953–64.

Karymov M, Chinnaraj M, Bogdanov A, Srinivasan AR, Zheng G, Olson WK, Lyubchenko Y. 2008a. Structure, dynamics, and branch migration of a DNA Holliday junction: a single-molecule fluorescence and modeling study. Biophys J. 95(9):4372–83

Karymov M, Daniel D, Sankey OF, Lyubchenko YL. 2005. Holliday junction dynamics and branch migration: single-molecule analysis. Proc Natl Acad Sci USA 102(23):8186–91.

Karymov MA, Bogdanov A, Lyubchenko YL. 2008b. Single molecule fluorescence analysis of branch migration of Holliday junctions: effect of DNA sequence. Biophys J 95(3):1239–47.

Koopmans WJ, Brehm A, Logie C, Schmidt T, van Noort J. 2007. Single-pair FRET microscopy reveals mononucleosome dynamics. J Fluoresc 17(6):785–95.

Kuzmenkina EV, Heyes CD, Nienhaus GU. 2005. Single-molecule Forster resonance energy transfer study of protein dynamics under denaturing conditions. Proc Natl Acad Sci USA 102(43):15471–6.

Lakowicz JR. 2006. *Principles of Fluorescence Spectroscopy*, 3rd ed. Springer, New York.

Lang MJ, Fordyce PM, Block SM. 2003. Combined optical trapping and single-molecule fluorescence. J Biol 2(1):6.

Laurence TA, Kong XX, Jager M, Weiss S. 2005. Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. Proc Natl Acad Sci USA 102(48):17348–53.

Lee NK, Kapanidis AN, Koh HR, Korlann Y, Ho SO, Kim Y, Gassman N, Kim SK, Weiss S. 2007. Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances. Biophys J 92(1):303–12.

Lee NK, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebright RH, Weiss S. 2005. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys J 88(4):2939–53.

Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. Science 299(5607):682–6.

Lipman EA, Schuler B, Bakajin O, Eaton WA. 2003. Single-molecule measurement of protein folding kinetics. Science 301(5637):1233–5.

Lippincott-Schwartz J, Snapp E, Kenworthy A. 2001. Studying protein dynamics in living cells. Nat Rev Mol Cell Biol 2(6):444–56.

Margeat E, Kapanidis AN, Tinnefeld P, Wang Y, Mukhopadhyay J, Ebright RH, Weiss S. 2006. Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. Biophys J 90(4):1419–31.

Margittai M, Widengren J, Schweinberger E, Schroder GF, Felekyan S, Haustein E, Konig M, Fasshauer D, Grubmuller H, Jahn R, et al. 2003. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. Proc Natl Acad Sci USA 100(26): 15516–21.

McKinney SA, Declais AC, Lilley DM, Ha T. 2003. Structural dynamics of individual Holliday junctions. Nat Struct Biol 10(2):93–7.

McKinney SA, Freeman AD, Lilley DM, Ha T. 2005. Observing spontaneous branch migration of Holiday junctions one step at a time. Proc Natl Acad Sci USA. 102(16):5715–5720.

McKinney SA, Joo C, Ha T. 2006. Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys J 91(5):1941–51.

Michalet X, Pinaud FF, Bentolila LA, Tsay JM, Doose S, Li JJ, Sundaresan G, Wu AM, Gambhir SS, Weiss S. 2005. Quantum dots for live cells, in vivo imaging, and diagnostics. Science 307(5709):538–44.

Michalet X, Weiss S, Jager M. 2006. Single-molecule fluorescence studies of protein folding and conformational dynamics. Chem Rev 106(5):1785–813.

Miyawaki A, Tsien RY. 2000. Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. Appl Chimeric Genes Hybrid Proteins B 327:472–500.

Moerner WE, Fromm DP. 2003. Methods of single-molecule fluorescence spectroscopy and microscopy. Rev Sci Instrum 74(8):3597–619.

Morgan MA, Okamoto K, Kahn JD, English DS. 2005. Single-molecule spectroscopic determination of lac repressor-DNA loop conformation. Biophys J 89(4):2588–96.

Mori T, Vale RD, Tomishige M. 2007. How kinesin waits between steps. Nature 450(7170):750–4.

Muller BK, Reuter A, Simmel FC, Lamb DC. 2006. Single-pair FRET characterization of DNA tweezers. Nano Lett 6(12):2814–20.

Muller BK, Zaychikov E, Brauchle C, Lamb DC. 2005. Pulsed interleaved excitation. Biophys J 89(5):3508–522.

Murakoshi H, Iino R, Kobayashi T, Fujiwara T, Ohshima C, Yoshimura A, Kusumi A. 2004. Single-molecule imaging analysis of Ras activation in living cells. Proc Natl Acad Sci USA 101(19):7317–22.

Myong S, Bruno MM, Pyle AM, Ha T. 2007. Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. Science 317(5837):513–6.

Myong S, Rasnik I, Joo C, Lohman TM, Ha T. 2005. Repetitive shuttling of a motor protein on DNA. Nature 437(7063):1321–5.

Nie S, Chiu DT, Zare RN. 1994. Probing individual molecules with confocal fluorescence microscopy. Science 266(5187):1018–21.

Noji H, Yasuda R, Yoshida M, Kinosita K, Jr. 1997. Direct observation of the rotation of F1-ATPase. Nature 386(6622):299–302.

Rasnik I, McKinney SA, Ha T. 2006. Nonblinking and longlasting single-molecule fluorescence imaging. Nat Methods 3(11):891–3.

Rasnik I, Myong S, Cheng W, Lohman TM, Ha T. 2004. DNA-binding orientation and domain conformation of the E. coli rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes. J Mol Biol 336(2):395–408.

Rhoades E, Gussakovsky E, Haran G. 2003. Watching proteins fold one molecule at a time. Proc Nat Acad Sci USA 100(6):3197–202.

Rice S, Lin AW, Safer D, Hart CL, Naber N, Carragher BO, Cain SM, Pechatnikova E, Wilson-Kubalek EM, Whittaker M, et al. 1999. A structural change in the kinesin motor protein that drives motility. Nature 402(6763): 778–84.

Ross J, Buschkamp P, Fetting D, Donnermeyer A, Roth CM, Tinnefeld P. 2007. Multicolor single-molecule spectroscopy with alternating laser excitation for the investigation of interactions and dynamics. J Phys Chem B 111(2):321–6.

Rothemund PW. 2006. Folding DNA to create nanoscale shapes and patterns. Nature 440(7082):297–302.

Rothwell PJ, Berger S, Kensch O, Felekyan S, Antonik M, Wohrl BM, Restle T, Goody RS, Seidel CA. 2003. Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. Proc Natl Acad Sci USA 100(4):1655–60.

Sabanayagam CR, Eid JS, Meller A. 2004. High-throughput scanning confocal microscopy for single molecule analysis. Appl Phys Lett 84(7):1216–8.

Sako Y, Minoghchi S, Yanagida T. 2000. Single-molecule imaging of EGFR signalling on the surface of living cells. Nat Cell Biol 2(3):168–72.

Sapsford KE, Berti L, Medintz IL. 2006. Materials for fluorescence resonance energy transfer analysis: beyond traditional donor-acceptor combinations. Angew Chem Int Ed 45(28):4562–88.

Schluesche P, Stelzer G, Piaia E, Lamb DC, Meisterernst M. 2007. NC2 mobilizes TBP on core promoter TATA boxes. Nat Struct Mol Biol 14(12):1196–201.

Schuler B, Eaton WA. 2008. Protein folding studied by single-molecule FRET. Curr Opin Struct Biol 18(1):16–26.

Schuler B, Lipman EA, Eaton WA. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. Nature 419(6908):743–7.

Seeman NC. 2003. DNA in a material world. Nature 421(6921):427–31.

Seeman NC. 2005. From genes to machines: DNA nanomechanical devices. Trends Biochem Sci 30(3):119–25.

Selvin PR. 1995. Fluorescence resonance energy transfer. Methods Enzymol 246:300–34.

Selvin PR. 2000. The renaissance of fluorescence resonance energy transfer. Nat Struct Biol 7(9):730–4.

Sharma S, Chakraborty K, Muller BK, Astola N, Tang YC, Lamb DC, Hayer-Hartl M, Hartl FU. 2008. Monitoring protein conformation along the pathway of chaperonin-assisted folding. Cell 133(1):142–53.

Shroff H, Reinhard BM, Siu M, Agarwal H, Spakowitz A, Liphardt J. 2005. Biocompatible force sensor with optical readout and dimensions of 6 nm(3). Nano Lett 5(7):1509–14.

Smiley RD, Collins TR, Hammes GG, Hsieh TS. 2007. Single-molecule measurements of the opening and closing of the DNA gate by eukaryotic topoisomerase II. Proc Natl Acad Sci USA 104(12):4840–5.

Stryer L, Haugland RP. 1967. Energy transfer: a spectroscopic ruler. Proc Natl Acad Sci USA 58(2):719–26.

Talaga DS, Lau WL, Roder H, Tang J, Jia Y, DeGrado WF, Hochstrasser RM. 2000. Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. Proc Natl Acad Sci USA 97(24):13021–6.

Tan E, Wilson TJ, Nahas MK, Clegg RM, Lilley DM, Ha T. 2003. A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. Proc Natl Acad Sci USA 100(16):9308–13.

Tarsa PB, Brau RR, Barch M, Ferrer JM, Freyzon Y, Matsudaira P, Lang MJ. 2007. Combined optical tweezers force measurement with simultaneous single molecule fluorescence resonance energy transfer (FRET) detection. Biophys J 654A.

Tomishige M, Stuurman N, Vale RD. 2006. Single-molecule observations of neck linker conformational changes in the kinesin motor protein. Nat Struct Mol Biol 13(10):887–94.

Vale RD. 2003. The molecular motor toolbox for intracellular transport. Cell 112(4):467–80.

Vogelsang J, Doose S, Sauer M, Tinnefeld P. 2007. Single-molecule fluorescence resonance energy transfer in nanopipets: improving distance resolution and concentration range. Anal Chem 79(19):7367–75.

Vogelsang J, Kasper R, Steinhauer C, Person B, Heilemann M, Sauer M, Tinnefeld P. 2008. A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. Angew Chem Int Ed Engl 47(29):5465–9.

Walter NG, Huang CY, Manzo AJ, Sobhy MA. 2008. Do-it-yourself guide: how to use the modern single-molecule toolkit. Nat Methods 5(6):475–89.

Webb SE, Roberts SK, Needham SR, Tynan CJ, Rolfe DJ, Winn MD, Clarke DT, Barraclough R, Martin-Fernandez ML. 2008. Single-molecule imaging and fluorescence lifetime imaging microscopy show different structures for high- and low-affinity epidermal growth factor receptors in A431 cells. Biophys J 94(3):803–19.

Weiss S. 1999. Fluorescence spectroscopy of single biomolecules. Science 283(5408):1676–83.

Weiss S. 2000. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. Nat Struct Biol 7(9):724–9.

Wu P, Brand L. 1994. Resonance energy transfer: methods and applications. Anal Biochem 218(1):1–13.

Wu PG, Brand L. 1992. Orientation factor in steady-state and time-resolved resonance energy-transfer measurements. Biochemistry 31(34):7939–47.

Yildiz A, Tomishige M, Vale RD, Selvin PR. 2004. Kinesin walks hand-over-hand. Science 303(5658):676–8.

Yurke B, Turberfield AJ, Mills AP, Jr., Simmel FC, Neumann JL. 2000. A DNA-fuelled molecular machine made of DNA. Nature 406(6796):605–8.

Zhuang X, Bartley LE, Babcock HP, Russell R, Ha T, Herschlag D, Chu S. 2000. A single-molecule study of RNA catalysis and folding. Science 288(5473):2048–51.

Zhuang X, Kim H, Pereira MJ, Babcock HP, Walter NG, Chu S. 2002. Correlating structural dynamics and function in single ribozyme molecules. Science 296(5572):1473–6.

Zhuang X, Rief M. 2003. Single-molecule folding. Curr Opin Struct Biol 13(1):88–97.

Zimmer M. 2002. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. Chem Rev 102(3):759–81.

Zimmermann B, Diez M, Borsch M, Graber P. 2006. Subunit movements in membrane-integrated EF0F1 during ATP synthesis detected by single-molecule spectroscopy. Biochim Biophys Acta 1757(5–6):311–9.

6

Single-Molecule Enzymology

Joseph J. Loparo and Antoine van Oijen

Abstract Understanding how enzymes function requires a thorough characterization of enzymatic dynamics. Traditional enzymatic assays average over an ensemble of molecules, making it difficult to detect reaction intermediates and conformational fluctuations of the enzyme. These problems can be overcome by observing enzymes functioning in real time on the single-molecule level. This chapter describes recent research efforts to measure single-molecule enzyme kinetics and observe the structural dynamics of enzymes and discusses new approaches to study multiprotein complexes on the single-molecule level.

6.1. Introduction

The high rates and specificities by which enzymes transform their substrates and the diversity of reactions they catalyze make enzymology a fascinating area of biochemical research. Chemists' efforts to develop commensurate man-made catalysts require a thorough understanding of how nature has optimized its own. Hypotheses proposed to explain enzyme catalysis have generally relied on structural complementarity between the enzyme and the reactants, intermediates, and products (Benkovic and Hammes-Schiffer, 2003; Garcia-Viloca et al., 2004). Early "lock-and-key" models proposed that the enzyme's structure was uniquely suited to position the substrate in a reactive conformation. Motivated by transition-state theory, later theories have generally built on the "lock-and-key" hypothesis by suggesting that an enzyme preferentially binds the transition state and therefore destabilizes the reactant or ground state. Providing a molecular explanation of enzyme catalysis remains an active and challenging area of research due to the high degree of structural complexity and large number of degrees of freedom present in proteins (Hammes-Schiffer and Benkovic, 2006). A true molecular understanding of enzyme catalysis requires knowledge of not only the structural conformations of the enzyme, but also of how these states affect catalytic function and how the intrinsic fluctuations of the enzyme lead to the interconversion between these states.

Traditionally, enzymology has combined biochemical methods, such as mutagenesis, with structural and kinetic studies to build up a mechanistic picture of enzyme function

J. J. Loparo and A. van Oijen • Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_6,

(Fersht, 1999). Structural studies have generally relied on obtaining "static" pictures of the enzyme either free or bound to a substrate analog through the use of crystallography or nuclear magnetic resonance spectroscopy. Although these studies are able to provide structural detail at the level of single atoms, they only do so for states that are stable and long-lived enough to allow structural characterization. On the other hand, kinetic studies are able to provide rates of conversion between species but do not provide access to high-resolution structural information. Ideally, one would be able to combine these two experimental strategies to simultaneously follow enzyme structure and catalytic activity, effectively obtaining a "molecular movie" of an enzyme transforming its substrate. As described elsewhere in this volume, imaging and mechanical manipulation techniques of individual molecules are beginning to provide such insights into a number of biological processes. In this chapter we describe how these single-molecule methods can be applied to answer long-standing questions regarding the dynamics of enzymes.

6.2. Enzyme Kinetics

Bulk-phase enzyme kinetic studies are generally performed by monitoring the loss of reactants or generation of products through the measurement of spectroscopic observables. Enzyme activity in such an ensemble-averaging approach is well described by the Michaelis-Menten (MM) mechanism (Fersht, 1999):

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^* + P \tag{6.1}$$

In this scheme, a substrate S binds to an enzyme E with a bimolecular rate constant k_1 to form the enzyme substrate complex ES. This complex can either dissociate to yield the starting species, or the catalytic activity of the enzyme transforms ES to the reaction product P and enzyme product E^*. These two processes are unimolecular, depending on only the ES concentration, and are described by the first-order rate constants k_{-1} *and* k_2, respectively. In most cases it is assumed that E^* is equivalent to E or that the reaction $E^* \rightarrow E$ is fast. To obtain a practical expression for enzyme velocity *dP/dt* from the MM scheme, one must consider the following rate equations:

$$\frac{d[E]}{dt} = -k_1[E][S] + k_{-1}[ES] \tag{6.2}$$

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] \tag{6.3}$$

$$\frac{d[P]}{dt} = k_2[ES] \tag{6.4}$$

By noting that the total enzyme concentration is equal to $[E_T] = [E] + [ES]$ and applying the steady-state approximation for ES,

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] = 0 \tag{6.5}$$

one obtains the equation for the enzyme velocity:

$$V = \frac{[E_T][S]k_{cat}}{K_M + [S]} \tag{6.6}$$

where k_{cat} is the rate constant describing the conversion of $[ES]$ to $[P]$ and is equivalent to k_2 in the MM scheme. K_M is the Michaelis constant and is given by

$$K_M = \frac{k_2 + k_{-1}}{k_1} \tag{6.7}$$

In the limit when $k_{-1} >> k_2$ the Michaelis constant reduces to the binding constant of the enzyme–substrate complex:

$$K_M \approx \frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]} \tag{6.8}$$

The MM mechanism predicts a hyperbolic dependence of V on $[S]$. At low $[S]$, V increases linearly with $[S]$. This linearity breaks down at high $[S]$, with S effectively saturating the available substrate binding sites and thus limiting V by k_{cat} and $[E]$.

With only two parameters, K_m and k_{cat}, the MM scheme accurately describes the steady-state activities of countless enzymes. It would be naïve, however, to think that these two parameters are able to fully describe the complexity of enzyme dynamics. For instance, k_{cat} is an average measure of the catalytic rate describing the behavior of a large ensemble of enzymes. Extracting the distribution of rates from an ensemble experiment is difficult if not impossible. This distribution is important, given the many conformational states occupied by a population of enzymes at room temperature. Similarly, the existence of additional enzyme–substrate intermediates beyond ES will cause steady-state measurements of k_{cat} to yield a rate constant that is a combination of the various rate constants.

By their very nature, steady-state experiments are unable to detect short-lived intermediates. As a result, experimentalists often turn to pre–steady-state kinetics in an effort to measure the individual rate constants of these intermediate states. Pre–steady-state experiments are more difficult because they require the synchronization of the enzyme–substrate population. In other words, there needs to be a well-defined time zero indicating the start of the experiment. At this point, the entire population of enzymes and substrates needs to be present in the same starting state. Furthermore, a time resolution is needed that exceeds the turnover time of the enzyme. This requires rapid initiation methods, such as rapid mixing of the enzyme and substrate, or relaxation methods, such as a temperature jump, that "instantaneously" shift the chemical equilibrium.

Even successful measurements of pre–steady-state kinetics can leave ambiguity. Often, a multiexponential population decay of a given reaction intermediate is observed. This phenomenon might be due to a static distribution of enzymatic states with different rate constants or caused by states that are interconverting on the timescale of the experiment (Xie, 2002). Even a single-exponential population decay does not necessarily yield an accurate estimate of the microscopic rate constant but could be due to fast (relative to the experimental time resolution) interconversion of different conformational states with different rate constants. Single-molecule measurements of enzymes are able to resolve these issues by following individual trajectories of an enzyme over many cycles. Analysis of these trajectories quantifies both the distribution of catalytic rates and the dynamic fluctuations in rates of a given enzyme.

Most single-molecule studies have examined enzyme kinetics by measuring the turnover time, τ, the time for a single enzyme to complete one entire reaction cycle. One can connect τ with ensemble experiments by noting that the average turnover time, $\langle\tau\rangle$, can be related to the ensemble enzyme velocity through $1/\langle\tau\rangle = V/[E_T]$ (English et al., 2006). Most experimental observations of the catalytic turnover of single enzymes have been performed using fluorescence-based assays. These experiments have measured τ by relying on enzymes containing intrinsically fluorescent cofactors (Lu et al., 1998) or through the use of fluorogenic substrates (Edman et al., 1999). Standard fluorescence microscopy methods, such as confocal or total internal reflection fluorescence (TIRF), can spatially resolve a sparse density of surface-immobilized enzymes and record their repetitive catalytic activities.

The enzyme cholesterol oxidase is an example of the use of a fluorescent cofactor as an experimental probe (Lu et al., 1998). Cholesterol oxidase utilizes flavin adenine dinucleotide (FAD), which is intrinsically fluorescent, in its active site. The reduction of FAD by cholesterol yields the nonfluorescent $FADH_2$. Subsequent oxidization of $FADH_2$ by molecular oxygen resets the enzyme for another catalytic cycle. From a photochemical perspective, the enzymatic cycle can be viewed in terms of two reactions: a "switching off" and a "switching back on" of the fluorescence. Using fluorescent cofactors that are stably incorporated into the enzyme ultimately limits the observation time due to photobleaching of the cofactor. Fluorogenic substrates use an enzyme to convert a nonfluorescent substrate into a fluorescent product. By using fluorogenic substrates the waiting times between events of product generation can be measured indefinitely, with the observation window limited only by dissociation of the enzyme from the surface or its degradation.

On their own, individual measurements of τ are not particularly useful. Because every chemical reaction is stochastic in nature, individual measurements of transition times will be seemingly random. Many instances taken together, however, yield distributions of turnover times, $f(\tau)$, that can be used to determine rate constants, detect reaction intermediates, and provide distributions of kinetic parameters. To understand how this information is extracted from $f(\tau)$, one must return to the MM formulation but now from a single-molecule perspective. Given that the experimental observations are made on the single-enzyme level, enzyme concentration is now a meaningless quantity. This prompted Xie and coworkers to reformulate the MM approach from the perspective of the probabilities of finding an individual enzyme in a specific state (Xie, 2001). Returning to Eq. (6.1) describing the MM scheme, and following the derivation of Xie and coworkers, we can write the following differential equations:

$$\frac{dP_E(t)}{dt} = -k_1^o P_E(t) + k_{-1} P_{ES}(t) \tag{6.9}$$

$$\frac{dP_{ES}(t)}{dt} = k_1^o P_E(t) - (k_{-1} + k_2) P_{ES}(t) \tag{6.10}$$

$$\frac{dP_{E^o}(t)}{dt} = k_2 P_{ES}(t) \tag{6.11}$$

where $k_1^o = k_1[S]$ and $P(t)$ is the probability of finding the enzyme in the specified state. The differential equations can be solved by noting that the sum of the probabilities must equal 1, $P_E(t) + P_{ES}(t) + P_{E^o}(t) = 1$, and that initially (i.e., at $t = 0$) the enzyme is in state E, $P_E(0) = 1$. A $f(\tau)$ distribution that corresponds to the time to complete the MM cycle is given by

$$f(\tau) = \left.\frac{dP_{E^o}(t)}{dt}\right|_{t=\tau} = k_2 P_{ES}(\tau) \tag{6.12}$$

Solving for $f(\tau)$ yields

$$f(\tau) = \frac{k_1^o k_2}{2a}\{\exp[(a+b)\tau] - \exp[(b-a)\tau]\} \tag{6.13}$$

where

$$a = \sqrt{\frac{1}{4}(k_1^o + k_{-1} + k_2)^2 - k_1^o k_2} \text{ and } b = -\frac{1}{2}(k_1^o + k_{-1} + k_2).$$

The behavior of $f(\tau)$ is sensitive to the kinetic parameters within the MM scheme. An examination of the behavior of $f(\tau)$ as a function of the number of rate-limiting steps present within the kinetic scheme reveals how mechanistic insight can be obtained from these distributions. In the limit in which there is only a single rate-limiting step, be it association of enzyme and substrate (i.e., $k_1 >> k_2$) or the conversion of *ES* to product (i.e., $k_2 >> k_1$), $f(\tau)$ reduces to a single-exponential decay, as one expects for a stochastic process. In the limit of two rate-limiting steps, for example, if *ES* is particularly stable ($k_{-1} = 0$), one needs to determine the sequential probability density of two stochastic processes. This is equivalent to the convolution of two exponentials, which yields

$$f(\tau) = \frac{k_1^o k_2}{k_2 - k_1^o}\{\exp[-k_1^o\tau] - \exp[-k_2\tau]\} \tag{6.14}$$

Therefore, in the presence of two rate-limiting steps, $f(\tau)$ exhibits an exponential rise and decay, indicating the presence of an enzyme–substrate intermediate. A number of N rate-limiting steps results in the convolution of N exponentials. Assuming the rates (k) for each intermediate step are equal, then $f(\tau)$ becomes equal to the gamma distribution:

$$f(\tau) = \frac{k^N \tau^{N-1} \exp[-k\tau]}{(N-1)!} \tag{6.15}$$

As N increases, the width of $f(\tau)$ decreases relative to the mean, and the shape approaches a Gaussian, a result of the Central Limit Theorem (see Figure 6.1).

Proper determination of $f(\tau)$ requires not only a large number of events, but also a dynamic range covering both short and long turnover times. Depending on the experimental conditions, this may be difficult or impossible to achieve; for example, photobleaching may limit the observation window. To circumvent this problem, Schnitzer and Block (1995) suggested the following randomness parameter, r, to characterize $f(\tau)$:

$$r = \frac{\langle\tau^2\rangle - \langle\tau\rangle^2}{\langle\tau\rangle^2} \tag{6.16}$$

If $f(\tau)$ is exponential, then $r = 1$, indicating that the spread in τ is large relative to the mean of τ. The value of r decreases with an increase in N due to the decreasing width of $f(\tau)$.

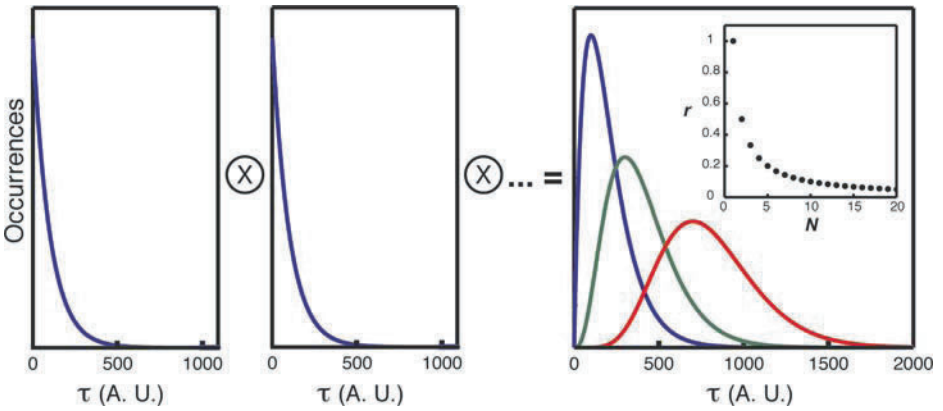


Figure 6.1. A convolution of N exponentials with rate constant k results in a turnover distribution that is described by the gamma distribution. As N increases (N = 2 [*blue*], N = 4 [*green*], and N = 8 [*red*]), the width of the distribution decreases relative to the mean, a result that is captured in the randomness parameter, r, shown in the inset.

Effectively, an enzyme with multiple intermediates is more "regular" in its generation of product than one with only a few.

So far, we have assumed that the kinetics of individual enzymes can be well described by a single set of rate constants. A number of single-molecule studies have concluded that single enzymes have different catalytic rates that can differ by an order of magnitude (Edman et al., 1999; English et al., 2006; Flomenbom et al., 2005; Lu et al., 1998; Rissin et al., 2008; van Oijen et al., 2003). To illustrate the strength of single-molecule experiments to reveal the heterogeneities within a population of enzymes, it is worth looking at a recent example from the literature. Xie and coworkers examined the kinetics of β-galactosidase using the fluorogenic substrate resorufin–β-D-galactopyranoside (RGP) (English et al., 2006). RGP releases fluorescent resorufin on hydrolysis by β-galactosidase, resulting in a fluorescent burst. Single enzymes were imaged confocally, and τ was defined as the time between fluorescent bursts. An intensity trajectory of a single β-galactosidase is shown in Figure 6.2A. The form of $f(\tau)$ was found to be dependent on substrate concentration as shown in Figure 6.2B. At low [S], $f(\tau)$ is a single-exponential decay because of the slow and rate-limiting nature of enzyme–substrate association. As substrate concentration was increased, the catalytic step k_2 became rate limiting. However, a multiexponential behavior of $f(\tau)$ was observed even at saturating substrate concentrations. Such multiexponential behavior is not anticipated with only a single rate-limiting step. The researchers attributed this behavior to a distribution of *ES* conformers whose interconversion is slower than the timescale of an enzymatic turnover. Even on the single-molecule level, fast interconversion of conformers, as compared to the turnover time, averages the measured catalytic rate across the population of enzyme conformations. The expression for $f(\tau)$ may be modified to capture relatively long-lived heterogeneity by integrating over the distribution of k_2, $w(k_2)$:

$$f(\tau) = \int_0^\infty dk_2 w(k_2) \frac{k_1^o k_2}{2a} \{\exp[(a+b)\tau] - \exp[(b-a)\tau]\} \tag{6.17}$$

Through global fitting of the $f(\tau)$ data, Xie and coworkers were able to extract $w(k_2)$.

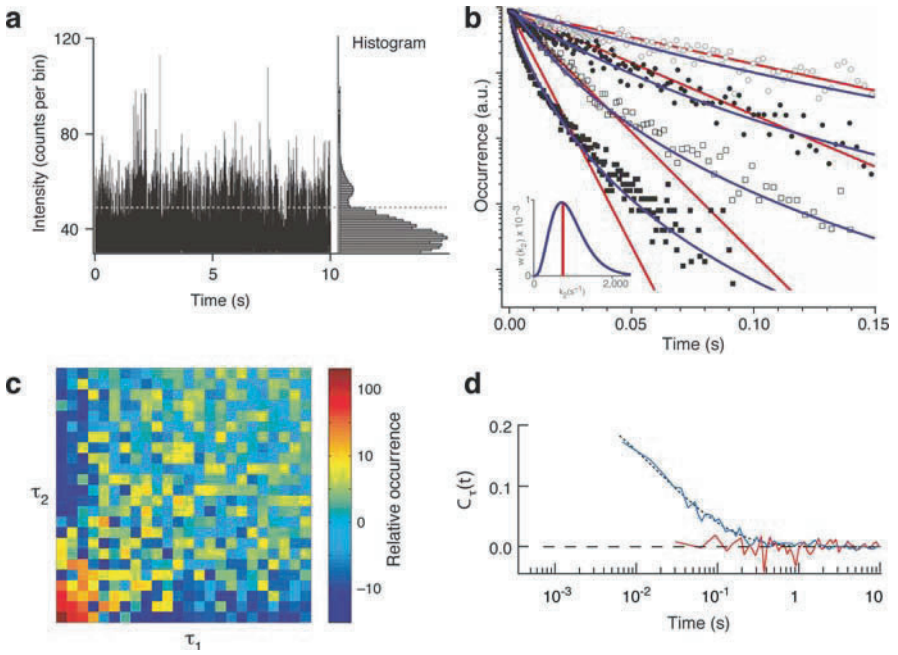


Figure 6.2. A. Turnover trajectory for a single β-galactosidase enzyme at a substrate concentration of 100 μM resorufin–β-D-galactopyranoside (RGP). Turnover times correspond to the time between intensity spikes that surpass the intensity threshold (*dotted line*). **B.** The turnover distribution is dependent on the substrate concentration. Concentrations of RGP are 10 μM (*open circles*), 20 μM (*filled circles*), 50 μM (*open squares*), and 100 μM (*filled squares*). Red lines correspond to single-exponential fits and show that $f(\tau)$ becomes multiexponential with increasing substrate concentration. Blue lines correspond to global fits with a model that includes a distribution of enzymes with different catalytic rates (shown in the inset). **C.** Two-dimensional difference turnover histogram showing that subsequent turnovers are generally correlated. **D.** The loss of correlation between turnovers is quantified in the turnover correlation function, $C_\tau(t)$. [Adapted from English et al. (2006).]

What are the timescales for the interconversion of these different conformers? Through analysis of individual enzymes one is able to observe how τ changes with time. If distinct conformers have different catalytic rates, one would anticipate that subsequent τ's of the same enzyme should be correlated. In other words, a relatively fast turnover time should be followed by another fast turnover. As one looks beyond a few turnovers, this correlation should begin to disappear as the conformers begin to interconvert. To look for correlations, Xie and coworkers plotted two-dimensional (2D) joint histograms of τ's as a function of the time between them. Correlation appears as an elongation of the 2D histogram along the diagonal axis, $\tau_1 = \tau_2$. In the limit of no correlation (i.e., at large separation times between the turnovers) the histogram is simply the product of the two turnover distributions, $h(\tau_1,\tau_2) = f(\tau_1) f(\tau_2)$. Clear broadening along the diagonal was observed for difference 2D histograms of β-galactosidase with short and long separations between turnovers (Figure 6.2C). Quantifying the loss of correlation as a function of time is difficult using 2D joint τ histograms, so Xie and coworkers turned to calculating turnover correlation functions:

$$C_\tau(m) = \frac{\langle \Delta\tau(0)\Delta\tau(m)\rangle}{\langle \Delta\tau^2\rangle} \tag{6.18}$$

where $\Delta\tau(m) = \tau(m) - \langle\tau\rangle$, with m as the index of turnovers. Converting from m to time was accomplished by noting that $t = m\langle\tau\rangle$. $C_\tau(m)$ quantifies the loss of "enzymatic memory" due to the interconversion of conformers and is shown in Figure 6.2D. For β-galactosidase, $C_\tau(m)$ was found to decay over multiple timescales from milliseconds to seconds, indicating that the structural fluctuations of the enzyme occur over multiple decades in time.

6.3. Conformational Fluctuations and Dynamics

In the previous section we described how single-molecule measurements are able to reveal distributions of catalytic activity. This heterogeneity in k_{cat} (also referred to as k_2 in the previous section) is due to a distribution of conformers that interconvert on a timescale slower than the time of a single enzymatic turnover. Although measurements of individual turnover events are able to show the existence of different conformers, they are unable to provide any direct structural information. In this section we describe single-molecule efforts aimed at understanding the interconversion of enzyme conformers and how these states affect catalytic function.

Protein structure and dynamics are so complicated due to the large number of degrees of freedom. In the context of protein folding, this leads to the well-known Levinthal paradox (Dill et al., 2008). How can a protein find its native state in a reasonable amount of time if it has to randomly search through the myriad other conformations? Modern answers to this question have generally focused on conformational searching through an energy landscape in which proteins are funneled from nonnative conformations into the global energy minimum in which the native fold resides (Frauenfelder et al., 1991). Such a framework minimizes the importance of specific folding pathways but instead emphasizes the heterogeneous nature of the folding process. Single-molecule techniques are particularly powerful, given their ability to quantify such distributions.

Roughness in the energy landscape leads to local minima corresponding to relatively stable conformers. The interconversion time of these conformers is directly related to the free energy barrier height between them, a scenario that is illustrated schematically in Figure 6.3. From the Arrhenius equation, the rate of interconversion of two conformers is given by $k = Ae^{-E_a/RT}$, where A is the prefactor and E_a is the activation energy. The prefactor A can be understood as the rate of the chemical process in the absence of an energy barrier, with its value depending on the nature of the reaction. Isolated chemical reactions (i.e., gas-phase reactions) generally have A^{-1} in the femtosecond range. Proteins are slower due to the numerous noncovalent interactions they are capable of forming within the polypeptide chain and with the surrounding solvent. Some studies have suggested "speed limits" for the formation of secondary structure on the order of $A^{-1} = 1\ \mu s$ (Kubelka et al., 2004). These estimates are largely based on polymer theory and can be understood as the time for a random coil to collapse on itself. The fluctuations that drive the interconversion of enzyme conformers can be estimated to be somewhere between these two limits. Given the timescales of memory loss in single-molecule enzyme turnover distributions and assuming a value for A^{-1} of 10^{-12} sec, Xie (2002) estimated that the barrier heights for the interconversion of enzyme conformers ranged between 0.5 kcal/mol ($\sim 1\ k_B T$), corresponding to picosecond fluctuations, to 16 kcal/mol ($\sim 30\ k_B T$) for dynamics on the second timescale.

Single-molecule experiments have generally measured structural fluctuations in proteins by measuring a distance between two residues. We will describe two common methods for determining distance: electron transfer (ET) (Lu and Xie, 1997; Neuweiler et al., 2003;

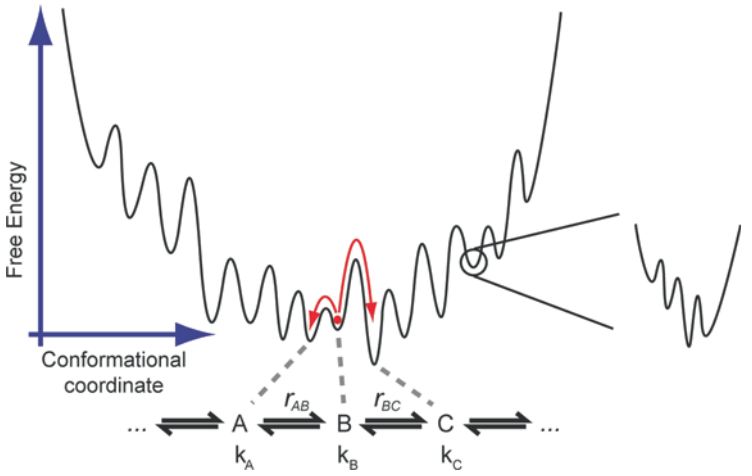


Figure 6.3. Enzyme dynamics occurs on a corrugated free energy landscape. Distinct enzyme conformations (A, B, C,. . .) interconvert through motion along the conformational coordinate, *r*. The timescale of this interconversion depends on the free energy barrier height, which can vary by orders of magnitude.

Yang et al., 2003) and fluorescence resonance energy transfer (FRET) (Roy et al., 2008). Electron transfer is observed as a decrease in the fluorescence lifetime. This reduction in lifetime occurs because ET provides another quenching pathway for the relaxation of the excited electronic state of the chromophore. Therefore, the measured fluorescence lifetime is equal to $\gamma^{-1}(t) = [\gamma_o + k_{ET}(t)]^{-1}$, where γ_o is the fluorescence lifetime in the absence of quenching and $k_{ET}(t)$ is the fluctuating ET rate. The rate of ET is described by the Marcus equation, the details of which are not important for our discussion (Stubbe et al., 2003). In the context of measuring conformational fluctuations, the distance dependence between the electron donor and acceptor of the ET rate constant, k_{ET}, is critical. ET is exponentially dependent on the distance $k_{ET}(t) = k_o e^{-\beta(r(t))}$, where β is $\sim$1.4 for proteins (Gray and Winkler, 1996). Experiments have generally used a naturally present amino acid such as tyrosine to be the electron acceptor. Back transfer from the donor to acceptor occurs on the order of a few nanoseconds. Fluctuations in the fluorescence lifetime can be attributed to changes in k_{ET} and therefore fluctuations in $r(t)$.

Yang et al. used single-molecule ET as a probe to measure the conformational fluctuations of Fre, a flavin reductase from *Escherichia coli* (Yang et al., 2003). Fre contains the fluorescent cofactor FAD, which can transfer an electron to a neighboring tyrosine residue. Fluorescence lifetimes were measured using a pulsed laser to excite FAD and single-photon counting methods to detect emitted fluorescence (see Figure 6.4A). In such an approach, the photon delay time, τ_p, is merely the time between the excitation pulse and the emitted photon (Figure 6.4B). Plotting a histogram of τ_p yields a single-molecule fluorescent decay similar to ensemble measurements of fluorescent lifetime.

A particular strength of this single-molecule approach is the ability to look at fluctuations in τ_p as a function of time. A "running" average fluorescence lifetime as a function of time can be calculated by binning the τ_p trajectory over 100 photons as is shown in Figure 6.4C. Given the photon counting rate in the Fre study, this bin size corresponded to 200 ms. Lifetimes for FAD in the Fre system were found to vary from 30 ps to $\sim$3 ns. Given the relationship between the electron donor and acceptor distance r and fluorescence lifetime γ,

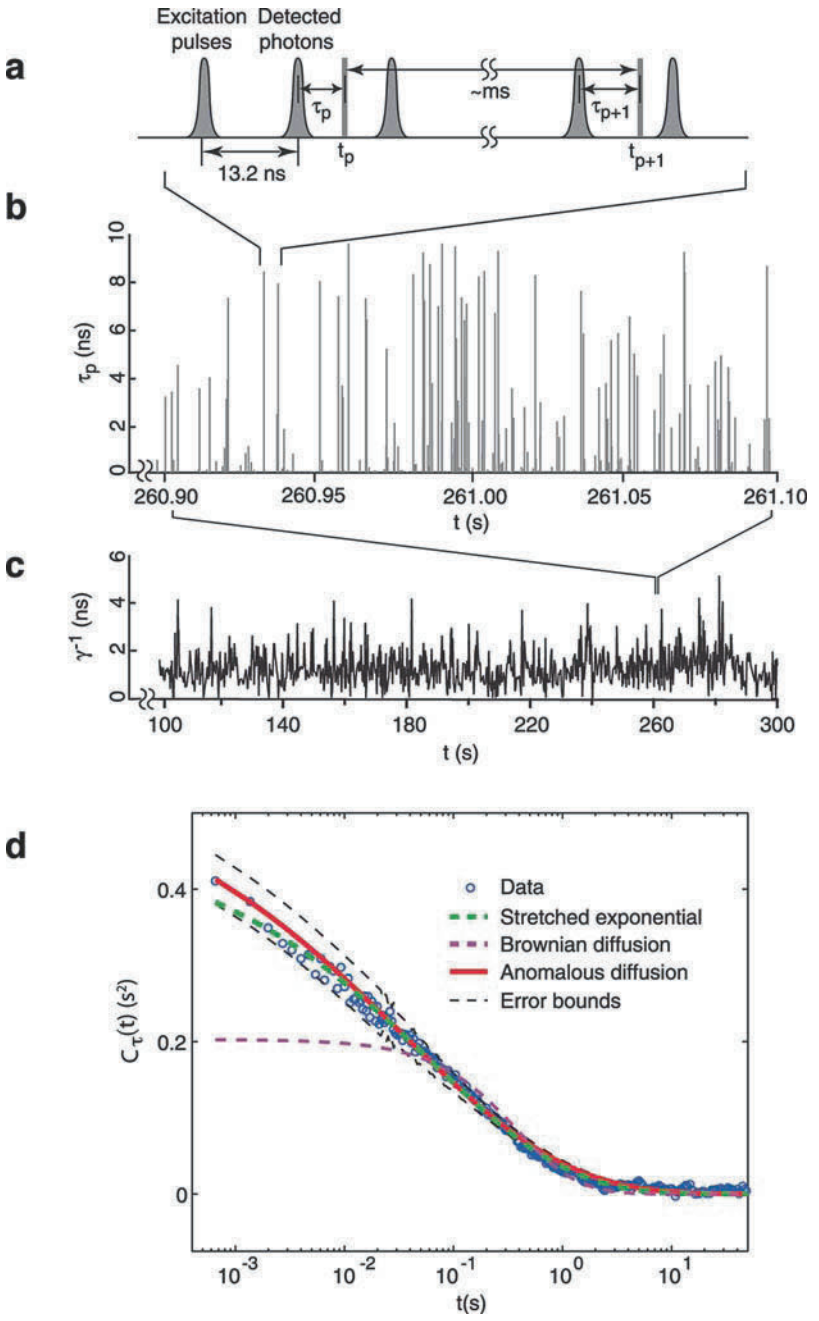


Figure 6.4. A. Diagram of excitation pulses and detected fluorescence photons. **B.** Plot of τ_p as a function of the photon arrival time. **C.** Fluctuations in the fluorescence lifetime for a single Fre/flavin adenine dinucleotide (FAD) complex. **D.** Fluctuations in the fluorescence lifetime are captured by the correlation function, $C_\gamma(t)$. [Adapted from Yang et al. (2003).]

one can build a probability distribution $P(r)$ for the distance from the lifetime histogram. The potential of mean force describing the dynamics of the FAD–Tyr pair can be obtained from $P(r)$ by using

$$V(r) = -k_B T \ln [P(r)] \tag{6.19}$$

The dynamics of $V(r)$ is captured by the lifetime correlation function, shown in Figure 6.4D: $C_\gamma(t) = \langle \delta\gamma^{-1}(t)\delta\gamma^{-1}(0)\rangle$, where $\delta\gamma^{-1}(t) = \gamma^{-1}(t) - \langle\gamma^{-1}\rangle$. The loss of correlation occurs over multiple decades in time, from hundreds of microseconds to tens of seconds, strongly supporting previous hypotheses that memory effects in enzyme turnover trajectories were due to conformational fluctuations.

Some caution is required in interpreting $V(r)$, in that it is a projection of the multidimensional free energy landscape onto only one coordinate, r. For the Fre system, $V(r)$ was found to be smoothly varying and nearly harmonic. Simulations of Brownian diffusion on this potential are unable to reproduce $C_\gamma(t)$; instead, anomalous diffusion was necessary to explain the data. This indicates that the "actual" potential is much more rugged than $V(r)$, which is time averaged over the time window necessary to sufficiently average τ_p. The study of Yang et al. demonstrates the ability of single-molecule ET to probe equilibrium fluctuations of an enzyme.

FRET experiments, described in greater detail in a separate chapter in this volume, monitor the nonradiative energy transfer between a donor and an acceptor dye. Much like ET, FRET serves as a competing pathway for the relaxation of the electronic excited state. The rate of energy transfer from the donor to acceptor is dependent on the distance and orientation of the dyes, as well as on the spectral overlap and quantum yield of the donor. For dyes that are freely rotating it is often the case that the orientational factor averages out to the value for isotropic rotation (Lackowicz, 1999). The FRET efficiency, E, is given by

$$E = \left[1 + \left(\frac{R}{R_o}\right)^6\right]^{-1} \tag{6.20}$$

where R_o is the Förster radius, the distance of 50% FRET efficiency, and R is the distance between the donor and acceptor dyes. Typical values of R_o are on the order of a few nanometers, which correspond to and describe roughly the range of distances over which FRET is most effective.

Experimental measures of E are performed ratiometrically by exciting the donor dye and subsequently monitoring the fluorescence (F) from both the donor and the acceptor:

$$E = \frac{F_A}{F_A + \chi F_D} \tag{6.21}$$

where $\chi = \frac{\eta_A \phi_A}{\eta_D \phi_D}$ corrects for the detection efficiencies (η) and quantum yields (ϕ) of the acceptor and donor. Given that the fluorescence intensities of donor and acceptor should be anticorrelated, measuring both signals helps to discern real changes in E from noise. Practically, donor and acceptor dyes are generally covalently coupled to the protein. Attaching dyes to specific residues of the protein and doing so in a manner that does not impair catalytic activity is one of the major challenges of a typical single-molecule FRET experiment.

FRET experiments have probed a number of systems, investigating phenomena such as protein folding and the equilibrium conformational fluctuations of proteins; a number of

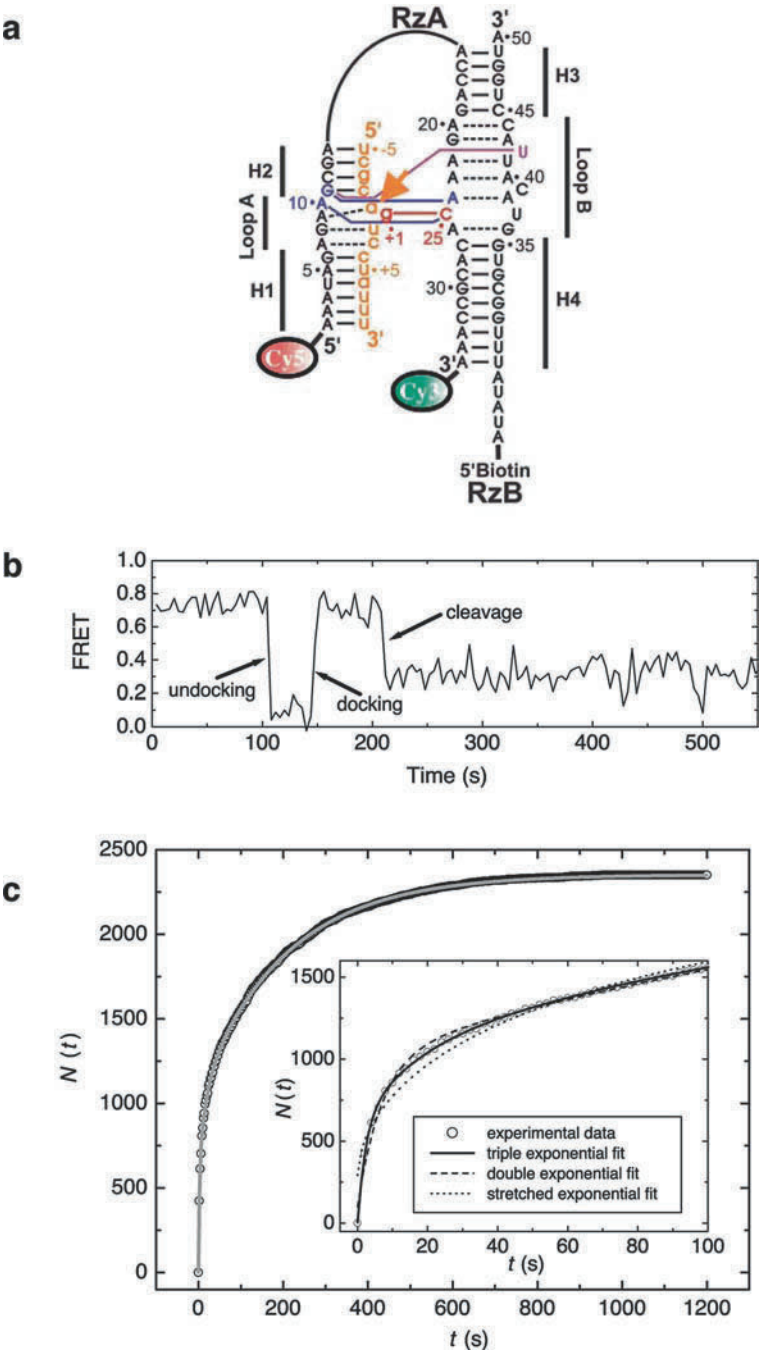


Figure 6.5. **A.** The structure of the hairpin ribozyme. The orange strand refers to the substrate, and the orange arrow indicates the cleavage site. **B.** Fluorescence resonance energy transfer (FRET) time trace of a single ribozyme undergoing a catalytic cycle. **C.** The dwell-time distribution in the docked state is complex, suggesting four distinct docked states. [Adapted from Zhuang et al. (2002).]

reviews have recently been published (Joo et al., 2008; Michalet et al., 2006; Toprak and Selvin, 2007). A demonstration of the utility of single-molecule FRET for probing the interplay of structural dynamics and function is provided by Zhuang et al. (2002), who studied a hairpin ribozyme. The ribozyme is the minimal active form derived from the autocatalytic negative strand of the tobacco ringspot virus satellite RNA, which cleaves RNA into 32P and 52P products. Approximately 50 nucleotides in length, the ribozyme is comprised of a two helix-loop-helix domains A and B. Previous biochemical studies showed that the hairpin ribozyme could form an extended "undocked" conformation with the two domains coaxially stacked or a bent "docked" conformation with tertiary contacts between loops A and B (Walter et al., 1998, 1999).

Single-molecule measurements were performed on the ribozyme by tethering the 52 end of domain B to the surface and attaching the Cy3 donor dye to the 32 end of B and Cy5 to the 52 end of domain A (see Figure 6.5A). Histograms of E were comprised of three clear distributions: the undocked low-FRET state, the docked high-FRET state, and an intermediate-FRET state that results when the substrate is not bound to the ribozyme. Trajectories of single ribozymes clearly showed the transition between these states with cleavage of the substrate only occurring from the docked state. A single ribozyme trajectory is shown in Figure 6.5B. Surprisingly, analysis of the dwell times in the docked state, shown in Figure 6.5C, revealed multiexponential behavior, indicating the presence of four distinct docked states. Previous ensemble experiments were unable to resolve these states due to their relatively short lifetime, which results in little accumulation of these species. In addition, individual ribozymes seemed to repeat similar dwell times in the docked state as if they "remembered" the structural features of the docked state in which they had previously resided. Mutation experiments suggested that these distinct docked states were the result of disruption in interloop hydrogen bonds, while the long memory time indicates that the interconversion of these states is slow.

6.4. Enzymology of Multiprotein Complexes

Up to this point our discussion has focused on observing the dynamics of single enzymes as they chemically transform their substrates. However, most biochemical pathways involve the concerted activity of multiple proteins working together. The requirement for sufficiently high protein concentrations to facilitate protein–protein interaction is seemingly incompatible with single-molecule experiments that traditionally have relied on reducing protein concentration to such a sparse level that individual fluorescent proteins can be resolved by optical means. Standard fluorescence microscopes operating at visible wavelengths can achieve a minimal, diffraction-limited spot size of $\sim$300 nm, which corresponds to an excitation volume of $\sim$0.1 femtoliter (fL) (femto $= 10^{-15}$). To have on average only one molecule in this volume, the concentration of fluorescent molecules must be on the order of $\sim$1 nM. Protein–protein affinity is quantified by the equilibrium dissociation constant K_d, which can be $\sim$1 μM or higher.

There are multiple ways to address this disparity in concentrations. Depending on the experiment, it is possible to fluorescently label only a fraction of a given protein population. Yang and Musser (2006) used this approach to examine the concentration dependence of transport across the nuclear pore complex. A component of the import complex, NLS-2xGFP(4C), was fluorescently labeled and mixed with unlabeled cargo. Rates of transport of the fluorescent cargo were found to be dependent on the total concentration of cargo.

However, this approach will only work in situations in which one only needs to visualize one of the protein components.

As a more generally applicable approach, one can attempt to reduce the optical excitation volume, effectively moving beyond the diffraction limit and thus reducing the number of fluorescent molecules observed. Both TIRF and confocal microscopy, which are well-established modalities in single-molecule imaging, rely on this approach to reduce background fluorescence and improve signal-to-noise ratios. Recent studies have demonstrated significant reductions of the excitation volume by using subwavelength nanostructures known as zero-mode waveguides (Moran-Mirabel and Craighead, 2008; Stark et al., 2007). Zero-mode waveguides consist of a sub–100 nm metal film, such as aluminum, deposited onto a glass substrate. Arrays of apertures anywhere from 30 to 80 nm in diameter can be patterned using electron-beam or ion-beam lithography techniques. Due to the aperture's subwavelength size, visible light does not propagate beyond the aperture, but instead an evanescent field is generated at the aperture entrance. Therefore, the excitation volume is determined by the aperture size and the length scale of the evanescent decay, amounting to volumes as small as attoliters (aL) (10^{-18}) to zeptoliters (zL) (10^{-21}). Webb and coworkers have demonstrated the power of this approach by using zero-mode waveguides in combination with fluorescence correlation spectroscopy to observe the incorporation of fluorescent nucleotides at micromolar concentrations by the T7 DNA polymerase (Levene et al., 2003).

Instead of confining the excitation light, one can confine the molecules of interest into a small volume and thus increase the local concentration. A number of recent studies have demonstrated the encapsulation of proteins within small lipid vesicles with a diameter on the order of 50–200 nm, resulting in a typical volume of $\sim$1 aL (Boukobza et al., 2001; Okumus et al., 2004; Rhoades et al., 2003). One molecule in this volume would correspond to a concentration of $\sim$1 μM. Although they are somewhat larger in diameter ($\sim$700 nm), one can also encapsulate proteins within water drops in an oil emulsion (Tang et al., 2008). Lipid vesicles can be surface-tethered and functionalized to include pores to allow for the diffusion of small molecules into the vesicle interior (Cisse et al., 2007). In addition, emulsified single aqueous nanodroplets have been optically trapped (Tang et al., 2008). Chen and coworkers demonstrated the utility of lipid encapsulation by characterizing the weak protein–protein interactions of two copper transport proteins (Benitez et al., 2008).

In a fourth approach, proteins can be maintained at high concentration and their substrates can be made sparse. This strategy has been used in mechanical assays that allow for the stretching of individual DNA molecules and the accurate measurement of their length or the position of proteins bound to the DNA. In these experiments, the DNA is kept at low concentrations, and the untagged and invisible proteins are introduced at concentrations sufficiently high to form multiprotein complexes on the DNA.

A recent example from van Oijen and coworkers is the single-molecule observation of the activity of the multienzyme machinery, known as the replisome, responsible for replicating DNA (Hamdan et al., 2007; Lee, 2006; Tanner et al., 2008). Mediating DNA replication in a highly synchronized manner, the replisome consists of a DNA helicase that unwinds the double-stranded DNA; DNA polymerases that synthesize DNA on each of the two template strands; a DNA primase that synthesizes RNA primers to initiate lagging-strand synthesis; and single-stranded DNA binding protein [for a review of prokaryotic DNA replication see Benkovic et al. (2001)]. Because the leading- and lagging-strand polymerases both synthesize DNA in a 52 to 32 direction, they are moving in opposite directions on two antiparallel DNA strands, making it difficult to coordinate the synthesis of both DNA strands during replication. Alberts et al. (1983) proposed a "trombone model" in which a replication loop

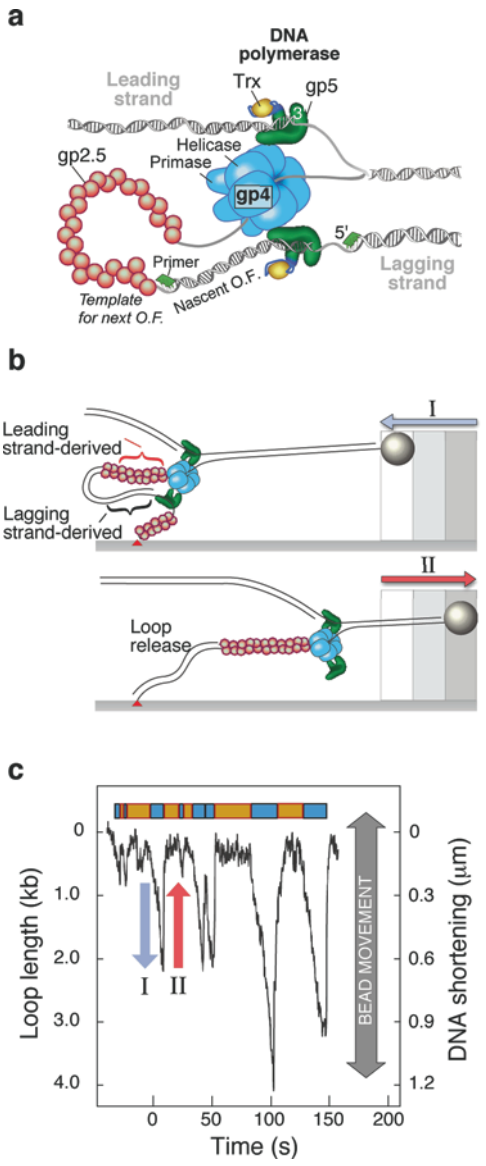


Figure 6.6. A. Cartoon of the T7 replisome, illustrating the formation of a replication loop on the lagging strand. **B.** Experimental scheme for watching replication loop dynamics. Imaging of the bead results in a real-time measure of DNA length. **C.** Trajectory of DNA length as a function of time. Loop growth and lag phases between looping events are denoted by the cyan and orange boxes, respectively. Phase I (*blue arrow*) refers to loop growth, and phase II (*red arrow*) indicates loop release. O.F., Okazaki fragment. [Adapted from Hamdan et al. (2008).]

(Figure 6.6A) formed on the lagging strand of the replication fork reorients the lagging-strand DNA polymerase so that it advances in parallel with the leading-strand polymerase. The replication loop grows and shrinks (like the slide of a trombone) during each cycle of DNA synthesis, producing short fragments of DNA called Okazaki fragments. Half of the

replication loop is the double-stranded product of DNA polymerase, and the other half is the single-template for the next Okazaki fragment, which is extruded by the helicase and coated with single-stranded DNA–binding protein. Decades of biochemical, genetic, and structural studies have elucidated the main components and activities of the replisome, but there is no clear understanding of the dynamics of replication loop formation and its coordination with the various enzymatic activities at the fork. Hamdan et al. (2008) used single-molecule techniques to visualize the real-time formation and release of replication loops by individual replisomes of bacteriophage T7. One end of a 48.5 kb-long duplex λ-phage DNA molecule is attached to the bottom surface of a glass flow cell and the other end to a bead. A constant laminar flow is applied above the surface such that the resultant drag on the beads stretches the DNA molecules with a well-defined force in the low-pN range. Changes in the lengths of the individual DNA molecules are measure by imaging the beads and tracking their positions. On introducing the four proteins that form the T7 replisome (the helicase-primase protein, the polymerase with processivity factor, and single-stranded DNA–binding protein) the DNA length changes, indicating replication. The gradual shortening of the DNA length corresponds to the formation of a replication loop (Figure 6.6B), followed by the sudden release of the loop and the "snapping back" of the DNA length to the original value. The discontinuous synthesis of Okazaki fragments on the lagging strand was observed as a repeated expansion and release of the replication loop (Figure 6.6C). Analyses of the distributions of loop sizes and lag times between loops reveal that initiation of primer synthesis and the completion of an Okazaki fragment each serves as a trigger for loop release. The presence of two triggers may represent a fail-safe mechanism ensuring the timely reset of the replisome after the synthesis of every Okazaki fragment.

In summary, single-molecule techniques have enabled new insights into the complexities of enzyme dynamics. Characterization of enzyme turnover kinetics has revealed that enzyme populations are surprisingly heterogeneous, displaying a distribution of catalytic rates. Studies of enzyme conformational fluctuations have indicated that structural dynamics occur over multiple decades in time, a finding consistent with the hypothesis that the distribution of catalytic rates is a result of interconverting conformers of varying catalytic activity. Future directions in single-molecule enzymology require new experiments that correlate enzyme structure with activity and that are capable of probing multiprotein complexes.

References

Alberts, B. M., J. Barry, P. Bedinger, T. Formosa, C. V. Jongeneel, and K. N. Kreuzer. 1983. Studies on DNA replication in the bacteriophage T4 *in vitro* system. *Cold Spring Harb Symp Quant Biol* 47 Pt 2:655–668.

Benitez, J. J., A. M. Keller, P. Ochieng, L. A. Yatsunyk, D. L. Huffman, A. C. Rosenzweig, and P. Chen. 2008. Probing transient copper chaperone-Wilson disease protein interactions at the single-molecule level with nanovesicle trapping. *J Am Chem Soc* 130(8):2446–2447.

Benkovic, S. J., and S. Hammes-Schiffer. 2003. A perspective on enzyme catalysis. *Science* 301:1196–1202.

Benkovic, S. J., A. M. Valentine, and F. Salinas. 2001. Replisome-mediated DNA replication. *Ann Rev Biochem* 70:181–208.

Boukobza, E., A. Sonnenfeld, and G. Haran. 2001. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *J Phys Chem B* 105:12165–12170.

Cisse, I., B. Okumus, C. Joo, and T. Ha. 2007. Fueling protein DNA interactions inside porous nanocontainers. *Proc Natl Acad Sci USA* 104(31):12646–12650.

Dill, K. A., S. B. Ozkan, M. S. Shell, and T. R. Weikl. 2008. The protein folding problem. *Annu Rev Biophys* 37: 289–316.

Edman, L., Z. Foldes-Papp, S. Wennmalm, and R. Rigler. 1999. The fluctuating enzyme: a single molecule approach. *Chem Phys* 247:11–22.

English, B. P., W. Min, A. M. van Oijen, K. T. Lee, G. Luo, H. Sun, B. J. Cherayil, S. C. Kou, and X. S. Xie. 2006. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat Chem Biol* 2(2):87–94.

Fersht, A. 1999. *Structure and Mechanism in Protein Science*. San Francisco: W. H. Freeman.

Flomenbom, O., K. Velonia, D. Loos, S. Masuo, M. Cotlet, Y. Engelborghs, J. Hofkens, A. E. Rowan, R. J. Nolte, M. Van der Auweraer, F. C. de Schryver, and J. Klafter. 2005. Stretched exponential decay and correlations in the catalytic activity of fluctuating single lipase molecules. *Proc Natl Acad Sci USA* 102(7):2368–2372.

Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science* 254(5038):1598–1603.

Garcia-Viloca, M., J. Gao, M. Karplus, and D. G. Truhlar. 2004. How enzymes work: analysis by modern rate theory and computer simulations. *Science* 303:186–195.

Gray, H. B., and J. R. Winkler. 1996. Electron transfer in proteins. *Annu Rev Biochem* 65:537–561.

Hamdan, S. M., D. E. Johnson, N. A. Tanner, J. B. Lee, U. Qimron, S. Tabor, A. M. van Oijen, and C. C. Richardson. 2007. Dynamic DNA helicase–DNA polymerase interactions assure processive replication fork movement. *Mol Cell* 27(4):539–549.

Hamdan, S. M, J. J. Loparo, M. Takahashi, C. C. Richardson, and A. M. van Oijen. 2009. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* 457:336–339.

Hammes-Schiffer, S., and S. J. Benkovic. 2006. Relating protein motion to catalysis. *Annu Rev Biochem* 75: 519–541.

Joo, C., H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha. 2008. Advances in single-molecule fluorescence methods for molecular biology. *Annu Rev Biochem* 77:51–76.

Kubelka, J., J. Hofrichter, and W. A. Eaton. 2004. The protein folding 'speed limit'. *Curr Opin Struct Biol* 14: 76–88.

Lackowicz, J.R. 1999. *Principles of Fluorescence Spectroscopy*. New York: Kluwer Academic/Plenum Press.

Lee, J.-B., Hite, R. K, Hamdan, S. M, Xie, X. S., Richardson, C. C., and van Oijen, A. M. 2006. DNA primase acts as a molecular brake during DNA replication. *Nature* 439(7076):621–624.

Levene, M. J., J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, and W. W. Webb. 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 299(5607):682–686.

Lu, H. P., and X. S. Xie. 1997. Single-molecule kinetics of interfacial electron transfer. *J Phys Chem B* 101(15): 2753–2757.

Lu, H. P., L. Xun, and X. S. Xie. 1998. Single-molecule enzymic dynamics. *Science* 282(5395):1877–1882.

Michalet, X., S. Weiss, and M. Jager. 2006. Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem Rev* 106:1785–1813.

Moran-Mirabel, J. M., and H. G. Craighead. 2008. Zero-mode waveguides: sub-wavelength nanostructures for single molecule studies at high concentrations. *Methods* 46(1):11–17.

Neuweiler, H., A. Schulz, M. Bohmer, J. Enderlein, and M. Sauer. 2003. Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *J Am Chem Soc* 125:5324–5330.

Okumus, B., T. J. Wilson, D. M. J. Lilley, and T. Ha. 2004. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys J* 87:2798–2806.

Rhoades, E., E. Gussakovsky, and G. Haran. 2003. Watching proteins fold one molecule at a time. *Proc Natl Acad Sci USA* 100:3197–3202.

Rissin, D. M., H. H. Gorris, and D. R. Walt. 2008. Distinct and long-lived activity states of single enzyme molecules. *J Am Chem Soc* 130(15):5349–5353.

Roy, R., S. Hohng, and T. Ha. 2008. A practical guide to single-molecule FRET. *Nat Methods* 6:507–516.

Schnitzer, M. J., and S. M. Block. 1995. Statistical kinetics of processive enzymes. *Cold Spring Harb Symp Quant Biol* 60:793–802.

Stark, P. R. H., A. E. Halleck, and D. N. Larson. 2007. Breaking the diffraction barrier outside of the optical near-field with bright, collimated light from nanometric apertures. *Proc Natl Acad Sci USA* 104:18902–18906.

Stubbe, J., D. G. Nocera, C. S. Yee, and M. C. Y. Chang. 2003. Radical Initiation in the class I ribonucleotide reductase: long-range proton-coupled electron transfer? *Chem Rev* 103:2167–2201.

Tang, J., A. M. Jofre, G. M. Lowman, R. B. Kishore, J. E. Reiner, K. Helmerson, L. S. Goldner, and M. E. Greene. 2008. Green fluorescent protein in inertially injected aqueous nanodroplets. *Langmuir* 24:4975–4978.

Tanner, N. A., S. M. Hamdan, S. Jergic, P. M. Schaeffer, N. E. Dixon, and A. M. van Oijen. 2008. Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat Struct Mol Biol* 15(2):170–176.

Toprak, E., and P. R. Selvin. 2007. New fluorescent tools for watching nanometer-scale conformational changes of single molecules. *Annu Rev Biophys Biomol Struct* 36:349–369.

van Oijen, A. M., P. C. Blainey, D. J. Crampton, C. C. Richardson, T. Ellenberger, and X. S. Xie. 2003. Single-molecule kinetics of lambda. Exonuclease reveal base dependence and dynamic disorder. *Science* 301(5637):1235–1239.

Walter, N. G., J. M. Burke, and D. P. Millar. 1999. Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction. *Nat Struct Biol* 6:544–549.

Walter, N. G., K. J. Hampel, K. M. Brown, and J. M. Burke. 1998. Tertiary structure formation in the hairpin ribozyme monitored by fluorescence energy transfer. *EMBO J* 17:2378–2391.

Xie, S. 2001. Single-molecule approach to enzymology. *Single Molecules* 2(4):229–236.

Xie, X. S. 2002. Single-molecule approach to dispersed kinetics and dynamic disorder: probing conformational fluctuation and enzymatic dynamics. *J Chem Phys* 117(24):11024–11032.

Yang, H., G. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie. 2003. Protein conformational dynamics probed by single-molecule electron transfer. *Science* 302(5643):262–266.

Yang, W., and S. M. Musser. 2006. Nuclear import time and transport efficiency depend on importin b concentration. *J Cell Biol* 174:951–961.

Zhuang, X., H. Kim, M. J. B. Pereira, H. P. Babcock, N. G. Walter, and S. Chu. 2002. Correlating structural dynamics and function in single ribozyme molecules. *Science* 296(5572):1473–1476.

7

Single-Molecule Studies of Rotary Molecular Motors

Teuta Pilizota, Yoshiyuki Sowa, and Richard M. Berry

Abstract Rotary molecular motors are protein complexes that transform chemical or electrochemical energy into mechanical work. There are five known rotary molecular motors in nature; the bacterial flagellar motor, and two motors in each of ATP-synthase and V-ATPase. Rotation of the flagellar motor drives a helical propeller that powers bacterial swimming. The function of the other rotary motors is to couple electrochemical ion gradients to synthesis or hydrolysis of ATP, and rotation is a detail of the coupling mechanism rather than the ultimate purpose of the motors. Much has been learned about the mechanism of the F_1 part of ATP-synthase and the flagellar motor by measuring the rotation of single motors with a variety of techniques under a wide range of conditions. This chapter will review the structures of ATP-synthase and the flagellar motor, and what has been learned about their mechanisms using single molecule techniques.

7.1. Introduction

Rotarymolecular motors are protein complexes that transform chemical or electrochemical energy into mechanical work. The work done is defined as the product of torque (force times the perpendicular distance to the axis of rotation) and the relative angle of rotation between a fixed part designated the stator and a moving part designated the rotor. There are five known rotary molecular motors in nature (we will not consider here DNA-processing motors with a rotary component to their motion). The bacterial flagellar motor (BFM), the F_o part of ATP-synthase, and the V_o part of ion-pumping V-ATPases are driven by electrochemical gradients (ion-motive forces [IMFs]) of H^+ ions (protonmotive force [PMF]) or Na^+ ions (sodium-motive force [SMF]) across energized membranes. The F_1 and V_1 parts of ATP-synthase and V-ATPases are driven by ATP hydrolysis. In the BFM, the stator is a ring of independent torque-generating units anchored to the cell wall, and the rotor is a set of rings $\sim$45 nm in diameter that spans the bacterial envelope. In ATP-synthase and V-ATPases, the stators of the F/V_o and F/V_1 parts are connected together by a peripheral stalk, and the rotors are connected together by a central stalk. The dimensions

T. Pillizota, Y. Sowa, and R. M. Berry • Department of Physics, Clarendon Lab, University of Oxford, Oxford OX1 3PU, UK

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_7,

of both enzymes are roughly 10 nm in diameter by 20 nm long. Whereas rotation of the BFM drives a helical propeller that powers bacterial swimming, the function of the other rotary motors is to couple the IMF to synthesis or hydrolysis of adenosine triphosphate (ATP), and rotation is a detail of the coupling mechanism rather than the ultimate purpose of the motors. Much has been learned about the mechanism of F_1 and the BFM by measuring the rotation of single motors with a variety of techniques under a wide range of conditions. Similar experiments on F_o, V_1, and V_o may provide further insights in the near future. This chapter will review the structures of F_1, F_o, and the BFM and what has been learned about their mechanisms using single-molecule techniques. Little will be said about V-ATPase, which is very similar to ATP-synthase and almost certainly shares the same rotary mechanisms.

ATP-synthase, also known as F_1F_o-ATPase, is a ubiquitous enzyme located in mitochondria and chloroplasts of eukaryotic cells and in the cytoplasmic membranes of bacteria. It synthesizes ATP from adenosine diphosphate (ADP) and inorganic phosphate (P_i), using the IMF as the source of free energy. ATP is the "energy currency" of living organisms. Its hydrolysis to ADP + P_i releases the energy needed for many cellular processes, and it is present in relatively large amounts inside cells; ATP-synthase synthesizes roughly 2 million kg of ATP in the 75-year lifespan of a typical 70-kg human (Yoshida et al. 2001). Despite the fact that different organisms generate IMFs by many different respiratory or photosynthetic mechanisms, ATP-synthase is one of the most evolutionarily conserved enzymes (Kanazawa et al. 1981, Walker et al. 1985, Hudson et al. 1987). Under normal conditions F_o generates more torque than F_1, and thus F_1 is driven backwards to synthesize rather than hydrolyze ATP. This relationship is reversed in V-ATPase, in which rotation of V_1 drives V_o backwards to pump ions against the IMF. V-ATPase is located in many intracellular compartments (endosomes, lysosomes, secretory vesicles) and has also been identified in the plasma membrane of certain cells (Nishi and Forgac 2002, Beyenbach and Wieczorek 2006). Given that it regulates pH, a carefully controlled parameter affecting many cellular processes, V-ATPase has an important role in, for example, receptor-mediated endocytosis, protein processing and degradation, acid secretion, and bone degradation (Nishi and Forgac 2002, Beyenbach and Wieczorek 2006).

At ~50 nm in diameter and with a molecular mass of ~11 MD, the BFM is one of the largest molecular machines in bacteria. It consists of about 13 different component proteins, and a further approximately 25 proteins are required for its expression and assembly, with a total of 40–50 genes involved in expression, assembly, and control (Macnab 2003). It is a rotary machine embedded in the cell envelope and connected to an extracellular helical filament. As with F_o and V_o, the motor is powered by the flow of ions down an electrochemical gradient across the cytoplasmic membrane into the cell. The rotating filament enables the cell to swim away from unfavourable and towards favourable external conditions. Bacteria achieve active control of their motion by regulating motor rotation, the final stage of processes such as chemotaxis, phototaxis, magnetotaxis, and thermotaxis by which many species of bacteria sense and navigate their environment (Blair 1995, Falke et al. 1997, Wadhams and Armitage 2004, Baker et al. 2006, Paster and Ryu 2008). The best-studied flagellar motors are those of peritrichiously flagellated (several flagella at random locations around the cell) enteric bacteria *Escherichia coli* and *Salmonella enterica Sv typhimurium*, and results described in this review refer to one or other of these motors unless stated otherwise.

7.2. Structure

7.2.1. ATP-Synthase

ATP-synthase has a very similar overall structure whatever the source—prokaryotes, chloroplasts, or mitochondria. The overall structure of F_1F_o-ATPase has been obtained from low-resolution cryo–electron microscopy (cryo-EM) studies of yeast mitochondrial F_1F_o (Stock et al. 1999) and *E. coli* F_1F_o (Wilkens et al. 2000). Numerous atomic structures of the cytoplasmic F_1 part from various species have been obtained by X-ray crystallography, as well as a small number of structures of the membrane-bound rotor of F_o and of various other components of the enzyme. These atomic structures have been docked into a low-resolution structure of bovine heart mitochondrial ATP-synthase (Rubinstein et al. 2003), and in combination with a range of biochemical studies (e.g., Schwem and Fillingame 2006) have led to the model shown in Figure 7.1, where the main figure shows the nomenclature of the bacterial enzyme and the inset shows the equivalent mitochondrial subunits. Unless otherwise stated, we use the bacterial nomenclature in this chapter. Synthesis and hydrolysis of ATP occur in F_1, in which the minimal rotary motor consists of the stator $\alpha_3\beta_3$ and the rotor γ.

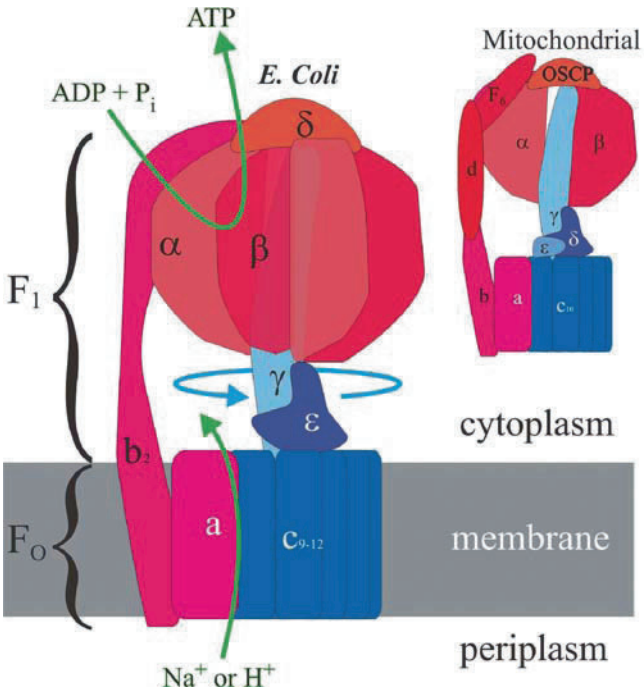


Figure 7.1. Schematic showing the structure and subunit nomenclature of *Escherichia coli* and mitochondrial **(inset)** F_1F_o-ATPase. The ε- and δ-subunits in the mitochondrial enzyme are equivalent to the ε-subunit of the *E. coli* enzyme. Rotor subunits are shown in shades of blue, stator subunits in shades of red. One α– and one β-subunit are removed in the inset to show the γ-subunit more clearly. Ion flux across the membrane in F_o drives rotation of the common rotor (blue arrow), which in turn drives synthesis of adenosine triphosphate (ATP) in F_1. ADP, adenosine diphosphate; OSCP, oligomycin-sensitivity conferring protein; P_i, inorganic phosphate

The rotor of F_o is a ring of 10–15 *c*-subunits, depending on the species (Figure 7.2), and the stator is composed of the *a*- and *b*-subunits. The peripheral stalk containing δ- and b-subunits connects the two stators, and ε helps to link the two rotors together.

F_1

Whereas the majority of single-molecule experiments on F_1 have used prokaryotic F_1, in particular from the thermophilic *Bacillus* PS3 (TF_1), most atomic structures are of the enzyme from bovine heart mitochondria. The first of these was obtained in 1994 by Walker, Leslie, and coworkers with a resolution of 2.8 Å and a crystal grown in the presence of 250 μM adenyl-5-yl imidodiphosphate (AMP-PNP; an analog of ATP), 250 μM ADP, and azide (protein data bank [PDB] code 1BMF) (Abrahams et al. 1994). The structure consists of three α-subunits, three β-subunits, and a γ-subunit. Nucleotide-binding sites were identified at the interfaces between α- and β-subunits. The three noncatalytic sites, formed mainly by the α-subunits, have AMP-PNP and Mg^{2+} bound. Of the three catalytic sites, formed mainly by β-subunits, one has ADP and Mg^{2+} bound (referred to as β_{DP}), one AMP-PNP and Mg^{2+} (β_{TP}), and one subunit is empty (β_E). The 1994 structure most likely represents the Mg-ADP inhibited form, that is, the state immediately after ATP binds the β_{TP} site, stalled by the failure of Mg-ADP release from the β_{DP} site (Hirono-Hara et al. 2001, Yasuda et al. 2003).

Several additional crystal structures were obtained subsequently, of which only a few differ significantly from the 1994 structure in terms of γ-subunit orientation or the number and location of nucleotides. One of them, bovine mitochondrial F_1 inhibited by dicyclohexylcarbodiimide (DCCD), obtained in 2000 (PDB code 1E79), has mitochondrial ε- and δ-subunits as well as the γ-subunit, forming a foot that interacts with F_o (Gibbons et al. 2000). In another bovine F_1 structure, obtained in 2001, the γ-subunit orientation is shifted by 15° compared to the 1994 structure (PDB code 1H8E) (Menz et al. 2001). This crystal was formed from aluminum fluoride–inhibited F_1 and has ADP, Mg^{2+}, and glycerol bound to all α-subunits; ADP, Mg, and SO_4^{2-} to the β_E-subunit; ADP, Mg^{2+}, and AlF_4^- to β_{DP}; and ADP, Mg^{2+}, AlF_4^-, and glycerol to the β_{TP}-subunit. It is speculated that the structure represents the post-hydrolysis, pre-product-release step, in that the β_E-subunit now binds ADP and sulfate (mimicking P_i) and adopts a "half-closed" conformation. Recently a crystal structure of bovine mitochondrial F_1 was obtained in conditions identical to those in the 1994 structure but in the absence of azide (PDB code 2CK3) (Bowler et al. 2007). The two structures differ only in the β_{DP} site. Whereas in the 1994 structure this site is occupied by ADP and azide, the 2007 structure has AMP-PNP bound. The authors take this as further evidence that the 1994 structure represents the Mg-ADP inhibited form and that the 2007 structure represents the most accurate available intermediate state in the catalytic cycle of hydrolysing F_1. In 1997, the nucleotide-free structure of the $\alpha_3\beta_3$ hexamer of TF_1 was obtained with a resolution of 3.20 Å, showing that the α- and β-subunits are essentially the same as in the bovine crystals (Shirakihara et al. 1997).

F_o

There are no high-resolution structures of the F_o stator *a*-subunit. An early X-ray structure of yeast mitochondrial ATP-synthase including F_1 and the F_o rotor showed a ten-fold C-ring attached to the base of F_1 (Figure 7.2, top) (Stock et al. 1999), but without atomic resolution. In 2005 a high-resolution crystal structure of the F_o rotor, the C-ring from Na^+-ATPase from *Ilyobacter tartaricus*, was determined (Figure 7.2, bottom; PDB code

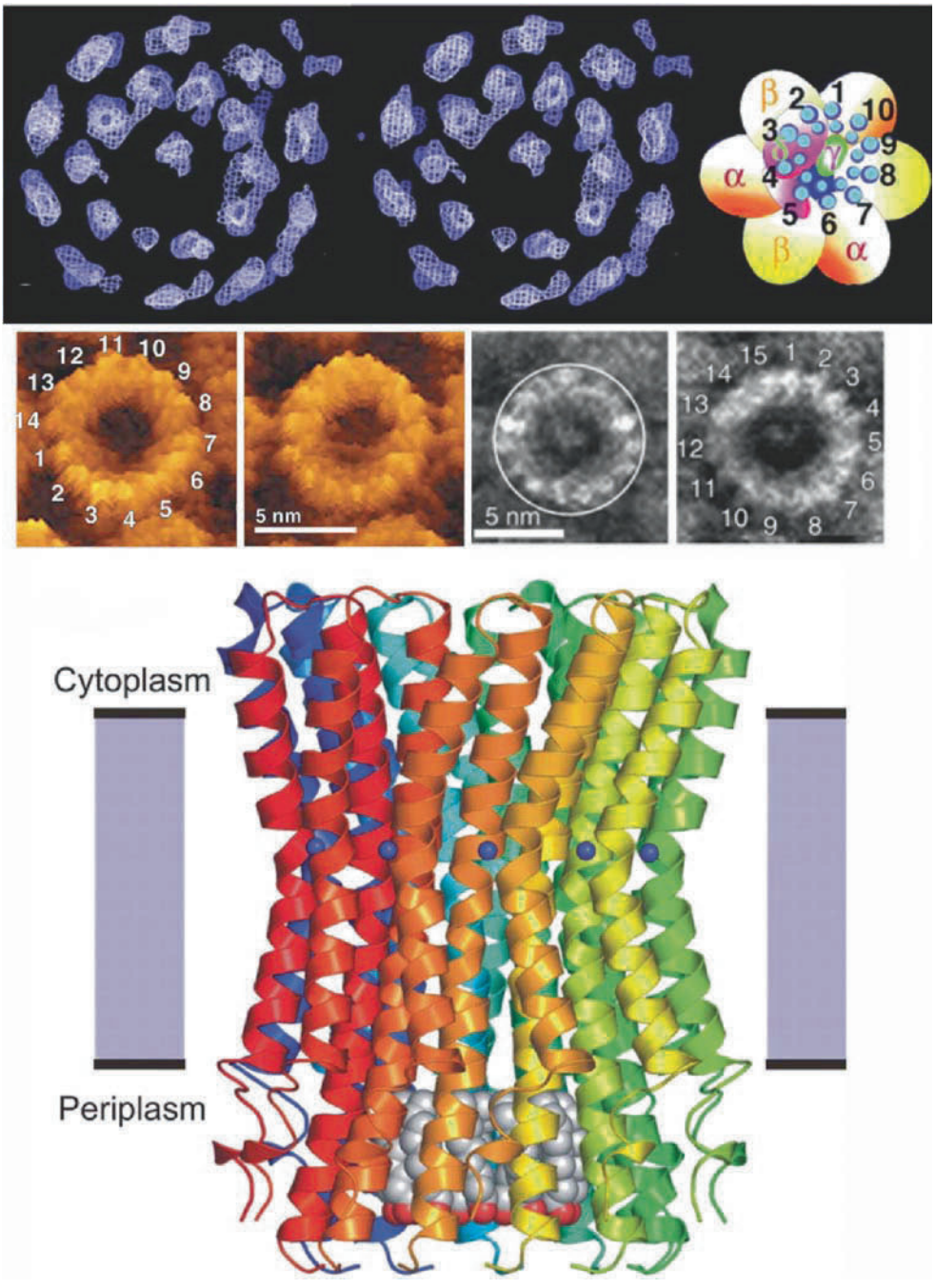


Figure 7.2. F_O rotors from various species. **Top:** Stereo views (*left*) from an electron density map of the yeast F_1c_{10} complex, and the location of the F_1 subunits in relation to the *c*-subunits (*right*). The membrane-spanning α-helices of the *c*-subunit are shown as blue circles. **Middle:** Atomic force microscope images of rotors from chloroplasts of different species, showing 11- and 15-fold symmetries. **Bottom:** X-ray crystal structure of the rotor from *I. tartaricus* showing 11 *c*-subunits. Blue spheres represent the bound Na^+ ions; detergent molecules inside the ring are shown in red and gray. [From Stock et al. (1999). top; Seelert et al. (2000), middle left; Pogoryelov et al. (2005), middle right; Meier et al. (2005), bottom; with permission.]

1YCE) (Meier et al. 2005). Each of the 11 symmetrical *c*-subunits showed a cylindrical, hourglass shape and had a Na^+ ion bound in the middle at a position predicted to be near the middle of the hydrophobic core of the membrane. A similar structure has been obtained for the rotor of a Na^+-driven V-ATPase (Murata et al. 2005), showing a tenfold symmetry, in contrast to earlier assumptions of a sixfold V_o rotor (the assumption was based on different experimental evidence—for example, chemical analysis and electron microscopy studies at lower resolution (Powell et al. 2000, Holzenburg et al. 1993, Arai et al. 1988)). The Na^+-binding sites in these structures both include a single acidic residue that could bind H^+ in a PMF-driven enzyme, indicating that H^+ rotor structures may be very similar to those of the Na^+ rotors.

One of the more interesting features of the F_o rotor is the variability of its symmetry among different species. The C-rings from yeast mitochondria, *Ilyobacter tartaricus*, *Propionigenium modestum*, *Synechocystis*, cyanobacterium, *Spirulina platensis*, and spinach chloroplasts reconstituted into two-dimensional crystals and observed with atomic force microscopy (AFM) or, in the case of *P. modestum*, observed with cryo–transmission electron microscopy (cryo-TEM), have symmetries of 10, 11, 11, 13, 14, and 15, respectively (Stock et al. 1999, Stahlberg et al. 2001, Meier et al. 2003, Pogoryelov et al. 2007, Seelert et al. 2000, Pogoryelov et al. 2005) (Figure 7.2, middle). It was suggested that the variable number of *c*-subunits within a certain species could be a regulatory mechanism (Schemidt et al. 1998), but experimental data for spinach chloroplast and *Acetobacterium woodii* ATP synthase show that the stoichiometry seems to be fixed for a given species (Fritz et al. 2008, Meyer Zu Tittingdorf et al. 2004). Given that a 360° rotation of the F_1 γ-subunit generates three ATP molecules and presumably requires each *c*-subunit to translocate one ion, it follows that the number of ions required to make one ATP molecule is not necessarily an integer. Symmetry mismatch may also be necessary for an elastic power transmission between F_1 and F_o as suggested by some (Junge et al. 2001), and it is a feature observed in the BFM as well (Thomas et al. 2006).

7.2.2. Bacterial Flagellar Motor

The bacterial flagellum spans the cell envelope, extending into the cytoplasm and the cell exterior. It consists of a hook and a flagellar filament, peptidoglycan P- and outer-membrane L-ring, rod, MS-ring, C-ring, stator units, and an export apparatus (Berg 2003, Macnab 2003). Atomic structures are available for the filament and hook, as well as for isolated protein fragments from the rotor. The overall arrangement of the motor has been derived from cryo-EM reconstructions and biochemical data and is shown in Figure 7.3. Whole rings have proved too large for X-ray crystallography. A recent report of crystallization of the membrane-spanning Mot B protein offers the hope of revealing the atomic structures of parts of the stator in the near future (O'Neill and Roujeinikova 2008).

The periplasmic P-ring and outer lipopolysaccharide membrane L-ring are composed of FlgI and FlgH proteins and transmit rotation from the motor through the cell envelope. It is not known whether they form part of the rotor or are fixed to the cell wall. The MS-ring is the first part of the motor to assemble and can be thought of as the platform on which the rest of the motor is built (Aizawa 1996, Macnab 2003). The name comes from the original idea that the ring is composed of two rings (membrane and supramembranous), but in 1992 it was shown and later confirmed that the ring consists of a single protein, FliF (Ueno et al. 1992, 1994, Suzuki et al. 2004). FliG, FliM, and FliN comprise the C-ring, which is thought to be the site of torque generation. These three proteins are often referred to as the "switch complex" because mutations lead to defects in switching (Yamaguchi et al. 1986a,b). The

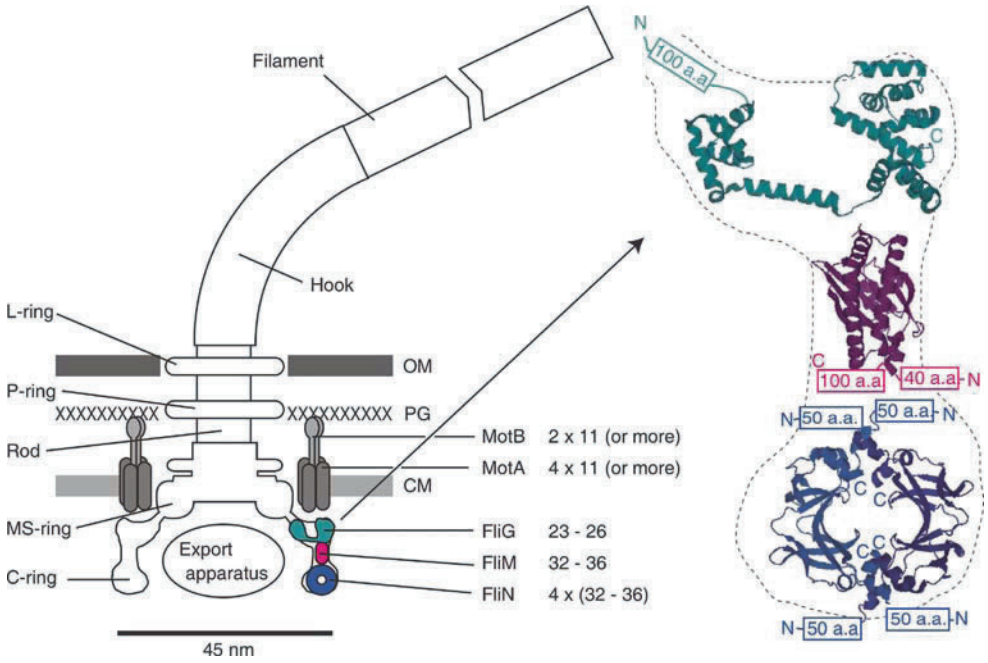


Figure 7.3. Left: Schematic side view of H^+-driven flagellar motor, with the proposed location and copy number of proteins involved in torque generation. MotA and MotB are thought to form stator complexes with stoichiometry A_4B_2, and FliN, a tetramer that has 1:1 stoichiometry with FliM. The motor spans the three layers of the of cell envelope: outer membrane (OM), peptidoglycan cell wall (PG), and cytoplasmic membrane (CM). **Right:** Detail of proposed location and orientation of rotor proteins. X-ray crystal structures of truncated rotor proteins, FliG (*cyan*; Protein Data Base [PDB] code 1LKV), FliM (*magenta*; PDB code 2HP7), and FliN (Blue; PDB code 1YAB), are shown docked into the rotor structure. The N- and C-termini and missing amino acids are indicated. a.a., amino acids. [Molecular graphics generated using PyMol (http://www.pymol.org).]

switch complex also serves as the housing for the export apparatus of the bacterial flagella. In 1999 a partial structure of *Thermotoga maritima* FliG was obtained (PDB code 1QC7) corresponding to the C-terminal domain of the protein, which functions specifically in torque generation (Lloyd et al. 1999, 1996, Irikura et al. 1993). This domain contains a set of charged residues that are essential for motor rotation and interaction with the stator (Lloyd and Blair 1997, Zhou et al. 1998a), and the structure showed these residues clustered along a ridge that is proposed to contact the stator and to change orientation to allow the motor to switch direction. The N-terminal and middle domains of FliG appear to be important for flagellar assembly and switching. In 2002 the atomic structure of a larger fragment of FliG from *T. maritima* was obtained containing the C-terminal and middle domains and an α-helix with an extended segment linking them together (Figure 7.3; PDB code 1LKV) (Brown et al. 2002). Mutational analysis from the same report indicated that FliM binds to both of these domains. An X-ray crystal structure of the middle part of *T. maritima* FliM, the protein that binds the switch-inducing signal protein CheY, was obtained in 2006 (PDB code 2HP7) (Park et al. 2006). Cross-linking studies based on this structure showed that FliM self-associates in a side-to-side arrangement, which would allow cooperative transitions during a motor switch (Park et al. 2006). Crystal structures of a fragment of FliN from *T. maritima* lacking 50 amino acids at the N-terminus were reported in 2003 (PDB code 1O6A) and 2005 (PDB code

1YAB) (Brown et al. 2005). FliN forms saddle-shaped dimers in both structures, and association studies from the same reports predict FliN tetramers and stable $FliN_4$–$FliM_1$ complexes. In the latter structure, pairs of dimers form a doughnut-shaped tetramer that has been fitted to the cryo-EM reconstruction of the bottom of the C-ring (Thomas et al. 2006). The FliG, FliM, and FliN structures from *T. maritima* provide a good model for the same proteins in *E. coli* motor due to high sequence homology overall and in particular in the segments most important for function.

Axially resolved cryo-EM of the rotor has revealed two different symmetries in the MS- and C-rings. The MS-ring and the inner part of the C-ring have 23- to 26-fold symmetry, whereas the bottom and outer parts of the C-ring are 32- to 36-fold (Thomas et al. 2006). The copy numbers of FliM and FliN have been estimated as 34 ± 6 and 111 ± 13, respectively (Thomas et al. 1999, Zhao et al. 1996), consistent with a $\sim$34 copies of an M_1N_4 unit. FliF–FliG fusion proteins produce functional motors (Kihara et al. 2000), suggesting that FliG shares the symmetry of the MS- rather than that of the C-ring. In one model, FliG is restricted to the inner part of the C-ring, with $\sim$26-fold symmetry (Thomas et al. 2006). In an alternative model (Brown et al. 2007), FliG spans the inner and outer parts of the C-ring, and the outer part has *n* defects, where *n* is the difference between the two symmetries and each defect is a missing FliG. A similar symmetry mismatch was proposed earlier in a model in which *n* was identified with eight stator units and each unit generated torque by propagating the defect along the ring (Thomas et al. 1999). The model predicts that the MS- and C-rings rotate at different speeds, an outlandish proposal that might be testable in future using fluorescent-labelled rotor proteins.

MotA and MotB are membrane proteins that form the stator of the H^+-driven motor, and PomA and PomB are the Na^+-driven equivalents, with additional proteins MotX and MotY required in wild-type Na^+ motors. Mot (or Pom) proteins can be inserted into the assembled structure by inducing the *mot* genes. Based on Na^+-driven–motor studies, PomA and PomB proteins form a complex of four PomA and two PomB (the same is likely for MotA and MotB) proteins (Sato and Homma 2000, Yorimitsu et al. 2004). Two thirds of MotA is in the cytoplasm (Blair and Berg 1991, Zhou et al. 1995), and MotB has one membrane-spanning α-helix and the rest of the molecule is in the periplasm (Chun and Parkinson 1988). MotB has a peptidoglycan-binding domain near the C-terminus, which anchors the stator unit to the cell wall (de Mot and Vanderleyden 1994). Mutational studies identified Asp32 in MotB (of *E. coli*) as an essential proton-accepting residue that is presumed to form part of the conduction pathway in the motor (Zhou et al. 1998b).

A recent complete EM structure *in situ* of the flagellar motor of the spirochete *Treponema primitia* at 7-nm resolution (Murphy et al. 2006) shows a motor of different shape and size from that previously reported in *S. typhimurium* (Thomas et al. 2006). Sixteen stator units are visible, each connected to the P-collar, and in three places to the rotor. Only one of these latter connections corresponds to the part of the C-ring where FliG is thought to be located in the rotor of *S. typhimurium* (Thomas et al. 2006).

7.3. Single-Molecule Methods for Measuring Rotation

The earliest measurement of the motion of a single molecular motor was the tethered bacterial cell experiment of Silverman and Simon (1974), in which a single BFM was tethered to a microscope coverslip by the filament and the whole cell was used as a visible marker of rotation. Since then, a range of markers has been used to measure rotation of the BFM, F_1,

and F_1F_o: polystyrene beads, fluorescent actin filaments, magnetic beads, gold nanoparticles, and single fluorescent molecules. In single-molecule experiments on rotary motors, motor torque M is balanced by viscous drag on the marker, $M = f\omega$, where f is the frictional drag coefficient of the marker and ω is its angular velocity (the Reynolds number for a spinning motor is much less than 1, and therefore inertial effects can be neglected). External torque has also been applied to tethered cells using electrorotation, to magnetic beads using magnetic fields, and to polystyrene beads and tethered cells using optical tweezers. These methods and their applications will be briefly summarized here.

7.3.1. ATP-Synthase

Although indirect evidence of rotation in F_1 had been obtained before (Rao and Senior 1987, Boyer 1993, Duncan et al. 1995, Sabbert et al. 1996, Abrahams et al. 1994), direct observation of rotation came in 1997 (Noji et al. 1997). A genetically engineered TF_1 with His tags on β-subunits and a Cys mutation on the γ-subunit was attached to a coverslip surface via nickel nitrilotriacetic acid (NTA), and a fluorescently labelled actin filament was attached to the γ-subunit through a biotin-streptavidin link (Figure 7.4A). Direct rotation was observed under an epifluorescence microscope. The same method was used for the observation of three ATP-dependent steps in motor rotation at low ATP concentration [ATP] (Yasuda et al. 1998). Greater temporal resolution was achieved with a much smaller marker of rotation, a 40- to 300-nm gold bead or bead pair attached to TF_1 (Figure 7.4B) observed with a high-speed video camera at 8,000 frames per second and laser dark-field (DF) microscopy (Yasuda et al. 2001). From these experiments it became clear that with an ∼100-nm gold bead pair or smaller the motor rotation is no longer impeded by the load and reaches a maximum speed in saturating [ATP] at room temperature of ∼130 Hz. By comparison, the motor speed with actin filaments attached varied from ∼0.1 to 7 Hz. The gold bead experiment allowed observation of ∼2-ms substeps that are [ATP] independent and therefore could not easily be slowed down by using low [ATP]. In a later experiment the angular resolution was increased by using 200-nm bead pairs instead of single beads (Shimabukuro et al. 2003). Various single-molecule fluorescence methods have also been used for low-resolution studies of TF_1 rotation—for example, polarized fluorescence microscopy of single fluorophores attached to the γ-subunit (Figure 7.4C) (Adachi et al. 2000) and fluorescence resonance energy transfer (FRET) between a donor attached to one of the β-subunits and an acceptor on a γ-subunit (Figure 7.4D) (Yasuda et al. 2003). Binding of a fluorescent ATP analog to F_1 has also been observed simultaneously with rotation of a bead pair attached to the γ-subunit (Nishizaka et al. 2004).

Rotation of F_1 from *E. coli* (EF_1) has been observed using fluorescent actin filaments (Omote et al. 1999, Iko et al. 2001), gold beads (Nakanishi-Matsui et al. 2006, 2007), and back-focal-plane interferometry of polystyrene bead pairs (Pilizota et al. 2007). The maximum observed speed of EF_1 with 60- and 40-nm gold beads attached to the γ-subunit is 400–500 Hz (Nakanishi-Matsui et al. 2006). The relatively low speed of TF_1 at room temperature reflects the high temperature of its native conditions and makes resolution of the details of stepping rotation considerably easier than in EF_1. EF_1 rotation has also been observed in laser DF microscopy via intensity changes in polarized light scattering by gold nanorods (Spetzler et al. 2006, York et al. 2007).

Synthesis of ATP by externally rotated F_1 was demonstrated in 2004 (Itoh et al. 2004) and again with improved resolution in 2005 (Rondelez et al. 2005). F_1 was attached to the coverslip surface by engineered His tags on β-subunits as before, and the γ-subunit was

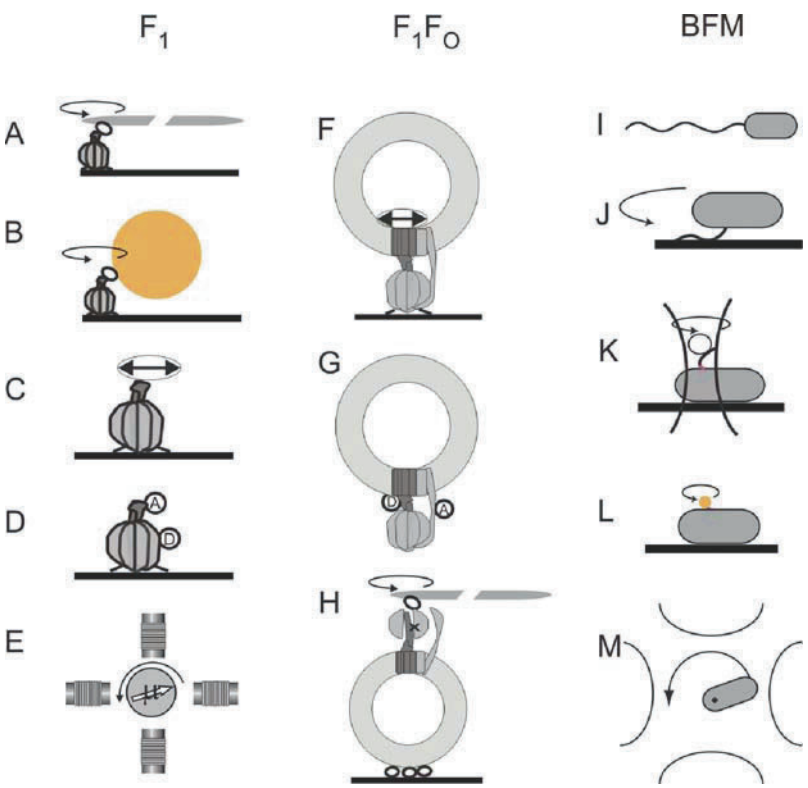


Figure 7.4. Single-molecule assays for rotary motors. The F_1 rotation has been measured by attaching α- or β-subunits to the surface via His-tags and attaching to the γ-subunit, via biotin-avidin, **(A)** fluorescent actin filaments, **(B)** beads, **(C)** single fluorophores, or **(E)** magnetic beads. Fluorescence resonance energy transfer (FRET) has also been used to determine the relative positions of γ- and β-subunits during rotation **(D)**. F_1F_o rotation driven by energized F_o in phospholipid vesicles has been observed via single fluorophores **(F)**, FRET **(G)**, and actin filaments **(H)**. Rotation of the BFM has been measured using **(I)** swimming cells, **(J)** tethered cells, **(K)** polystyrene beads attached to flagella, and **(L)** gold beads attached to hooks. External torque has been applied to F_1 using magnetic fields **(E)** and to the bacterial flagellar motor (BFM) using tethered cells and either electrorotation **(M)** or optical tweezers. More details of these methods can be found in the text.

labelled with a superparamagnetic bead attached via a biotin-streptavidin link (Figure 7.4E). The magnetic bead was rotated using magnetic tweezers, resulting in production of ATP. Typical resolutions reported using magnetic beads are several seconds and 5° (Hirono-Hara et al. 2005). Pairs of polystyrene beads each 500 nm in diameter have also been used as handles to apply external torque to F_1 with an optical trap in an "angle clamp" feedback mode (Pilizota et al. 2007) with a bandwidth up to 1.6 kHz and angular precision of 2°. The drawback of the method is the considerable variation in the trap stiffness with angle and bead pair. Whereas in principle optical tweezers are easier to calibrate due to the good reproducibility of polystyrene beads compared to magnetic beads, the method in fact gave no great improvement in calibration of absolute torque compared to magnetic tweezers.

The rotation of isolated F_o has not been measured directly. There are several reports of F_1F_o rotation measured in a similar manner to the TF_1 and EF_1 experiments (Sambongi et al. 1999, Nishio et al. 2002). Although F_o was present in these experiments, it was not energized, making them effectively F_1 rotation measurements, confirming the co-rotation of F_1 and

F_o rotors. Low-resolution measurements of F_1F_o rotation driven by F_o have been reported. The first observation of F_1F_o rotation during ATP synthesis or hydrolysis came in 2002 (Kaim et al. 2002). Na^+-translocating ATP-synthase of *P. modestum* specially labelled with a single fluorophore on one *c*-subunit was used. The enzyme was immobilized via His tags on β-subunits and reconstituted into proteoliposomes (Figure 7.5F). Rotation was observed with polarization-resolved confocal microscopy during synthesis. In a later experiment F_1F_o from *E. coli* was engineered for FRET, with a donor on the γ- or ε-subunit and an acceptor on the *b*-subunit, and incorporated into liposomes (Figure 7.5G). This allowed the detection of rotation during both synthesis and hydrolysis (Borsch et al. 2002, Diez et al. 2004, Zimmermann et al. 2005). A variant of the F_1 actin experiment has also been reported in which TF_1 was reconstituted with bacteriorhodopsin and *E. coli* F_1F_o in phospholipid vesicles. TF_1 was locked using azide and freed from the F_o stator by removal of the δ-subunit and labelled with an actin filament on the β-subunit (Figure 7.4H) (Zhang et al. 2005). When the vesicles were illuminated, the actin filaments rotated, indicating rotation of F_o driven by the PMF generated by proton pumping in bacteriorhodopsin.

7.3.2. Bacterial Flagellar Motor

True rotation of bacterial flagella, as opposed to propagation of helical waves, was demonstrated in the 1970s (Berg and Anderson 1973, Silverman and Simon 1974). Cells were tethered to a microscope coverslip by filaments (Figure 7.4 J), and rotation of the cell body was observed in a light microscope (Silverman and Simon 1974). A nonswimming strain with straight filaments was used to exclude the possibility that cells were swimming about an inert point of attachment rather than by a rotating motor. To observe the faster rotation of the motor (tethered cells rotate at up to 20 Hz) when driving smaller loads, several techniques have been used.

The rotating filaments of stuck or swimming cells (Figure 7.4I) have been visualized with conventional DF, laser DF, differential interference contrast (DIC), and fluorescence microscopy. Conventional DF and DIC studies have been limited to video rates (Block et al. 1991, Hotani 1976, Macnab 1976). Laser DF has achieved higher time resolution by recording the oscillating light intensity passing through a slit perpendicular to the image of a single filament, which appears as a series of bright spots in this method—one spot for each turn of the filament helix (Kudo et al. 1990, Muramoto et al. 1995). The maximum recorded speed of any molecular motor, 1,700 Hz in the Na^+-driven motor in *Vibrio alginolyticus* at 37ºC, was measured using this technique (Magariyama et al. 1994). Fluorescent labelling of flagellar filaments combined with stroboscopic laser illumination and high-speed video microscopy has revealed polymorphic transitions of filaments (Turner et al. 2000) both in swimming cells (Darnton et al. 2007) and in response to external forces applied with optical tweezers (Darnton and Berg 2007).

The preferred method of measuring fast rotation since 2000 has been to attach polystyrene beads, 200 nm to 2 μm in diameter, to truncated flagellar filaments of immobilized cells. Beads have most commonly been tracked by back-focal-plane interferometry (Figure 7.4 K) (Ryu et al. 2000, Chen and Berg 2000a, b, Sowa et al. 2003, 2005, Reid et al. 2006, Lo et al. 2006, 2007, Inoue et al. 2008). In addition, 200-nm fluorescent beads have been tracked by fluorescence microscopy, and bead tracking has allowed resolution of 26 steps per revolution in very slow flagellar rotation (Sowa et al. 2005). The angular resolution of bead or tethered assays is estimated to be better than 1°, but time resolution is limited to milliseconds or worse by the relaxation time of the flexible hook (Block et al. 1989, Sowa et al. 2005). In

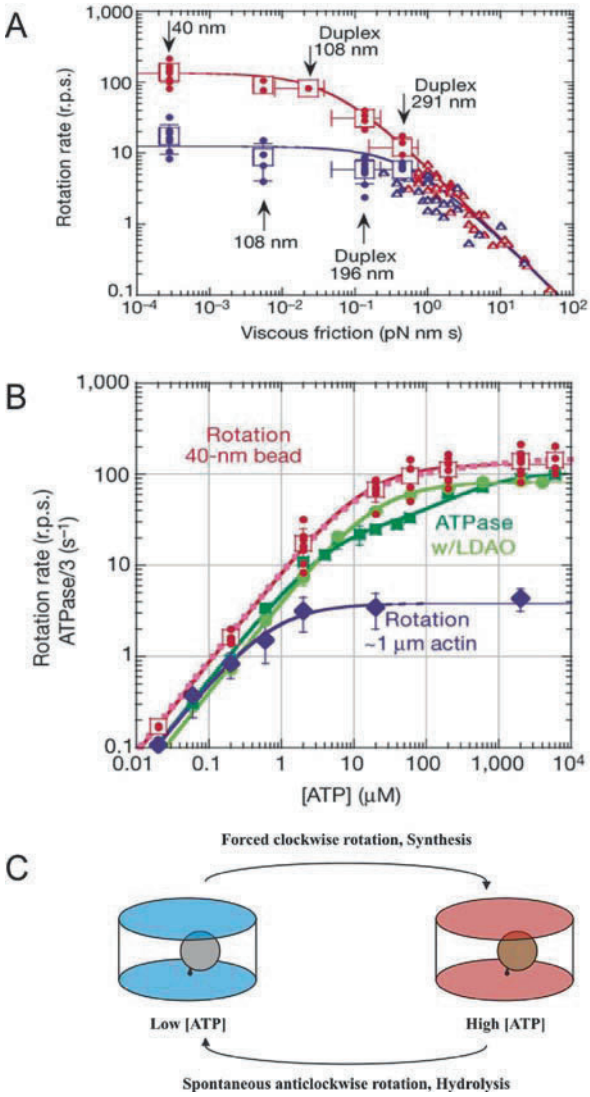


Figure 7.5. **A**. Average speeds of different rotating markers attached to F_1 at saturating concentration of adenosine triphosphate [ATP] (2 mM; *red*) and low [ATP] (2 μM; *blue*) versus rotational frictional drag coefficient. Lines show the expected speeds for a motor producing a constant torque of 40 pN nm rad^{-1}. **B**. Average speeds of 40-nm gold beads (*red*) and actin filaments (*blue*) attached to F_1 versus [ATP]. ATP hydrolysis rates by free F_1, divided by 3 for comparison with rotation rates assuming three ATP molecules hydrolysed per revolution, are shown in green with and without LDAO that prevents occupation of the ADP-inhibited state. Curves show fits with Michaelis-Menten kinetics. **C**. ATP synthesis by single F_1 molecules enclosed in microchambers and driven to rotate by magnetic beads. ATP synthesis was detected as an increase in rotation speed when the magnetic field was removed, due to the increase in [ATP] within the chamber. [Panels A and C from Yasuda et al. (2001), with permission.]

2008, higher time resolution was achieved by using 60-nm gold beads attached to the flagellar motor via the hook in a bacterial strain lacking filaments (Figure 7.4L). Rotation of the bead was observed with laser DF in a manner similar to the F_1 experiments (Yuan and Berg 2008). The viscous drag coefficient of the gold bead is at least two orders of magnitude smaller than any previously measured, making it essentially indistinguishable from zero load. Although the relaxation time of the hook with a 60-nm gold bead attached is estimated to be less than 6 μs, the angular resolution in these experiments was too low to reveal details of rotation within a single revolution.

Early attempts to use magnetic particles to apply external torque to tethered cells were abandoned because they were unable to apply sufficiently large torque (Berg HC, personal communication). Instead, a technique called electrorotation has been used (Washizu et al. 1993, Berg and Turner 1993). Microelectrodes arranged in a cross with a gap of ~70 μm generate a rotating electric field that polarizes a cell located in the gap (Figure 7.4 M). The polarization of the cell falls behind the rotating field, and a torque is generated as the dipole tries to line up with the field. Using substantial voltages, it was possible to spin the cell body in both directions at speeds up to ~1 kHz. Early electrorotation experiments indicated a ratchet-like mechanism in which considerably more torque is needed to force the motor backwards than to stop it rotating forwards (Berg and Turner 1993), but later work showed this to be an artefact of the method (Berry and Berg 1996, 1999, 1997, Berry et al. 1995). Optical tweezers have also been used to exert external torque on the motor with a tethered cell body or beads (Berry and Berg 1997, Pilizota et al. 2007). The motor torque of tethered cells was similar when they were rotated very slowly in either direction by optical tweezers, confirming the electrorotation results (Berry and Berg 1997). The BFM has also been controlled with optical tweezers and pairs of 0.5-μm beads, as described for F_1 in the previous section (Pilizota et al. 2007).

7.4. Energy Transduction

7.4.1. ATP-Synthase

ATP-synthase uses IMF as its free energy input and produces ATP from ADP + P_i as its output. However, single-molecule studies on the whole coupled enzyme are very few, and we will therefore consider the motors F_1 and F_o separately in this section. Considering each as a rotary motor, the inputs are IMF for F_o and ATP hydrolysis for F_1, and the output in both cases is torque. In ATP-synthase it is predicted that the torque generated by F_o should be higher than that generated by F_1, so that F_1 is driven in reverse and ATP is synthesized. In addition, the difference in torques should be small, so that most of the work done by F_o is used to synthesize ATP rather than being wasted as heat. This appears to be true across a range of species, in which the C-ring symmetry and thus the number of ions coupled to each ATP molecule vary proportionally with the ratio of free energies of ATP hydrolysis and ion transit (D. Muller, personal communication). A compliant elastic coupling between F_1 and F_o is believed to be necessary to smooth over the different step sizes of the two motors (see later discussion), transiently storing the energy output of three to five ions passing through F_o for each ATP synthesized.

F_o

F_o uses IMF as the source of free energy needed to drive the rotation of its rotor parts (Mitchell 1961, 1966, Jagendorf and Uribe 1966). The free energy available when an ion such

as H^+ or Na^+ crosses an energized membrane is defined as the product of the IMF and the ionic charge. IMF includes electrical and chemical potentials and is defined as follows:

$$\mathrm{IMF} = \Delta\psi + \frac{k_B T}{q} \ln\left(\frac{C_i}{C_o}\right) \tag{7.1}$$

where $\Delta\psi$ is the difference in electrical potential between the inside and outside of the cell (i.e., the membrane potential), k_B is Boltzmann's constant, T is the absolute temperature, $k_B T$ is the thermal energy, q is the charge of the ion, and C_i and C_o are the respective activities of the ions inside and outside the cell.

Early estimates of the proton conductance of F_o varied over four orders of magnitude due to difficulties in estimating the number of active molecules (Negrin et al. 1980, Friedl and Schairer 1981, Schneider and Altendorf 1982, Sone et al. 1981, Lill et al. 1986, Althoff et al. 1989, Wagner et al. 1989). The difficulty was resolved by preparing small chromatophore vesicles from the photosynthetic bacterium *Rhodobacter capsulatus*, which would on average contain less then one copy of F_o (Feniouk et al. 2001, 2002, 2004). Thus, it was determined that maximum conductance of F_o is 10 fS (at pH 8), equivalent to translocating 6,500 protons sec^{-1} at 100 mV (Feniouk et al. 2004). In the bacterial cytoplasmic membrane $\Delta\psi$ is typically in the range 120–150 mV (Nicholls and Ferguson 2002, Lo et al. 2007, Berg 2003, Felle et al. 1980, Kashket et al. 1980, Shioi et al. 1980), corresponding to (4–6)$k_B T$ per ion, and the total IMF is typically in the range 150–200 mV. For different systems, for example, mitochondria, IMF varies within the same range (Nicholls and Ferguson 2002). Given these values, an upper limit for the torque produced by F_o can be estimated as $M = (Nq \cdot \mathrm{IMF})/2\pi$, where $N \approx 10$ is the number of ions coupled to one revolution, assumed to be the same as the F_o rotor symmetry. Thus we estimate that F_o can generate up to ~50 pN nm rad^{-1} of torque. The only reported measurement of F_o rotation, using an actin filament and reconstituted F_1F_o in anchored phospholipid vesicles (Zhang et al. 2005) (Figure 7.4H), estimated a torque of ~28 pN nm rad^{-1}. However the PMF in this experiment was not known, and it is also possible that the speed of the actin filament, from which the torque was estimated, is slower than the speed of F_o rotation due to rotation of F_o in the membrane or shearing rotational flow within the vesicle.

F_1

The first direct proof of rotation in F_1 coupled to ATP hydrolysis came in 1997, with the rotating actin filament experiment described in Section 7.3.1. Similar results were subsequently obtained using an actin filament to label the ε-subunit of TF_1 (Kato-Yamada et al. 1998), the γ-subunit of EF_1 (Iko et al. 2001), the *c*-, α-, and β-subunits of EF_1F_o (Sambongi et al. 1999, Tanabe et al. 2001), and the *β*- or *a*-subunit of EF_1F_o in a nonenergized membrane fragment (Nishio et al. 2002). These experiments all confirm the rotational mechanism illustrated in Figure 7.1.

The free energy available from the hydrolysis of one ATP molecule is given by

$$\Delta G = \Delta G_0 + k_B T \ln\left(\frac{[\mathrm{ATP}]}{[\mathrm{ADP}] \cdot [P_i]}\right) \tag{7.2}$$

where ΔG_0 is the standard free energy and [X] represents the molar concentration of chemical species *X*. The energy released by one ATP molecule is $\Delta G = 57$ kJ/mol ≈ 95 pN nm (Berg

et al. 2002), setting the upper limit to the torque generated by F_1 as $M_{max} = (3 \cdot \Delta G)/2\pi \approx 45$ pN nm rad^{-1}. As expected, this is a little lower than the predicted maximum torque of F_o. The 1997 experiment (Noji et al. 1997) found that actin filaments a few microns long attached to the γ-subunit of TF_1 rotated at approximately up to ~4 Hz in saturating [ATP]. The average torque generated by each filament was estimated by multiplying this speed by the rotational drag coefficient of the filament, ζ, estimated as (Yasuda et al. 1998)

$$\zeta = \frac{(4\pi/3)\eta L^3}{\ln(L/2r) - 0.447}$$

where η is the viscosity of the medium, r is the radius, and L is the length of the filament. Torques up to ~40 pN nm rad^{-1} were estimated, consistent with the predictions based on the energetics of ATP hydrolysis and IMF. In 2001 the torque of the motor was calculated from the actin filament curvature, based on the finding that the major resistance of the actin filament to the rotation of the motor comes from viscous and other interactions with the surface. The value obtained was ~50 pN nm rad^{-1} (Pänke et al. 2001, Cherepanov and Junge 2001).

These results are therefore consistent with the hypothesis that one ATP is hydrolysed per 120° revolution and the motor operates near ~100% efficiency at low speeds when driving a large viscous load with nanomolar to millimolar [ATP] (Sakaki et al. 2005). In 2001 a much improved F_1 rotation experiment using gold beads instead of actin filaments (Yasuda et al. 2001) mapped the dependence of rotation rate upon viscous load and [ATP] (Figure 7.5A, B). In saturating [ATP] with gold beads 100 nm or less in diameter, speed was independent of both bead size and [ATP], indicating a rate-limiting chemical step linked to ATP hydrolysis. With larger viscous loads speeds were reduced, indicating rate-limiting mechanical rotation of the γ-subunit and attached load under these conditions. For the smallest loads, the speed versus [ATP] relationship followed Michaelis-Menten kinetics with similar V_{max} and K_m as measured for ATP hydrolysis in bulk by unlabelled F_1, demonstrating that the attached gold particle did not affect the kinetics of ATP hydrolysis. Compared to rotation rates, the bulk hydrolysis rate is ~40% too low, presumably due to an inactive fraction in the bulk assay.

Direct evidence of mechanically driven ATP synthesis, with F_1 rotation driven externally by a magnetic bead rather than by F_o, came in 2004 (Itoh et al. 2004) and 2005 (Rondelez et al. 2005). Detecting the small quantities of ATP synthesized proved to be the most difficult part of these experiments due to contaminating ATP in buffers nominally containing only ADP. The key was to confine the experiment to a small volume so that the small number of ATP molecules synthesized constituted a measurable increase in ATP concentration. In 2004, ATP was detected by the luciferase–luciferin reaction in droplets 3 μm high by 30 μm in diameter. The 2005 experiment used a silicone device presenting a large array of cavities 1.5 μm high by 1.1 μm in diameter, and ATP was detected via the concentration-dependent increase in speed of the very same F_1 molecule that was used to synthesize it, after the external magnetic field was switched off (Figure 7.5c). Tight coupling between ATP synthesis and rotation in this experiment was dependent upon the ε-subunit: magnetic beads attached to $\alpha_3\beta_3\gamma\varepsilon$ synthesized an average of 2.3 ± 1.6 ATP molecules per rotation, a coupling efficiency of ~77%, whereas with $\alpha_3\beta_3\gamma$ the coupling efficiency fell to 17%. The torque required to synthesize ATP was not measured.

7.4.2. Bacterial Flagellar Motor

Input

The BFM, like F_o, is powered by the IMF across the bacterial cytoplasmic membrane. The first direct evidence for this was the observation of flagellar rotation in starved *Streptococcus* or *Bacillus subtilis* cells provided with an artificial membrane potential or pH gradient and no ATP (Manson et al. 1977, Matsuura et al. 1977), confirming earlier indications that the motor was ion driven (Larsen et al. 1974). The existence of Na^+-driven motors in alkalophilic *Bacillus* and in *Vibrio* species was demonstrated shortly after. Flagellar motility of these species was sensitive to Na^+ concentration changes and not sensitive to proton ionophores that collapse the PMF (Chernyak et al. 1983, Hirota and Imae 1983, Hirota et al. 1981). Numerous functional chimeric motors that mix components from motors with different driving ions support the conclusion that the mechanisms of Na^+- and H^+-driven motors are very similar (Asai et al. 2003, Yorimitsu et al. 2003). The pattern of ion selectivity of functional chimeric motors shows that no single component of the motor uniquely determines ion selectivity (Sowa and Berry 2008).

Single-molecule experiments in the BFM are made easier than in F_o by the very same factors that make them harder than in F_1. Live bacteria present a motor that is assembled in an energized membrane, anchored to the cell wall, and relatively easy to label via the hook or filament that extends from the cell. Thus, in contrast to F_o, ion-driven rotation in the BFM can be observed simply by sticking live cells to a surface and measuring the rotation of a marker attached to hook or filament. However, unlike with F_1, where ATP, ADP and P_i can be added at whatever concentrations are desired, the energy input to the BFM is linked to the metabolic state of the cell and therefore relatively difficult to measure and control. There have been two approaches to this problem: Externally impose an IMF using voltage clamping, diffusion potentials, and specified concentration gradients in combination with ionophores; or disrupt the natural energization of the cell and attempt to measure the resulting IMF.

The most definitive experiment to impose an IMF on the flagellar motor was performed by Fung and Berg (1995). Filamentous cells were drawn halfway into micropipettes, the cytoplasmic membrane of the part inside the pipette was made permeable by exposure to a proton ionophore, and the external part of the membrane was energised by voltage clamping the pipette. Dead cells were used to mark motors in the external part, imposing a high viscous load equivalent to a tethered cell, and speed was found to be proportional to $\Delta\psi$ up to –150 mV (Figure 7.6A). This result confirmed earlier measurements of the speed of the tethered gram-positive bacteria *Streptococcus* and *Bacillus* energized by a K^+ diffusion potential (Figure 7.6B) (Khan et al. 1985, Manson et al. 1980, Meister and Berg 1987). The independent effects of each term in the IMF, Eq. (7.1), were investigated in tethered *Streptococcus* cells in 1980 using ionophores and uncouplers (Manson et al. 1980). Both $\Delta\psi$ and pH gradient were found to be equivalent in driving the motor (Manson et al. 1977, 1980); the speed depended only on the size of the PMF, not how it was composed. The dependence of motor speed on PMF at low load has been measured only by disrupting natural energization. Gabel and Berg (2003) attached polystyrene beads 0.4 μm in diameter to the free rotating filament of tethered cells and gradually collapsed the PMF by adding low concentrations of the proton ionophore carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) (Figure 7.6C). In each cell the speed of the motor rotating the bead (low load) was found to be proportional to that of the tethered motor (high load). Tethered cell speed is proportional to PMF, and thus it was concluded that

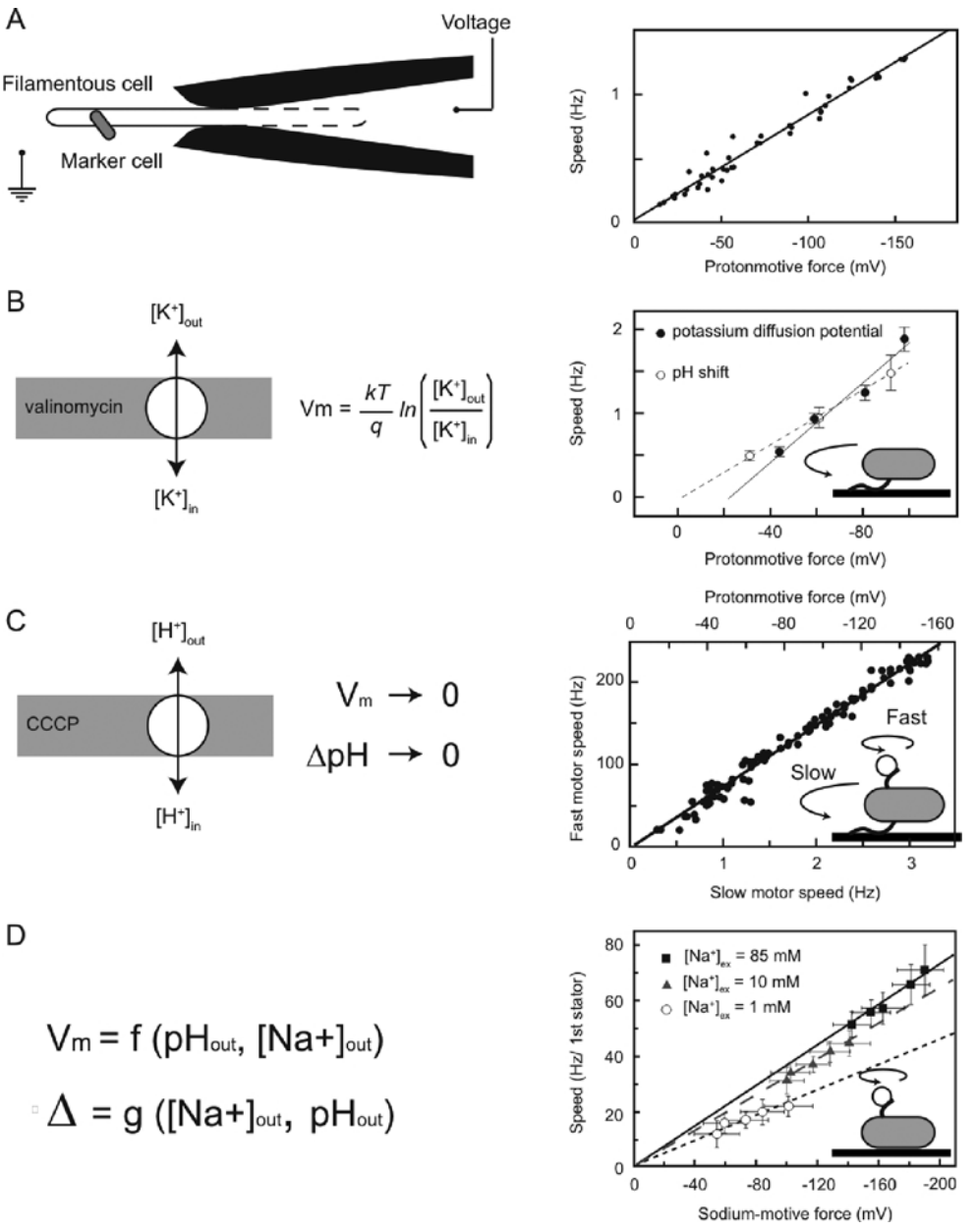


Figure 7.6. Torque versus ion-motive force in the BFM. **A. Left:** Schematic of a voltage clamp method using filamentous *Escherichia coli* cells held in custom-made micropipettes. The part of the membrane inside the pipette (*dashed line*) is made permeable using the ionophore gramicidin S. Motor speed was monitored by videomicroscopy of a dead cell attached to the motor. **Right:** Motor speed is proportional to membrane voltage (= protonmotive force [PMF]) between 0 and –150 mV. **B. Left:** Membrane voltages in *Streptococcus* can be controlled by a K^+ diffusion potential in the presence of valinomycin. **Right:** The speed of tethered *Streptococcus* cells is proportional to the PMF, and membrane voltage is equivalent to the pH gradient under these conditions. **C. Left:** The PMF can be varied slowly from –150 mV down to zero by adding small concentrations of carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) or sodium azide. **Right:** Based on the result of panel A, the speed of a tethered *E. coli* motor (lower axis)

motor speed at low load is also proportional to PMF, although the relative contributions of $\Delta\psi$ and pH gradient were not known.

In 2006 and 2007, the effects of both components of SMF on motors were measured at high (1-μm beads) and low (0.36-μm beads) load (Lo et al. 2006, 2007) in chimeric sodium-driven flagellar motors in *E. coli* (Asai et al. 2003, Sowa et al. 2005, Lo et al. 2006). The $\Delta\psi$ was varied by changing external pH and the [Na^+] gradient (ΔpNa) by changing external [Na^+]. The $\Delta\psi$ and intracellular [Na^+] were measured in single cells using two different fluorescent dyes (Lo et al. 2006, 2007). At high load, motor speed was proportional to SMF, with equivalent contributions from ΔpNa and $\Delta\psi$ as for PMF in previous work. The linear dependence of speed and IMF in the high-load regime, where load rather than electrochemical processes such as the arrival of ions is rate limiting, is consistent with high efficiency and tight coupling. At low load, however, the components of SMF were not equivalent (Lo et al. 2007). At a given external [Na^+], speed was proportional to SMF, but the proportionality constant changes depending on the sodium concentration. Higher [Na^+], corresponding to a larger relative ΔpNa contribution to the SMF, gave a larger constant of proportionality (Figure 7.6D). This result is consistent with an earlier experiment on the sodium-driven *V. alginolyticus* motor (Sowa et al. 2003) and indicates that in the low-load regime ion binding is the rate-limiting step.

The only measurement of the ion flux through the flagellar motor was based on shifts in the rate of pH change of a weakly buffered dense suspension of swimming *Streptococcus* cells when motors were stopped by cross-linking their filaments with antifilament antibody (Meister and Berg 1987). The estimated flux was on average 1,240 $\pm$ 240 protons per revolution per motor over a speed range of $\sim$20–60 Hz (Meister and Berg 1987).

Output

Torque and rotation are the output of the BFM. The recent detection of steps in motor rotation (see section Interactions between Rotor and Stator) promises to offer deep insight into the motor's mechanochemical cycle in the near future. However, at the time of writing, the best way of understanding the motor has been by investigating its torque–speed relationship. Torque measurements have been made at a range of speeds, using several different techniques. The results are summarized in Figure 7.7. Measurements on tethered *Streptococcus* cells gave a linear torque–speed relationship (Manson et al. 1980, Lowe et al. 1987). However, later experiments using polystyrene beads attached to truncated filaments achieved higher temporal resolution with a broader range of applicable loads. Load on the motor was changed by varying the size of the bead and using unlabelled filaments for the lowest load point, or more reliably by using a small bead and varying the viscosity of the medium (Ryu et al. 2000, Chen and Berg 2000a, Sowa et al. 2003, Inoue et al. 2008). The motor torque is approximately

Figure 7.6. (continued) was used as a proxy for PMF (upper axis, absolute value shown). The speed of a second motor on the same cell, attached to a submicron bead, was found to be proportional to PMF. **D. Left:** Both components of the sodium-motive force (SMF) in *E. coli* can be varied using external pH and Na^+ concentration and quantified using fluorescence methods. **Right:** The speed of single-stator chimeric motors driving small loads is proportional to the SMF at a given external Na^+ concentration, but motors spin faster in high Na^+ even at the same SMF. [Data adapted from (A) Fung and Berg (1995), (B) Manson et al. (1980), (C) Gabel and Berg (2003), and (D) Lo et al. (2007).]

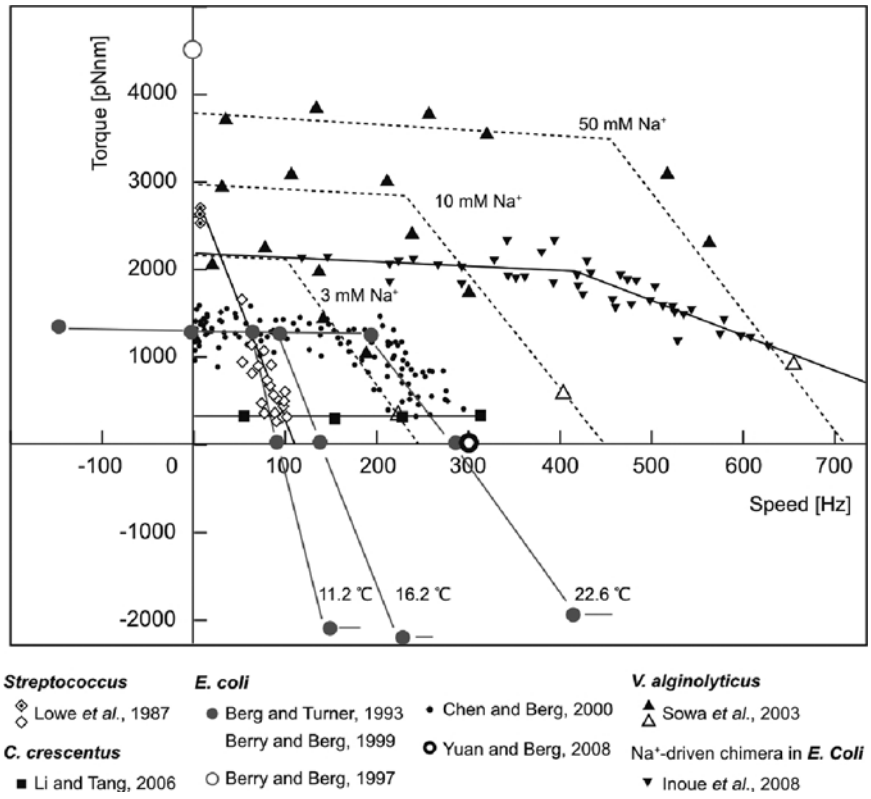


Figure 7.7. Torque–speed relationships for flagellar motors of various species with wild-type numbers of stator units. Upright triangles and dashed lines are data for the Na^+ motor of *Vibrio alginolyticus* at three different external Na^+ concentrations; inverted triangles and solid line are data for the Na^+-driven chimera in *Escherichia coli* at high sodium-motive force. The data for the H^+ motor of *E. coli* are all in agreement, except for the grey open circle at zero speed, which is probably too high due to systematic errors in the calibration of the optical trap. The low-speed data point for *Streptococcus* (*filled diamond*) may also be too high due to a systematic error. Except where indicated, all measurements were made at room temperature. For more details see references indicated in the legend. The *E. coli* experiments using electrorotation (*grey filled circles*) and beads (*black filled circles*) did not report absolute torques—these curves have been scaled to a stall torque of 1,260 pN nm (Reid et al. 2006). *C. crescentus*, *Caulobacter crescentus.*

constant (slowly declining) up to several 100 Hz. At higher speeds the motor torque declines more steeply and approximately linearly. The transition between these two limits has been called the "knee" and is temperature dependent. Torque–speed curves obtained by electrorotation are in agreement with those found by other methods and extend the measurements backwards to ∼100 Hz and forwards beyond the zero-load speed of the motor. The torque generated at high load is independent of temperature (Berg and Turner 1993, Chen and Berg 2000a), hydrogen isotope (Chen and Berg 2000b), and the composition of the IMF (Manson et al. 1980), consistent with tight coupling, constant efficiency, and mechanical rate limitation. By contrast, the torque at low load depends on all three factors (Chen and Berg 2000a, b), indicating rate-limiting kinetic steps such as ion binding and transit through the motor or conformational changes.

The most reliable estimate of the absolute magnitude of torque generated by the *E. coli* motor, 1,260 ± 190 pN nm rad^{-1}, comes from the measurements with polystyrene beads of diameter 1 μm (for which the uncertainty introduced by unknown filament lengths is negligible) (Reid et al. 2006). The torque in the chimeric motor is a little higher. As in F_o, conservation of energy sets a lower limit to the number of ions that must flow in the motor per revolution, $N_{min} = (2\pi M)/(q \cdot \text{IMF})$. For a single-stator chimeric motor at high SMF driving a 1-μm bead, $N_{min} = 36 \pm 6$ (Lo et al. 2007).

7.5. Mechanism

7.5.1. ATP-Synthase

F_1F_o

A substantial effort from several laboratories around the world has gone into investigating rotation powered by F_o on a single-molecule level. Because F_o is membrane bound and requires an energized membrane to work, these experiments are difficult, progress is slow, and experimental data are noisy. All attempts have used F_1F_o reconstituted into phospholipid vesicles. Coupled rotation of F_1F_o has been demonstrated using both polarization microscopy of single fluorophores (Figure 7.4F) and single-pair FRET (Figure 7.4G). The polarization experiment showed rotation of the enzyme during synthesis upon applying a diffusion potential but not with a Na^+ concentration gradient alone, and during hydrolysis but only in the presence of Na^+ (Kaim et al. 2002). Rotation was sensitive to *N*,*N*′- dicyclohexylcarbodiimide (DCCD), a known F_o rotation inhibitor, further confirming that F_o was indeed rotating. The frequency of observed oscillations in the polarization signal, corresponding to rotation, was consistent with expected rates for a given [ATP] based on previously observed F_1 rotation. The first FRET experiment (Figure 7.4G) showed three distinct FRET levels between fluorophores on the γ- and *b*-subunits that were consistent with 120° rotations of γ, and cycling between these levels during ATP hydrolysis (Borsch et al. 2002). Later work by the same group using similar methods showed cycling among the same three FRET levels but in opposite directions during hydrolysis and synthesis (Diez et al. 2004, Zimmermann et al. 2005) (Figure 7.8). The PMF for synthesis consisted of a K^+ diffusion potential plus a pH gradient; FRET was between an acceptor on the *b*-subunit and a donor on either the γ- or the ε-subunit. The later experiments also showed a breaking of the threefold symmetry of F_1, presumably due to interactions with the second stalk (*b*-subunits) and/or the fluorophores.

Stepping in F_1F_o is complicated by the different rotational symmetries of the two component motors. Whereas isolated F_o would be expected to take about ten steps per revolution, one per subunit of the C-ring, coupling of the C-ring to the F_1 rotor in F_1F_o is expected to introduce threefold symmetry to the rotation. Thus the number of steps observed is expected to vary from about ten if relative rotation of *c*- and *a*-subunits is measured to three if relative rotation of the rotor and stator of F_1 is measured. One possible explanation of the threefold stepping observed in the FRET experiments is that an elastic compliance between the F_1 and F_o rotors, or between the *b*- and *a*-subunits in F_o, filters out the F_o steps. Alternatively, even a perfect F_o FRET-label pair would be expected to show three steps at low time resolution if step times in F_o are not equally distributed but instead controlled by interaction with F_1. Thus the experiment would resolve only the dwell before a critical rate-limiting step in which ATP in synthesized, with all other F_o steps happening too fast to resolve. Single-pair FRET

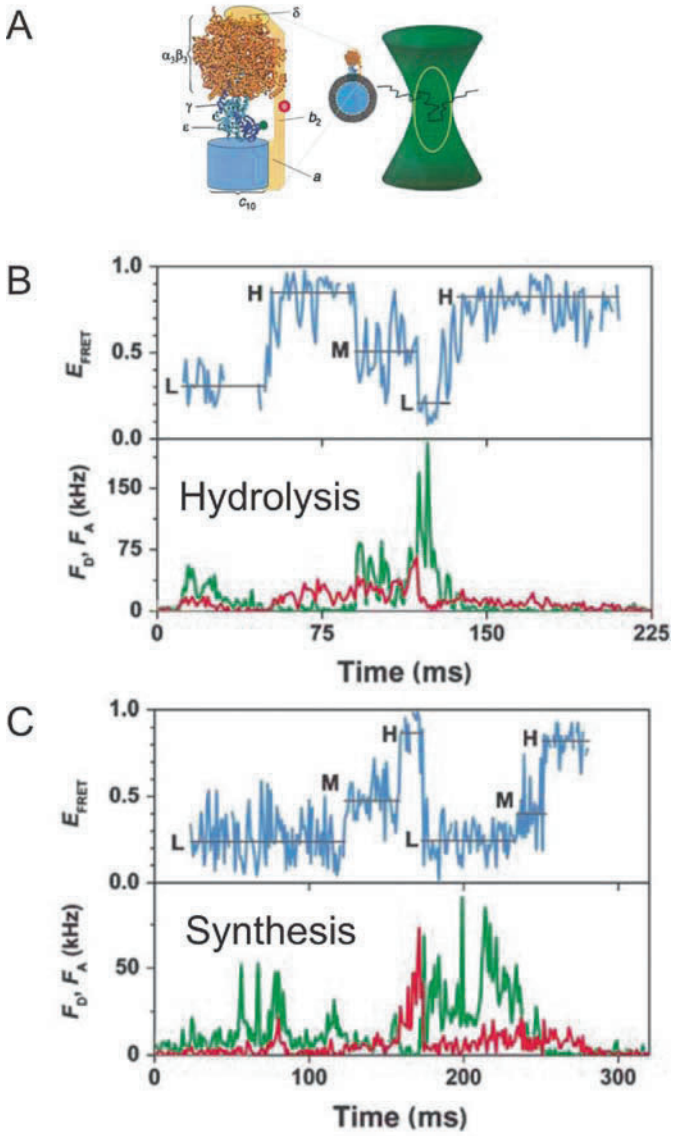


Figure 7.8. Detection of stepping rotation in F_1F_o using fluorescence resonance energy transfer (FRET) between fluorophores attached to the ε- and *b*-subunits. **A.** FRET-labelled F_1F_o is reconstituted into vesicles, which give bursts of photons each time they diffuse freely through a confocal laser excitation spot. The distance between fluorophores, and thus the FRET efficiency (E_{FRET}), cycles through three values, indicating 120° steps. The direction of cycling is opposite during adenosine triphosphate hydrolysis **(B)** and synthesis **(C)**. The lower plots in panels B and C show donor (*green*) and acceptor (*red*) fluorescence intensities. [From Zimmermann et al. (2005), with permission.]

is an ideal technique for probing this question because fluorophore pairs can in principle be engineered to monitor relative motions of different parts of the enzyme.

F_1

Compared to F_o and F_1F_o, single-molecule experiments on F_1 have been numerous and the data quality exceptionally high, allowing a detailed picture of the mechanism of F_1 to be built. The majority of the work has used TF_1, which has the advantages that it is probably more robust to denaturation and other mechanical damage and furthermore that at room temperature it is operating far below its normal temperature, resulting in slow kinetics that are an aid to direct detection with single-molecule methods. Unless otherwise stated, this section refers to TF_1. In 1998 the actin filament rotation assay revealed discrete 120° steps in F_1 rotation (Yasuda et al. 1998) in nanomolar [ATP]. The "transient" torque in discrete stepping traces was defined as $M = (2\pi/3)\zeta/\tau_{\text{step}}$, where τ_{step}, the time required to complete a step, was estimated by aligning many steps and averaging. The transient torque thus obtained was $\sim$40 pN nm rad^{-1}, similar to the average torque generated by F_1 driving a large viscous load. In 2000, stepping rotation was confirmed using angle-resolved single-fluorophore imaging, ruling out the possibility that interactions between the actin filament and F_1, rather than the mechanism of F_1, were responsible for the steps (Adachi et al. 2000).

Since 2001, polystyrene and gold nanoparticles with diameters ranging from 40 to 300 nm have been the label of choice for measuring F_1 rotation because their high light-scattering cross section and small size combine to give the best possible combination of time and angular resolution. The first experiment to use gold labels (Yasuda et al. 2001) allowed a detailed analysis of stepping rotation that gave a great insight into the mechanism of ATP hydrolysis by F_1. At high [ATP] only 120° steps were seen, whereas at low [ATP] substeps of 80° and 40° could be distinguished (Figure 7.9A). The initial report actually estimated 90° and 30° substeps. A later experiment using bead pairs rather than single beads to improve angular resolution gave improved estimates of 80° and 40° (Shimabukuro et al. 2003). To avoid confusion, we will use the later, improved estimates of 80° and 40° throughout this section. The duration of the 120° step at saturating [ATP] was roughly equal to that of the 40° substeps at low concentrations, indicating that the 80° substep becomes fast at high [ATP], but that the 40° substep is ATP independent. Because the 80° substep gets longer as [ATP] decreases, it is natural to assume that it corresponds to a step in which the ATP molecule binds to F_1. An analysis of the distribution of times taken to rotate 120° confirmed this. At low [ATP], ATP binding is rate limiting, and the 80° substep takes far longer than the 40° substep. Under these conditions the distributions were well fitted by a single exponential with a rate directly proportional to [ATP], giving a second-order rate constant $k_{\text{on}} = (3.0 \pm 0.1) \times 10^7$ M^{-1} sec^{-1} for ATP binding. At saturating [ATP] the 40° substep dominates, and distributions fitted a double-exponential with rates $k_1 = (1.64 \pm 0.06)$ msec^{-1} and $k_2 = (0.71 \pm 0.02)$ msec^{-1}, indicating that two sequential processes with approximately millisecond rates are required before a 40° substep. These results were consistent with the average rotation speeds and bulk ATP hydrolysis rates described in the subsection on F_1 in Section 7.4.1 and outlined the essential features of the catalytic cycle illustrated in Figure 7.9B.

In 2003, one of the two reactions in the 40° substep was identified as ATP cleavage, using a mutant that slowly hydrolyses ATP and a slowly hydrolysable substrate ATPγS (Shimabukuro et al. 2003). Further correspondence between chemical reactions and rotational

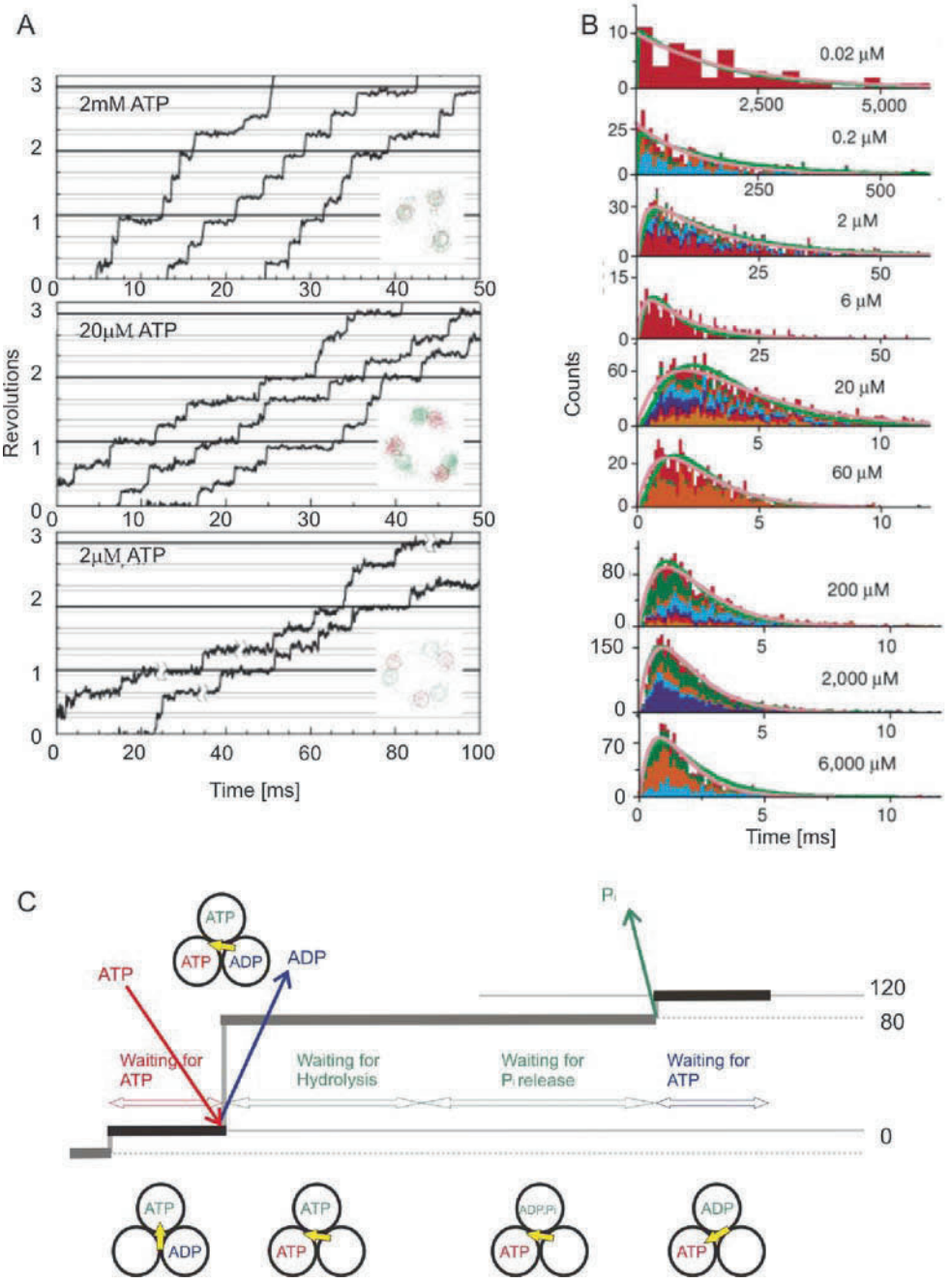


Figure 7.9. A. Angle versus time traces for single 40-nm gold beads attached to the γ-subunit of immobilized F_1, with three different adenosine triphosphate concentrations [ATP]. The larger substep is not resolved at the highest [ATP] **(panel)**. The insets show the traces of the bead positions from which angles were calculated. **B.** A schematic of the mechanochemical cycle of F_1. The inorganic phosphate (P_i) release from the most recently hydrolysed ATP molecule is shown as the trigger of the smaller substep, although it is also possible that each P_i is retained for a further 120°. ADP, adenosine diphosphate. [Panel A, from Yasuda et al. (2001), with permission. Panel C, adapted from Adachi et al. (2007).]

steps was investigated in 2004 by simultaneous observation of the binding of fluorescently labelled ATP and the rotation of bead pairs (Nishizaka et al. 2004). It was found that (1) ATP (or ADP) stays bound to the F_1 during two 120° steps, (2) the 80° substep happens immediately after the fluorescently labelled ATP binds to the motor, and (3) the bead angle and polarization of the fluorescent dye molecule were correlated, implying that the orientation of the γ-subunit dictates which β-subunit will bind the next ATP. In 2007 further observation of fluorescently labelled ATP with a total internal reflection fluorescence microscope (TIRFM) revealed that ADP is released during the 80° substep after it has been bound for two 120° steps, not, for example, during the last 40° substep of the two 120° steps during which it stays bound (Adachi et al. 2007). These results were also taken as an indication that the affinity of ADP decreases with rotation and that the ADP release contributes in energy to the 80° rotation induced by the ATP binding (Adachi et al. 2007). In addition, by adding 200–500 mM [P_i] to the TF_1 rotation observation chamber, the experiments of 2007 observed an increased in dwell time after the 80° substep and thus identifying the second of the two ∼1-ms reactions in the 40° substep with P_i release (Adachi et al. 2007). In terms of site occupancy, two of the three catalytic sites were occupied at high [ATP], whereas at lower [ATP] a variety of irregular behaviours was observed, most likely representing nonmajor reaction pathways (Adachi et al. 2007). The current best estimate of the mechanochemical coupling scheme in which all three of the β-subunits participate in one 120° step (Ariga et al. 2007) is given in Figure 7.9C (Adachi et al. 2007). One open question is whether P_i is released immediately after ATP hydrolysis, as illustrated in Figure 7.9C, or stays bound for an extra 120° step (Adachi et al. 2007).

The following conclusions were drawn from measurements of average and transient torque: (1) The so-called "DELSEED" region, a conserved acidic region on the β-subunit in contact with the γ-subunit, does not have a direct role (Hara et al. 2000). (2) Torque generated by mutants with one, two, or three altered catalytic sites that bind ATP slowly is unchanged, despite the fact that the binding affinity (k_{on} rate) is reduced (Ariga et al. 2002). The torque generated by the wild-type F_1 when using other nucleotide triphosphates whose k_{on} rate is lower than that of ATP was also unchanged (Noji et al. 2001). This can be explained with the "binding zipper" model proposed by Oster and Wang (2000), which divides the ATP-binding process into docking of ATP (described by k_{on}) and the torque generation as a result of the zipping of hydrogen bounds between the P_i moiety of ATP and residues of the catalytic site. (3) The rotor tip inside the $\alpha_3\beta_3$ hexamer of TF_1 is not strictly necessary for torque generation (Hossain et al. 2006, Iwamoto et al. 1990, Sokolov et al. 1999, Müller et al. 2002). Recently the role of the rotor shaft in the generation of torque was investigated further (Furuike et al. 2008). In 2008 the rotor tip placed inside of the stator hexamer was truncated up to the point where what remained of the γ-subunit was almost all outside the stator cavity. All of the mutants rotated in the right direction, with lower speeds and increasingly irregular motion observed as the γ-subunit was shortened (Furuike et al. 2008). It seems that the top surface of the $\alpha_3\beta_3$ hexamer provides the interactions with γ required to produce torque, and the fixed pivot and rigid axle are not needed but stabilize the torque generation, making if more effective. However, the speed of rotation in the truncated mutants was the same as would be predicted from the ATP hydrolysis rate of F_1 lacking the γ-subunit, indicating that care should be taken in interpreting these results. Perhaps the truncated γ-subunit merely reports the intrinsic residual activity of the $\alpha_3\beta_3$ hexamer rather than retaining any of its natural coordinating function.

7.5.2. Bacterial Flagellar Motor

Independent Torque-Generating Units

Proton flux in the bacterial flagellar motor passes through torque-generating stator units, probably consisting of four MotB and two MotA proteins (Sato and Homma 2000, Yorimitsu et al. 2004). Early studies showed that *mot* mutants of *E. coli* do not rotate (Armstrong and Adler 1969) but can be "resurrected" through protein synthesis by lambda-transducing phages (Silverman et al. 1976). This resurrection has also been studied on the level of single motors. Cells of a *motB* strain carrying plasmids with wild-type *motB* genes controlled by the *lac* promoter, as well as cells with *motA* and *motB* point and deletion mutations carrying plasmids with wild-type *motA* and *motB* genes, were tethered to the surface and resurrected. Stepwise increments in motor speed rotation were observed after addition of the inducer, showing that bacterial flagellar motors contain several independent torque-generating units (Block and Berg 1984, Blair and Berg 1988). The maximum number of speed increments observed in these early experiments was 8 (Blair and Berg 1988). More recently, using polystyrene beads attached to the sheared filaments of *E. coli* wild-type and Na^+-driven chimeric motors, the number of discrete steps observed has increased to at least 11 (Reid et al. 2006). Stepwise decreases in motor rotation speed upon activation of an irreversible Na^+-channel inhibitor were observed in Na^+-driven alkalophilic *Bacillus* as well, indicating that this species has up to 9 independent stator units (Muramoto et al. 1994). Transient speed changes observed in wild-type motor rotation suggested the possibility that the units are not fixed but in a process of constant turnover (Reid et al. 2006). This was confirmed in 2006 by observing green-fluorescent protein (GFP)–labelled MotB in live cells, using TIRFM (Leake et al. 2006). Motors were seen as bright spots, which were estimated by quantitative photobleaching to contain an average of ~22 GFP molecules, consistent with the previous estimate of at least 11 units each containing two copies of MotB. A freely diffusing pool of ~200 GFP-MotB molecules was observed in the cell membrane, and the exchange between the pool and stator units within the motor on a timescale of minutes was demonstrated using fluorescence recovery after photobleaching (Leake et al. 2006).

Motor speed is proportional to the number of units in the high-load regime. A slight reduction in the speed per unit at higher number was seen for H^+- but not for Na^+-driven motors (Reid et al. 2006). This is attributed either to steric hindrance once a large number of units is present or possibly to local depletion of H^+ ions. Torque–speed curves for fully energized H^+ motors containing small numbers of units have been obtained using polystyrene beads 0.3–1.0 μm in diameter attached to flagellar stubs (Ryu et al. 2000) and 60-nm gold beads attached to hooks (Yuan and Berg 2008). The experiments with polystyrene beads showed that the torque generated at a given speed by a motor with N units is simply N times the torque generated by a single unit at that speed. This means that one unit rotates the motor as fast as many as the load tends to zero, a prediction that was confirmed by the gold bead experiment: Under experimental conditions that gave stepwise resurrection with larger loads, gold beads showed only sudden jumps from 0 Hz to the zero-load speed of ~300 Hz. These results can be reproduced by a model in which the duty ratio of each unit is high—that is, it seldom if ever detaches from the rotor—and in which the rate-limiting step at low load is insensitive to the torque exerted upon one unit by the action of other units upon the rotor (Ryu et al. 2000).

Interactions between Rotor and Stator

By combining atomic structures of fragments of the rotor proteins FliG, FliM, and FliN with cross-linking and mutational studies, the following picture of rotor–stator interactions involved in torque generation has emerged. Stator proteins form ion channels, with a single

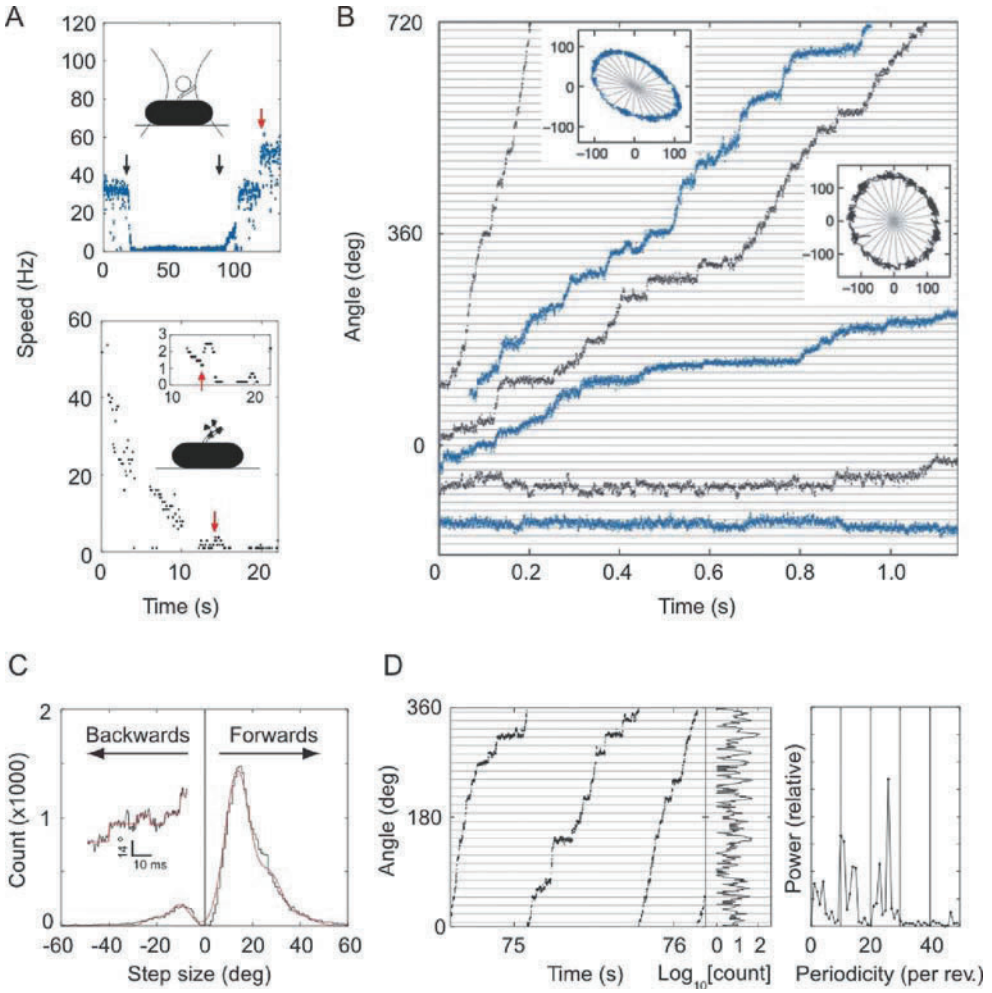


Figure 7.10. Steps in slow flagellar rotation. **A.** Reducing sodium-motive force (SMF) and motor speed of a chimeric Na^+-driven flagellar motor in *Escherichia coli* **(upper)** by lowering the external Na^+ concentration (from 5 mM to 0.1 mM and back; *black arrows*) and **(lower)** by photodamage. The speed doublings marked by red arrows indicate a probable change from one to two stators. **B.** Stepping rotation of flagellar motors with a range of average speeds depending on different SMF. Insets show the positions of beads attached to flagellar filaments; scales are in nanometres. Horizontal and radial lines indicate 1/26 revolution. **C.** Step-size distribution (*black*) with multiple Gaussian fit (*red*). The peak of forwards steps is 13.7°, indicating 26 steps per revolution. An example of steps identified by a step-finding algorithm is shown in the inset. **D.** Plot of angle against time during three revolutions, a histogram of dwell angles for the same revolutions, and the power spectrum of that histogram. The peak at 26 per revolution corresponds to a step size of 13.8° and shows that the motor stops at the same angles on successive revolutions. Speeds shown in black were measured using optical interferometry, those in blue using high-speed fluorescence microscopy (insets in panel A). [Adapted from Sowa et al. (2005).]

essential conserved acidic residue in MotB (Asp32 in *E. coli*) identified as the key ion-binding site (Sharp et al. 1995a, b). Five charged residues in the C-terminal domain of FliG interact with two charged residues in the cytoplasmic domain of MotA (Lloyd and Blair 1997, Zhou and Blair 1997, Zhou et al. 1998a) to generate torque. FliG and PomA in the Na^+-driven motor of *V. alginolyticus* interact via different charged residues but in a similar pattern (Yakushi et al. 2006, Yorimitsu et al. 2002, 2003). Single mutations of any of these residues reduce but do not abolish rotation, and charge-reversing second mutations can reverse the effect of single mutations. Therefore FliG–MotA interaction during torque generation is probably an electrostatic interaction between extended parts of the two proteins. The interaction is thought to be driven by Asp32 of MotB accepting an ion and inducing conformational changes in MotA (Braun et al. 1999, Kojima and Blair 2001).

Stepping Rotation

Unlike F_1 and several other ATP-driven motors discussed in this book, direct observation of steps corresponding to single mechanochemical cycles of the BFM has been limited by the high speed and small step size of the motor. Analysis of speed fluctuations of motors with different numbers of torque-generating units, combined with electrorotation and measurements with broken motors, demonstrated that the BFM is a stepping motor rather than a "fluid drive" (Samuel and Berg 1995, 1996). Assuming Poisson statistics for steps led to an estimate of ~50 steps per unit per revolution. The proportionality between step number and unit number is consistent with a high duty ratio. It should be noted, however, that the "steps" here need not correspond to actual physical movements of the rotor; rather they are strictly defined as "independent sequential stochastic events." Just as the passage of several ions in F_o might be stored up to produce a single observable 120° step in F_1F_o, so might the events identified by the fluctuation analysis be linked only indirectly to observable steps.

Such steps were first observed in 2005 in a Na^+-driven chimeric flagellar motor in *E. coli* (Sowa et al. 2005). At low sodium motive force and with controlled expression of a small number of torque-generating units, 26 steps per revolution were seen (Figure 7.10). This identifies the inner lobe of the C-ring as the site of stepping and thus presumably of torque generation. However, many questions remain unanswered. It has proved difficult to reconcile 26 steps per revolution with the estimate of 36 ± 6 ions per stator per revolution estimated from the energetics of flagellar rotation in the chimeric motor and the assumption that each step should correspond to an integer number of ions crossing the membrane. Future investigations of stepping under well-controlled conditions of known SMF and number of stator units will be necessary to reveal the details of the mechanism of the BFM in the way that measurements of steps and substeps have done for F_1.

References

Abrahams JP, Leslie AG, Lutter R, Walker JE (1994) Structure at 2.8 A resolution of F_1-ATPase from bovine heart mitochondria. Nature 370: 621–628.

Adachi K, Yasuda R, Noji H, Itoh H, Harada Y, Yoshida M, Kinosita K (2000) Stepping rotation of F_1-ATPase visualized through angle-resolved single-fluorophore imaging. Proc Natl Acad Sci USA 97: 7243–7247.

Adachi K, Oiwa K, Nishizaka T, Furuike S, Noji H, Itoh H, Yoshida M, Kinosita K (2007) Coupling of rotation and catalysis in F(1)-ATPase revealed by single-molecule imaging and manipulation. Cell 130: 309–321.

Aizawa SI (1996) Flagellar assembly in *Salmonella typhimurium*. Mol Microbiol 19: 1–5.

Althoff G, Lill H, Junge W (1989) Proton channel of the chloroplast ATP synthase, CF_0: its time-averaged single-channel conductance as function of pH, temperature, isotopic and ionic medium composition. J Membr Biol 108: 263–271.

Arai H, Terres G, Pink S, Forgac M (1988) Topography and subunit stoichiometry of the coated vesicle proton pump. J Biol Chem 263: 8796–8802.

Ariga T, Masaike T, Noji H, Yoshida M (2002) Stepping rotation of F(1)-ATPase with one, two, or three altered catalytic sites that bind ATP only slowly. J Biol Chem 277: 24870–24874.

Ariga T, Muneyuki E, Yoshida M (2007) F_1-ATPase rotates by an asymmetric, sequential mechanism using all three catalytic subunits. Nat Struct Mol Biol 14: 841–846.

Armstrong JB, Adler J (1969) Complementation of nonchemotactic mutants of *Escherichia coli*. Genetics 61: 61–66.

Asai Y, Yakushi T, Kawagishi I, Homma M (2003) Ion-coupling determinants of Na+-driven and H+-driven flagellar motors. J Mol Biol 327: 453–463.

Baker MD, Wolanin PM, Stock JB (2006) Signal transduction in bacterial chemotaxis. Bioessays 28: 9–22.

Berg HC, Anderson RA (1973) Bacteria swim by rotating their flagellar filaments. Nature 245: 380–382.

Berg HC, Turner L (1993) Torque generated by the flagellar motor of *Escherichia coli*. Biophys J 65: 2201–2216.

Berg HC (2003) The rotary motor of bacterial flagella. Annu Rev Biochem 72: 19–54.

Berg JM, Tymoczko JL, Stryer L (2002) *Biochemistry*. WH Freeman, New York.

Berry RM, Berg HC (1996) Torque generated by the bacterial flagellar motor close to stall. Biophys J 71: 3501–3510.

Berry RM, Berg HC (1997) Absence of a barrier to backwards rotation of the bacterial flagellar motor demonstrated with optical tweezers. Proc Natl Acad Sci USA 94: 14433–14437.

Berry RM, Berg HC (1999) Torque generated by the flagellar motor of *Escherichia coli* while driven backward. Biophys J 76: 580–587.

Berry RM, Turner L, Berg HC (1995) Mechanical limits of bacterial flagellar motors probed by electrorotation. Biophys J 69: 280–286.

Beyenbach KW, Wieczorek H (2006) The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. J Exp Biol 209: 577–589.

Blair DF, Berg HC (1988) Restoration of torque in defective flagellar motors. Science 242: 1678–1681.

Blair DF (1995) How bacteria sense and swim. Annu Rev Microbiol 49: 489–522.

Blair DF, Berg HC (1991) Mutations in the MotA protein of *Escherichia coli* reveal domains critical for proton conduction. J Mol Biol 173: 4049–4055.

Block SM, Berg HC (1984) Successive incorporation of force-generating units in the bacterial rotary motor. Nature 309: 470–472.

Block SM, Blair DF, Berg HC (1989) Compliance of bacterial flagella measured with optical tweezers. Nature 338: 514–518.

Block SM, Fahrner KA, Berg HC (1991) Visualization of bacterial flagella by video-enhanced light microscopy. J Bacteriol 173: 933–936.

Borsch M, Diez M, Zimmermann B, Reutera R, Graber P (2002) Stepwise rotation of the γ-subunit of EF_0F_1-ATP synthase observed by intramolecular single-molecule fluorescence resonance energy transfer. FEBS Lett 527: 147–152.

Bowler MW, Montgomery MG, Leslie AG, Walker JE (2007) Ground state structure of F1-ATPase from bovine heart mitochondria at 1.9 A resolution. J Biol Chem 282: 14238–14242.

Boyer PD (1993) The binding change mechanism for ATP synthase-some probabilities and possibilities. Biochim Biophys Acta 1140: 215–250.

Braun TF, Poulson S, Gully JB, Empey JC, Van Way S, Putnam A, Blair DF (1999) Function of proline residues of MotA in torque generation by the flagellar motor of *Escherichia coli.* J Bacteriol 181: 3542–3551.

Brown PN, Hill CP, Blair DF (2002) Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG. EMBO J 21: 3225–3234.

Brown PN, Mathews MA, Joss LA, Hill CP, Blair DF (2005) Crystal structure of the flagellar rotor protein FliN from *Thermotoga maritime*. J Bacteriol 187: 2890–2902.

Brown PN, Terrazas M, Paul K, Blair DF (2007) Mutational analysis of the flagellar protein FliG: sites of interaction with FliM and implications for organization of the switch complex. J Bacteriol 189: 305–312.

Chen X, Berg HC (2000a) Torque-speed relationship of the flagellar rotary motor of *Escherichia coli*. Biophys J 78: 1036–1041.

Chen X, Berg HC (2000b). Solvent-isotope and pH effects on flagellar rotation in *Escherichia coli*. Biophys J 78: 2280–2284.

Cherepanov DA, Junge W (2001) Viscoelastic dynamics of actin filaments coupled to rotary F-ATPase: curvature as an indicator of the torque. Biophys J 81: 1234–1244.

Chernyak BV, Glagolev AN, Sherman MY, Skulachev VP (1983) A novel type of energetics in a marine alkalitolerant bacterium. DmNa-driven motility and sodium cycle. FEBS Lett 164: 38–42.

Chun SY, Parkinson JS (1988) Bacterial motility: membrane topology of the *Escherichia coli* MotB protein. Science 239: 276–278.

Darnton NC, Berg HC (2007) Force-extension measurements on bacterial flagella: triggering polymorphic transformations. Biophys. J 92(6): 2230–2236.

Darnton NC, Turner L, Rojevsky S, Berg HC (2007) On torque and tumbling in swimming *Escherichia coli*. J Bacteriol 189: 1756–1764.

Diez M, Zimmermann B, Börsch M, König M, Schweinberger E, Steigmiller S, Reuter R, Felekyan S, Kudryavtsev V, Seidel CA, Gräber P (2004) Proton powered subunit rotation in single membrane-bound F_0F_1-ATP synthase. Nat Struct Mol Biol 11: 135–141.

Duncan TM, Bulygin VV, Zhou Y, Hutcheon ML, Cross RL (1995) Rotation of subunits during catalysis by *Escherichia coli* F1-ATPase. Proc Natl Acad Sci USA 92: 10964–10968.

Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA (1997) The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. Annu Rev Cell Dev Biol 13: 457–512.

Felle H, Porter JS, Slayman CL, Kaback HR (1980) Quantitative measurements of membrane potential in *Escherichia coli*. Biochemistry 19: 3585–3590.

Feniouk BA, Cherepanov DA, Junge W, Mulkidjanian AY (2001) Coupling of proton flow to ATP synthesis in *Rhodobacter capsula*tus: F(0)F(1)-ATP synthase is absent from about half of chromatophores. Biochim Biophys Acta 1506: 189–203.

Feniouk BA, Cherepanov DA, Voskoboynikova NE, Mulkidjanian AY, Junge W (2002) Chromatophore vesicles of *Rhodobacter capsulatus* contain on average one F(0)F(1)-ATP synthase each. Biophys J 82: 1115–1122.

Feniouk BA, Kozlova MA, Knorre DA, Cherepanov DA, Mulkidjanian AY, Junge W (2004) The proton driven rotor of ATP synthase: ohmic conductance (10 fS), and absence of voltage gating. Biophys J 86: 4094–4109.

Friedl P, Schairer HU (1981) The isolated F_0 of *Escherichia coli* ATP-synthase is reconstitutively active in H+-conduction and ATP-dependent energy-transduction. FEBS Lett 128: 261–264.

Fritz M, Klyszejko AL, Morgner N, Vonck J, Brutschy B, Muller DJ, Meier T, Müller V (2008) An intermediate step in the evolution of ATPases—a hybrid Fo–Vo rotor in a bacterial Na+ F1F0 ATP synthase. FEBS J 275: 1999–2007.

Fung DC, Berg HC (1995) Powering the flagellar motor of *Escherichia coli* with an external voltage source. Nature 375: 809–812.

Furuike S, Hossain MD, Maki Y, Adachi K, Suzuki T, Kohori A, Itoh H, Yoshida M, Kinosita K (2008) Axle-less F1-ATPase rotates in the correct direction. Science 319: 955–958.

Gabel CV, Berg HC (2003) The speed of the flagellar rotary motor of *Escherichia coli* varies linearly with protonmotive force. Proc Natl Acad Sci USA 100: 8748–8751.

Gibbons C, Montgomery MG, Leslie AG, Walker JE (2000) The structure of the central stalk in bovine F_1-ATPase at 2.4 A resolution. Nat Struct Biol 7: 1055–1061.

Hara KY, Noji H, Bald D, Yasuda R, Kinosita K, Yoshida M (2000) The role of the DELSEED motif of the beta subunit in rotation of F1-ATPase. J Biol Chem 275: 14260–14263.

Hirono-Hara Y, Noji H, Nishiura M, Muneyuki E, Hara KY, Yasuda R, Kinosita K, Yoshida M (2001) Pause and rotation of F(1)-ATPase during catalysis. Proc Natl Acad Sci USA 98: 13649–13654.

Hirono-Hara Y, Ishizuka K, Kinosita K, Yoshida M, Noji H (2005) Activation of pausing F_1 motor by external force. Proc Natl Acad Sci USA 102: 4288–4293.

Hirota N, Kitada M, Imae Y (1981) Flagellar motors of alkalophilic *Bacillus* are powered by an electrochemical potential gradient of Na^+. FEBS Lett 132: 278–280.

Hirota N, Imae Y (1983) Na+-driven flagellar motors of an alkalophilic *Bacillus* strain YN-1. J Biol Chem 258: 10577–10581.

Holzenburg A, Jones PC, Franklin T, Pali T, Heimburg T, Marsh D, Findlay JB, Finbow ME (1993) Evidence for a common structure for a class of membrane channels. Eur J Biochem 213: 21–30.

Hossain MD, Furuike S, Maki Y, Adachi K, Ali MY, Huq M, Itoh H, Yoshida M, Kinosita K (2006) The rotor tip inside a bearing of a thermophilic F_1-ATPase is dispensable for torque generation. Biophys J 90: 4195–4203.

Hotani H (1976) Light microscope study of mixed helices in reconstituted *Salmonella* flagella. J Mol Biol 106: 151–166.

Hudson GS, Mason JG, Holton TA, Koller B, Cox GB, Whitfeld PR, Bottomley W (1987) A gene cluster in the spinach and pea chloroplast genomes encoding one CF1 and three CF0 subunits of the H^+-ATP synthase complex and the ribosomal protein S2. J Mol Biol 196: 283–298.

Iko Y, Sambongi Y, Tanabe M, Iwamoto-Kihara A, Saito K, Ueda I, Wada Y, Futai M (2001) ATP synthase F_1 sector rotation. Defective torque generation in the beta subunit Ser-174 to Phe mutant and its suppression by second mutations. J Biol Chem 276: 47508–47511.

Inoue Y, Lo CJ, Fukuoka H, Takahashi H, Sowa Y, Pilizota T, Wadhams GH, Homma M, Berry RM, Ishijima A (2008) Torque–speed relationships of Na+-driven chimeric flagellar motors in *Escherichia coli*. J Mol Biol 376: 1251–1259.

Irikura VM, Kihara M, Yamaguchi S, Sockett H, Macnab RM (1993) *Salmonella typhimurium* fliG and fliN mutations causing defects in assembly, rotation, and switching of the flagellar motor. J Bacteriol 175: 802–810.

Itoh H, Takahashi A, Adachi K, Noji H, Yasuda R, Yoshida M, Kinosita K (2004) Mechanically driven ATP synthesis by F_1-ATPase. Nature 427: 465–468.

Iwamoto A, Miki J, Maeda M, Futai M (1990) H(+)-ATPase gamma subunit of *Escherichia coli*. Role of the conserved carboxyl-terminal region. J Biol Chem 265: 5043–5048.

Jagendorf AT, Uribe E (1966) ATP formation caused by acid-base transition of spinach chloroplasts. Proc Natl Acad Sci USA 55: 170–177.

Junge W, Pänke O, Cherepanov DA, Gumbiowski K, Müller M, Engelbrecht S (2001) Inter-subunit rotation and elastic power transmission in F0F1-ATPase. FEBS Lett 504: 152–160.

Kaim G, Prummer M, Sick B, Zumofen G, Renn A, Wild UP, Dimroth P (2002) Coupled rotation within single F0F1 enzyme complexes during ATP synthesis or hydrolysis. FEBS Lett 525: 156–163.

Kanazawa H, Mabuchi K, Kayano T, Noumi T, Sekiya T, Futai M (1981) Nucleotide sequence of the genes for F0 components of the proton-translocating ATPase from *Escherichia coli*: prediction of the primary structure of F_0 subunits. Biochem Biophys Res Commun 103: 604–612.

Kashket ER, Blanchard AG, Metzger WC (1980) Proton motive force during growth of *Streptococcus lactis* cells. J Bacteriol 143: 128–134.

Kato-Yamada Y, Noji H, Yasuda R, Kinosita K, Yoshida M (1998) Direct observation of the rotation of epsilon subunit in F_1-ATPase. J Biol Chem 273: 19375–19377.

Khan S, Meister M, Berg HC (1985) Constraints on flagellar rotation. J Mol Biol 184: 645–656.

Kihara M, Miller GU, Macnab RM (2000) Deletion analysis of the flagellar switch protein FliG of *Salmonella*. J Bacteriol 182: 3022–3028.

Kojima S, Blair DF (2001) Conformational change in the stator of the bacterial flagellar motor. Biochemistry 40: 13041–13050.

Kudo S, Magariyama Y, Aizawa S (1990) Abrupt changes in flagellar rotation observed by laser dark-field microscopy. Nature 346: 677–680.

Larsen SH, Adler J, Gargus JJ, Hogg RW (1974) Chemomechanical coupling without ATP: the source of energy for motility and chemotaxis in bacteria. Proc Natl Acad Sci USA 71: 1239–1243.

Leake MC, Chandler JH, Wadhams GH, Bai F, Berry RM, Armitage JP (2006) Stoichiometry and turnover in single, functioning membrane protein complexes. Nature 443: 355–358.

Lill H, Engelbrecht S, Schönknecht G, Junge W (1986) The proton channel, CF0, in thylakoid membranes. Only a low proportion of CF_1-lacking CF_0 is active with a high unit conductance (169 fS). Eur J Biochem 160: 627–634.

Lloyd SA, Tang H, Wang X, Billings S, Blair DF (1996) Torque generation in the flagellar motor of *Escherichia coli*: evidence of a direct role for FliG but not for FliM or FliN. J Bacteriol 178: 223–231.

Lloyd SA, Blair DF (1997) Charged residues of the rotor protein FliG essential for torque generation in the flagellar motor of *Escherichia coli*. J Mol Biol 266: 733–744.

Lloyd SA, Whitbey FG, Blair DF, Hill CP (1999) Structure of the C-terminal domain of FliG, a component of the rotor in the bacterial flagellar motor. Nature 400: 472–475.

Lo CJ, Leake MC, Berry RM (2006) Fluorescence measurement of intracellular sodium concentration in single *Escherichia coli* cells. Biophys J 90: 357–365.

Lo CJ, Leake MC, Pilizota T, Berry RM (2007) Nonequivalence of membrane voltage and ion-gradient as driving forces for the bacterial flagellar motor at low load. Biophys J 93: 294–302.

Lowe G, Meister M, Berg HC (1987) Rapid rotation of flagellar bundles in swimming bacteria. Nature 325: 637–640.

Macnab RM (1976) Examination of bacterial flagellation by dark-field microscopy. J Clin Microbiol 4: 258–265.

Macnab RM (2003) How bacteria assemble flagella. Annu Rev Biochem 57: 77–100.

Magariyama Y, Susgiyama S, Muramoto K, Maekawa Y, Kawagishi I, Imae Y, Kudo S (1994) Very fast flagellar rotation. Nature 371: 752.

Manson MD, Tedesco P, Berg HC, Harold FM, Van Der Drift C (1977) A protonmotive force drives bacterial flagella. Proc Natl Acad Sci USA 74: 3060–3064.

Manson MD, Tedesco PM, Berg HC (1980) Energetics of flagellar rotation in bacteria. J Mol Biol 138: 541–561.

Matsuura S, Shioi J, Imae Y (1977) Motility in *Bacillus subtilis* driven by an artificial protonmotive force. FEBS Lett 82: 187–190.

Meier T, Matthey U, von Ballmoos C, Vonck J, Krug von Nidda T, Kühlbrandt W, Dimroth P (2003) Evidence for structural integrity in the undecameric *c*-rings isolated from sodium ATP synthases. J Mol Biol 325: 389–397.

Meier T, Polzer P, Diederichs K, Welte W, Dimroth P (2005) Structure of the rotor ring of F-type Na+-ATPase from *Ilyobacter tartaricus*. Science 308: 659–662.

Meister M, Berg HC (1987) The stall torque of the bacterial flagellar motor. Biophys J 52: 413–419.

Menz RI, Walker JE, Leslie AG (2001) Structure of bovine mitochondrial F_1-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. Cell 106: 331–341.

Meyer Zu Tittingdorf JM, Rexroth S, Schäfer E, Schlichting R, Giersch C, Dencher NA, Seelert H (2004) The stoichiometry of the chloroplast ATP synthase oligomer III in *Chlamydomonas reinhardtii* is not affected by the metabolic state. Biochim Biophys Acta 1659:92–99.

Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature 191: 144–148.

Mitchell P (1966) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. Biol Rev Camb Philos Soc 1966 41: 445–502.

Mitome N, Ono S, Suzuki T, Shimabukuro K, Muneyuki E, Yoshida M (2002) The presence of phosphate at a catalytic site suppresses the formation of the MgADP-inhibited form of F_1-ATPase. Eur J Biochem 269: 53–60.

de Mot R, Vanderleyden J (1994) A conserved surface-exposed domain in major outer membrane proteins of pathogenic *Pseudomonas* and *Branhamella* species shares sequence homology with the calcium-binding repeats of the eukaryotic extracellular matrix protein thrombospondin. Mol Microbiol 13: 379–380.

Müller M, Pänke O, Junge W, Engelbrecht S (2002) F_1-ATPase, the C-terminal end of subunit gamma is not required for ATP hydrolysis-driven rotation. J Biol Chem 277: 23308–23313.

Muramoto K, Sugiyama S, Cragoe EJ, Imae Y (1994) Successive inactivation of the force-generating units of sodium-driven bacterial flagellar motors by a photoreactive amiloride analog. J Biol Chem 269: 3374–3380.

Muramoto K, Kawagishi I, Kudo S, Magariyama Y, Imae Y, Homma M (1995) High-speed rotation and speed stability of the sodium-driven flagellar motor in *Vibrio alginolyticus*. J Mol Biol 251: 50–58.

Murata T, Yamato I, Kakinuma Y, Leslie AG, Walker JE (2005) Structure of the rotor of the V-Type Na+-ATPase from *Enterococcus hirae*. Science 308: 654–659.

Murphy GE, Leadbetter JR, Jensen GJ (2006) *In situ* structure of the complete *Treponema primitia* flagellar motor. Nature 442: 1062–1064.

Nakanishi-Matsui M, Kashiwagi S, Hosokawa H, Cipriano DJ, Dunn SD, Wada Y, Futai M (2006) Stochastic high-speed rotation of *Escherichia coli* ATP synthase F_1 sector: the epsilon subunit-sensitive rotation. J Biol Chem 281: 4126–4131.

Nakanishi-Matsui M, Kashiwagi S, Ubukata T, Iwamoto-Kihara A, Wada Y, Futai M (2007) Rotational catalysis of *Escherichia coli* ATP synthase F_1 sector. Stochastic fluctuation and a key domain of the beta subunit. J Biol Chem 282: 20698–20704.

Negrin RS, Foster DL, Fillingame RH (1980) Energy-transducing H+-ATPase of *Escherichia coli*. Reconstitution of proton translocation activity of the intrinsic membrane sector. J Biol Chem 255: 5643–5648.

Nicholls DG, Ferguson SJ (2002) *Bioenergetics*. Academic Press, San Diego, CA.

Nishi T, Forgac M (2002) The vacuolar (H+)-ATPases—nature's most versatile proton pumps. Nat Rev Mol Cell Biol 3: 94–103.

Nishio K, Iwamoto-Kihara A, Yamamoto A, Wada Y, Futai M (2002) Subunit rotation of ATP synthase embedded in membranes: a or beta subunit rotation relative to the c subunit ring. Proc Natl Acad Sci USA 99: 13448–13452.

Nishizaka T, Oiwa K, Noji H, Kimura S, Muneyuki E, Yoshida M, Kinosita K (2004) Chemomechanical coupling in F_1-ATPase revealed by simultaneous observation of nucleotide kinetics and rotation. Nat Struct Mol Biol 11: 142–148.

Noji H, Yasuda R, Yoshida M, Kinosita K (1997) Direct observation of the rotation of F1-ATPase. Nature 386: 299–302.

Noji H, Bald D, Yasuda R, Itoh H, Yoshida M, Kinosita K (2001) Purine but not pyrimidine nucleotides support rotation of F_1-ATPase. J Biol Chem 276: 25480–25486.

Omote H, Sambonmatsu N, Saito K, Sambongi Y, Iwamoto-Kihara A, Yanagida T, Wada Y, Futai M (1999) The gamma-subunit rotation and torque generation in F_1-ATPase from wild-type or uncoupled mutant *Escherichia coli*. Proc Natl Acad Sci USA 96: 7780–7784.

O'Neill J, Roujeinikova A (2008) Cloning, purification and crystallization of MotB, a stator component of the proton-driven bacterial flagellar motor. Acta Cryst F64: 561–563.

Oster G, Wang H (2000) Reverse engineering a protein: the mechanochemistry of ATP synthase. Biochim Biophys Acta 1458: 482–510.

Pänke O, Cherepanov DA, Gumbiowski K, Engelbrecht S, Junge W (2001) Viscoelastic dynamics of actin filaments coupled to rotary F-ATPase: angular torque profile of the enzyme. Biophys J 81:1220–1233.

Park SY, Lowder B, Bilwes AM, Blair DF, Crane BR (2006) Structure of FliM provides insight into assembly of the switch complex in the bacterial flagella motor. Proc Natl Acad Sci USA 103: 11886–11891.

Paster E, Ryu WS (2008) The thermal impulse response of *Escherichia coli*. PNAS 105: 5373–5377.

Pilizota T, Bilyard T, Bai F, Futai M, Hosokawa H, Berry RM (2007) A programmable optical angle clamp for rotary molecular motors. Biophys J 93: 264–275.

Pogoryelov D, Yu J, Meier T, Vonck J, Dimroth P, Muller DJ (2005) The C15 ring of the Spirulina platensis F-ATP synthase: F1/F0 symmetry mismatch is not obligatory. EMBO Rep 6: 1040–1044.

Pogoryelov D, Reichen C, Klyszejko AL, Brunisholz R, Muller DJ, Dimroth P, Meier T (2007) The oligomeric state of C rings from cyanobacterial f-ATP synthases varies from 13 to 15. J Bacteriol 189: 5895–5902.

Powell B, Graham LA, Stevens TH (2000) Molecular characterization of the yeast vacuolar H+-ATPase proton pore. J Biol Chem 275: 23654–23660.

Rao R, Senior AE (1987) The properties of hybrid F1-ATPase enzymes suggest that a cyclical catalytic mechanism involving three catalytic sites occurs. J Biol Chem 262: 17450–17454.

Reid SW, Leake MC, Chandler JH, Lo CJ, Armitage JP, Berry RM (2006) The maximum number of torque-generating units in the flagellar motor of *Escherichia coli* is at least 11. Proc Natl Acad Sci USA 103: 8066–8071.

Rondelez Y, Tresset G, Nakashima T, Kato-Yamada Y, Fujita H, Takeuchi S, Noji H (2005) Highly coupled ATP synthesis by F1-ATPase single molecules. Nature 433: 773–777.

Rubinstein JL, Walker JE, Henderson R (2003) Structure of the mitochondrial ATP synthase by electron cryomicroscopy. EMBO J 22: 6182–6192.

Ryu WS, Berry RM, Berg HC (2000) Torque-generating units of the flagellar motor of *Escherichia coli* have a high duty ratio. Nature 403: 444–447.

Sabbert D, Engelbrecht S, Junge W (1996) Intersubunit rotation in active F-ATPase. Nature 381: 623–625.

Sakaki N, Shimo-Kon R, Adachi K, Itoh H, Furuike S, Muneyuki E, Yoshida M, Kinosita K (2005) One rotary mechanism for F_1-ATPase over ATP concentrations from millimolar down to nanomolar. Biophys J 88: 2047–2056.

Sambongi Y, Iko Y, Tanabe M, Omote H, Iwamoto-Kihara A, Ueda I, Yanagida T, Wada Y, Futai M (1999) Mechanical rotation of the c subunit oligomer in ATP synthase (F0F1): direct observation. Science 286: 1722–1724.

Samuel AD, Berg HC (1995) Fluctuation analysis of rotational speeds of the bacterial flagellar motor. Proc Natl Acad Sci USA 92: 3502–3506.

Samuel AD, Berg HC (1996) Torque-generating units of the bacterial flagellar motor step independently. Biophys J 71: 918–923.

Sato K, Homma M (2000) Multimeric structure of PomA, a component of the Na+-driven polar flagellar motor of *Vibrio alginolyticus*. J Biol Chem 275: 20223–20228.

Schemidt RA, Qu J, Williams JR, Brusilow WS (1998) Effects of carbon source on expression of F0 genes and on the stoichiometry of the c subunit in the F_1F_0 ATPase of *Escherichia coli*. J Bacteriol 180: 3205–3208.

Schneider E, Altendorf K (1982) ATP synthetase (F_1F_0) of *Escherichia coli* K-12. High-yield preparation of functional F0 by hydrophobic affinity chromatography. Eur J Biochem 126: 149–153.

Schwem BE, Fillingame RH (2006) Cross-linking between helices within subunit a of *Escherichia coli* ATP synthase defines the transmembrane packing of a four helix bundle. J Biol Chem 281: 37861–37867.

Seelert H, Poetsch A, Dencher NA, Engel A, Stahlberg H, Müller DJ (2000) Structural biology: proton-powered turbine of a plant motor. Nature 405: 418–419.

Sharp LL, Zhou J, Blair DF (1995a) Features of MotA proton channel structure revealed by tryptophan-scanning mutagenesis. Proc Natl Acad Sci USA 92: 7946–7950.

Sharp LL, Zhou J, Blair DF (1995b) Tryptophan-scanning mutagenesis of MotB, an integral membrane protein essential for flagellar rotation in *Escherichia coli*. Biochemistry 34: 9166–9171.

Shimabukuro K, Yasuda R, Muneyuki E, Hara KY, Kinosita K, Yoshida M (2003) Catalysis and rotation of F1 motor: cleavage of ATP at the catalytic site occurs in 1 ms before 40 degree substep rotation. Proc Natl Acad Sci USA 100: 14731–14736.

Shioi JI, Matsuura S, Imae Y (1980) Quantitative measurements of proton motive force and motility in *Bacillus subtilis*. J Bacteriol 144: 891–897.

Shirakihara Y, Leslie AG, Abrahams JP, Walker JE, Ueda T, Sekimoto Y, Kambara M, Saika K, Kagawa Y, Yoshida M (1997) The crystal structure of the nucleotide-free alpha 3 beta 3 subcomplex of F1-ATPase from the thermophilic *Bacillus* PS3 is a symmetric trimer. Structure 5: 825–836.

Silverman M, Simon M (1974) Flagellar rotation and the mechanism of bacterial motility. Nature 249: 73–74.
Silverman M, Matsumura P, Simon M (1976) The identification of the mot gene product with *Escherichia coli*–lambda hybrids. Proc Natl Acad Sci USA 73: 3126–3130.
Sokolov M, Lu L, Tucker W, Gao F, Gegenheimer PA, Richter ML (1999) The 20 C-terminal amino acid residues of the chloroplast ATP synthase gamma subunit are not essential for activity. J Biol Chem 274: 13824–13829.
Sone N, Hamamoto T, Kagawa Y (1981) pH dependence of H+ conduction through the membrane moiety of the H+-ATPase (F_0xF_1) and effects of tyrosyl residue modification. J Biol Chem 256: 2873–2877.
Sowa Y, Hotta H, Homma M, Ishijima A (2003) Torque–speed relationship of the Na+-driven flagellar motor of *Vibrio alginolyticus*. J Mol Biol 327: 1043–1051.
Sowa Y, Rowe AD, Leake MC, Yakushi T, Homma M, Ishijima A, Berry RM (2005) Direct observation of steps in rotation of the bacterial flagellar motor. Nature 437: 916–919.
Sowa Y and Berry RM (2008) Bacterial flagellar motor. Q Rev Biophys (In press).
Spetzler D, York J, Daniel D, Fromme R, Lowry D, Frasch W (2006) Microsecond time scale rotation measurements of single F_1-ATPase molecules. Biochemistry 45: 3117–3124.
Stahlberg H, Müller DJ, Suda K, Fotiadis D, Engel A, Meier T, Matthey U, Dimroth P (2001) Bacterial Na(+)-ATP synthase has an undecameric rotor. EMBO Rep 2: 229–233.
Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. Science 286: 1700–1705.
Suzuki H, Yonekura K, Namba K (2004) Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis. J Mol Biol 337: 105–113.
Tanabe M, Nishio K, Iko Y, Sambongi Y, Iwamoto-Kihara A, Wada Y, Futai M (2001) Rotation of a complex of the gamma subunit and C ring of *Escherichia coli* ATP synthase. The rotor and stator are interchangeable. Biol Chem 276: 15269–15274.
Thomas DR, Morgan DG, DeRosier DJ (1999) Rotational symmetry of the C ring and a mechanism for the flagellar rotary motor. Proc Natl Acad Sci USA 96: 10134–10139.
Thomas DR, Francis NR, Xu C, DeRosier DJ (2006) The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of *Salmonella enterica serovar Typhimurium*. J Bacteriol 188: 7039–7048.
Turner L, Ryu WS, Berg HC (2000) Real-time imaging of fluorescent flagellar filaments. J Bacteriol 182: 2793–2801.
Ueno T, Oosawa K, Aizawa S (1992) M ring, S ring and proximal rod of the flagellar basal body of *Salmonella typhimurium* are composed of subunits of single protein, FliF. J Mol Biol 227: 672–677.
Ueno T, Oosawa K, Aizawa S (1994) Domain structures of the MS ring component protein (FliF) of the flagellar basal body of *Salmonella typhimurium*. J Mol Biol 236: 546–555.
Wadhams GH, Armitage JP (2004) Making sense of it all: bacterial chemotaxis. Nat Rev Mol Cell Biol 5: 1024–1037.
Wagner R, Apley EC, Hanke W (1989) Single channel H+ currents through reconstituted chloroplast ATP synthase CF_0-CF_1. EMBO J 8: 2827–2834.
Walker JE, Fearnley IM, Gay NJ, Gibson BW, Northrop FD, Powell SJ, Runswick MJ, Saraste M, Tybulewicz VL (1985) Primary structure and subunit stoichiometry of F1-ATPase from bovine mitochondria. J Mol Biol 184: 677–701.
Washizu M, Kurahashi Y, Iochi H, Kurosawa O, Aizawa S, Kudo S, Magariyama Y, Hotani H (1993) Dielectrophoretic measurement of bacterial motor characteristics. IEEE Trans Ind Appl 29: 286–294.
Wilkens S, Zhou J, Nakayama R, Dunn SD, Capaldi RA (2000) Localization of the delta subunit in the *Escherichia coli* F_1F_0-ATPsynthase by immuno electron microscopy: the delta subunit binds on top of the F_1. J Mol Biol 295: 387–391.
Yakushi T, Yang J, Fukuoka H, Homma M, Blair DF (2006) Roles of charged residues of rotor and stator in flagellar rotation: comparative study using H+-driven and Na+-driven motors in *Escherichia coli*. J Bacteriol 188: 1466–1472.
Yamaguchi S, Aizawa S, Kihara M, Isomura M, Jones CJ, Macnab RM (1986a) Genetic evidence for a switching and energy-transducing complex in the flagellar motor of *Salmonella typhimurium*. J Bacteriol 168: 1172–1179.
Yamaguchi S, Fujita H, Ishihara A, Aizawa S, Macnab RM (1986b) Subdivision of flagellar genes of *Salmonella typhimurium* into regions for assembly, rotation and switching. J Bacteriol 166: 187–193.
Yasuda R, Noji H, Kinosita K, Yoshida M (1998) F_1-ATPase is a highly efficient molecular motor that rotates with discrete 120 degree steps. Cell 93: 1117–1124.
Yasuda R, Noji H, Yoshida M, Kinosita K, Itoh H (2001) Resolution of distinct rotational substeps by submillisecond kinetic analysis of F_1-ATPase. Nature 410: 898–904.
Yasuda R, Masaike T, Adachi K, Noji H, Itoh H, Kinosita K (2003) The ATP waiting conformation of rotating F_1-ATPase revealed by single-pair fluorescence resonance energy transfer. Proc Natl Acad Sci USA 100: 9314–9318.

Yorimitsu T, Sowa Y, Ishijima A, Yakushi T, Homma M (2002) The systematic substitutions around the conserved charged residues of the cytoplasmic loop of Na+-driven flagellar motor component PomA. J Mol Biol 320: 403–413.

Yorimitsu T, Mimaki A, Yakushi T, Homma M (2003) The conserved charged residues of the C-terminal region of FliG, a rotor component of the Na+-driven flagellar motor. J Mol Biol 334: 567–583.

Yorimitsu T, Kojima M, Yakushi T, Homma M (2004) Multimeric structure of the PomA/PomB channel complex in the Na+-driven flagellar motor of *Vibrio alginolyticus*. J Biochem 135: 43–51.

York J, Spetzler D, Hornung T, Ishmukhametov R, Martin J, Frasch WD (2007) Abundance of *Escherichia coli* F_1-ATPase molecules observed to rotate via single-molecule microscopy with gold nanorod probes. J Bioenerg Biomemb 39: 435–439.

Yoshida M, Muneyuki E, Hisabori T (2001) ATP synthase–a marvellous rotary engine of the cell. Nat Mol Cell Biol 2:669–677.

Yuan J, Berg HC (2008) Resurrection of the flagellar rotary motor near zero load. Proc Natl Acad Sci USA 105: 1182–1185.

Zhang Y, Wang J, Cui Y, Yue J, Fang X (2005) Rotary torque produced by proton motive force in F_0F_1 motor. Biochem Biophys Res Commun 331: 370–374.

Zhao R, Pathak N, Jaffe H, Reese TS, Khan S (1996) FliN is a major structural protein of the C-ring in the *Salmonella typhimurium* flagellar basal body. J Mol Biol 261: 195–208.

Zhou J, Fazzio RT, Blair DF (1995) Membrane topology of the MotA protein of *Escherichia coli*. J Mol Biol 251: 237–242.

Zhou J, Blair DF (1997) Residues of the cytoplasmic domain of MotA essential for torque generation in the bacterial flagellar motor. J Mol Biol 273: 428–439.

Zhou J, Lloyd SA, Blair DF (1998a) Electrostatic interactions between rotor and stator in the bacterial flagellar motor. Proc Natl Acad Sci USA 95: 6436–6441.

Zhou J, Sharp LL, Tang HL, Lloyd SA, Billings S, Braun TF, Blair DF (1998b) Function of protonatable residues in the flagellar motor of *Escherichia coli*: a critical role for Asp32 of MotB. J Bacteriol 180: 2729–2735.

Zimmermann B, Diez M, Zarrabi N, Gräber P, Börsch M (2005) Movements of the epsilon-subunit during catalysis and activation in single membrane-bound H(+)-ATP synthase. EMBO J 24: 2053–2063.

8

Fluorescence Correlation Spectroscopy in Living Cells

Thomas Weidemann and Petra Schwille

Abstract In fluorescence correlation spectroscopy (FCS), information about molecular dynamics is extracted by recording the fluctuating signal of a pico- to micromolar concentration of fluorescent molecules in solution. As primary parameters, FCS provides time constants of the fluorescence emission, as well as numbers and dwell times of the observed particles diffusing through the open volume. A biochemical reaction or macromolecular rearrangement causing changes in these parameters, when linked to fluorescence readout, can be quantified by FCS. Since the measurements are now routinely performed in a laser-illuminated confocal spot, making measurements in living cells is straightforward. Different cellular compartments, such as the nucleus, the cytoplasm, or the plasma membrane, can be easily distinguished and addressed. In addition to biochemical reactions, the local environment of macromolecules, for example, ion concentrations, pH, or viscosity, can be probed. Thus, FCS is a versatile and attractive technique for researchers striving for a quantitative understanding of interactions and dynamics of biological and in particular cellular systems.

8.1. Introduction

During the last 15 years, fluorescence correlation spectroscopy (FCS) has matured from a concept to a fully commercially developed technique [1–3]. There is a growing community of researchers in the life sciences with access to FCS as part of extended microscopy approaches. This chapter is meant to serve as an introduction for those who are considering using FCS as a method for addressing the physicochemical properties of biological systems. In the next two sections we provide a general introduction to the method and develop key aspects of the theory. In Section 8.4 we discuss some instrumental details. Finally, in Section 8.5, we review some biologically motivated studies that may guide the interested reader to particular applications.

T. Weidemann and P. Schwille • BIOTEC, Biophysics, Technische Universität Dresden, Tatzberg 47-51, 01307 Dresden, Germany

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_8,

8.2. Measurement Principle

As in time-resolved fluorescence spectroscopy, in FCS one records the fluorescence signal over time, although from a very small microscopic volume in solution (Figure 8.1). For example, measuring a 100 nM solution of dye in a macroscopic cuvette (125 mm^3) with a conventional fluorospectrometer integrates the fluorescence emitted by approximately 10^{13} molecules. In FCS, the fluorescent particles are excited in a diffraction-limited laser spot (focus). A confocal pinhole is used to further reduce the detection volume. Typical dimensions are 0.4 μm perpendicular and 1.6 μm parallel to the optical axis, thus shaping a sub-femtoliter (fL) ellipsoid. A volume of 0.3 fL placed in a 100 nM solution of dye will be occupied by, on average, 20–30 particles. Due to their Brownian motion, the particles will randomly enter and exit the observation volume. Neglecting time-dependent instabilities in emission, these diffusive movements will still lead to signal fluctuations, which, in contrast to noise, are correlated. The kinds of information contained in these so-called number fluctuations are their average deviation from the mean, the amplitude, and their frequencies. The amplitude reflects linearly a reciprocal concentration, while the time domain contains information about the diffusion behaviors and thus the size and shape of the fluorescent particles in conjunction

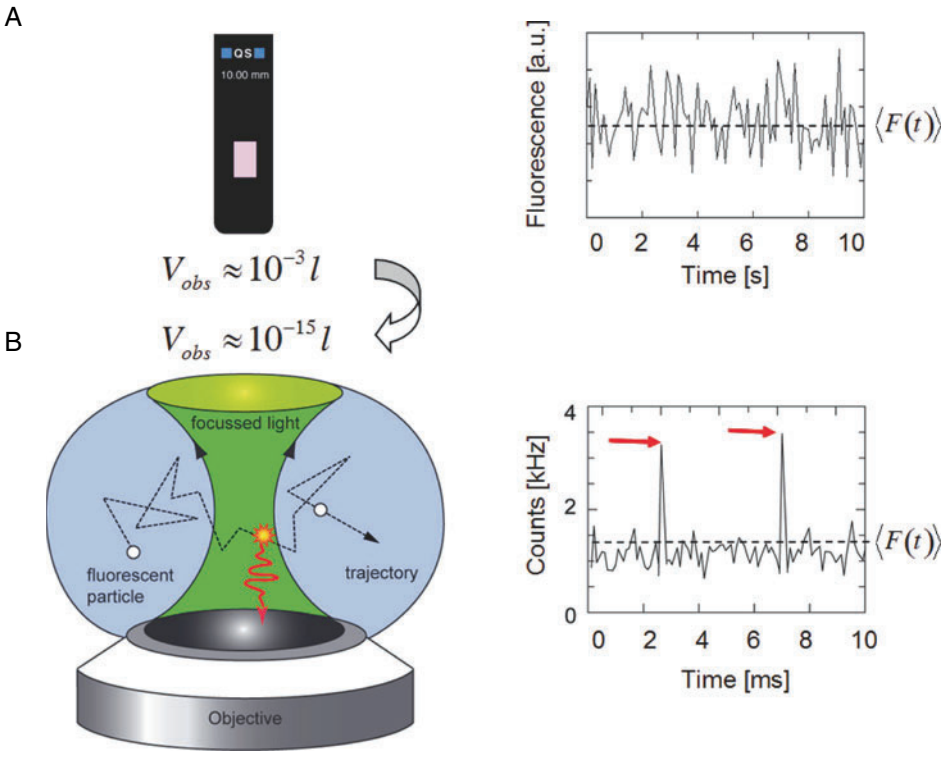


Figure 8.1. Measurement principle of fluorescence correlation spectroscopy (FCS). Downscaling of the observation volume V_{obs} from milliliters (cuvette-based) **(upper)** to femtoliters (FCS) **(lower)**. The fluorescence signal fluctuations in conventional bulk measurements indicate the noise of the system (detectors, stray light, etc.). In FCS, the transient residence of individual molecules gives rise to additional fluctuations. Individual peaks reflecting individual molecular dwellings become visually apparent at very low concentrations (*red arrows*).

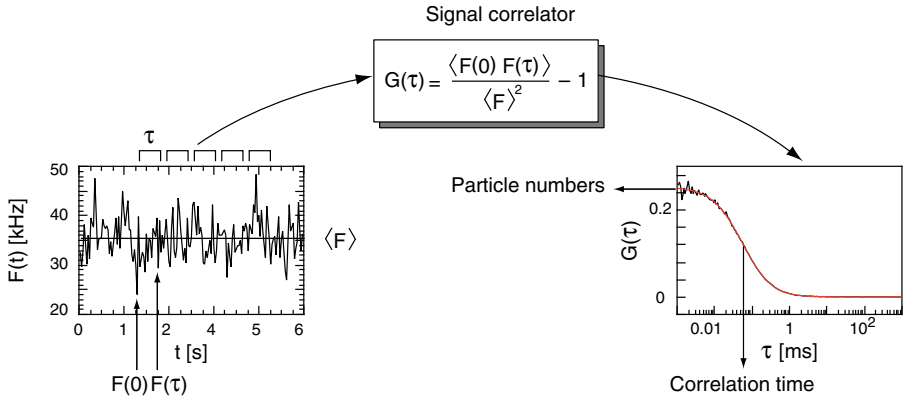


Figure 8.2. Autocorrelation analysis. Although individual peaks in signal fluctuations measured at more highly concentrated solutions cannot be resolved by the unaided eye, the intensity trace nevertheless contains information about concentration fluctuations of individual particles.

with the viscosity of the solution. Slow particles give rise to slower signal fluctuations and conversely.

A mathematical way of extracting these parameters is to perform an autocorrelation analysis (Figure 8.2). The signal is usually measured in photon counting mode with submicrosecond resolution, that is, binned in channels along the intensity trace (also called the "run"). For autocorrelation, the product of two signal channels separated by a certain lag-time is averaged over the whole trace, normalized by the square of the mean and plotted for increasing lag-times on a logarithmic scale [see Eq. (8.1)]. The resulting autocorrelation function (ACF)[1] is a sigmoidal decay with characteristic correlation times and an intercept larger than 1. This can be understood qualitatively by considering a very diluted solution. In such a case, individual peaks of the measured intensity trace correspond to the transits of individual particles. If the lag-time is much smaller than a typical dwell time, there will be many pairs of values that belong to the same molecular transit, and those products add up when averaged over the whole trace (see Figure 8.1B) to something larger than the square of the mean. However, with increasing lag-times the typical peak width will be exceeded, rendering the signal channels statistically independent, and thus their product will approach asymptotically the square of the mean. It is quite common to subtract the asymptotic offset 1 from the curves. Using this approach to presenting a display, one finds that the intercept directly equals the inverse occupation number for the given observation volume, while the decay approaches 0 for longer lag-times (Figure 8.2).

Autocorrelation curves, the "FCS raw data," are usually generated in real time by a hardware-based correlator board during data acquisition. The user has to evaluate the measured ACF with appropriate model functions. For example, fitting a model for translational diffusion returns up to three characteristic diffusion times and their relative contributions to the amplitude. Because the number of particles, as well as the diffusion times, depend critically on the shape and the size of the observation volume, the values have to be compared to an otherwise characterized standard. This subsequently transforms diffusion times into

[1] The acronym ACF can denote both the measured autocorrelated data and the model function that is used for fitting. We try to specify these according to the context.

diffusion coefficients and particle numbers into concentrations. It is important to stress that FCS in this respect is a comparative technique relying on proper calibration of the optical system.[2]

FCS is a "single-molecule" technique because it works at low concentrations and can resolve conventionally invisible molecular dynamics in the time range of microseconds to seconds. The average number of observed molecules ranges from less than one to several hundred, which is, compared to fractions of moles, a small number. However, FCS does not monitor processes at "individual" molecules as in single-molecule detection (SMD). The main difference between SMD and FCS is that in SMD the molecules under study are mostly immobilized (attached to a support) or slowly moving. Because an individual fluorophore can emit around 10^6 photons before it is irreversibly bleached, SMD of immobilized fluorophores is limited to observation times of less than 1 sec, and many events have to be compiled either in parallel or repeatedly. FCS usually works in solution, in which only a small fraction of fluorophores is excited. This fraction is in constant dynamic exchange with a large nonilluminated reservoir. FCS data are based on averaging hundreds to thousands of particle transits of the same kind. Although the constant renewal of the observed "individuals" prolongs the measurement time, this sets (quasi-) equilibrium conditions as an important prerequisite for FCS.

8.3. Theoretical Framework

Many excellent reviews have been published about FCS theory [1,2,4–8]. Here we compile some of the essential results to help the biochemically motivated reader to understand the underlying approach and to facilitate orientation among the different applications.

8.3.1. Diffusion

As stated in the previous subsection, during the measurement the normalized second-order autocorrelation function (ACF), $G(\tau)$, is generated from the fluorescence signal $F(t)$ by computing[3]

$$G(\tau) = \frac{\langle F(t) \cdot F(t+\tau)\rangle}{\langle F\rangle^2} - 1 = \frac{\langle \delta F(t) \cdot \delta F(t+\tau)\rangle}{\langle F\rangle^2} \quad *(8.1)$$

The second equation can be verified by expressing the signal as temporal deviations from the mean $F(t) = \langle F(t)\rangle + \delta F$ and using the fact that $\langle \langle F\rangle \delta F(t)\rangle = 0$. Angular brackets indicate time averaging of the respective observable O over all channels with binning times t_i,

$$\langle O(t)\rangle = \frac{1}{\sum t_i} \sum_{i=0}^{N} O(t_i) \quad (8.2)$$

$G(\tau)$ is an empirically measured, dimensionless function that contains no information about the underlying physical nature of the signal fluctuations. Therefore, we need an analytical expression for $G(\tau)$ in which the expected parameters are modeled and implemented. For the FCS user it is important to appreciate that any process leading to signal fluctuations will show

[2] The issue of calibrating diffusion times has been recently overcome by scanning FCS (S-FCS).

[3] Relations that are important for the application-oriented user are indicated by an asterisk.

up in the autocorrelated curve with its characteristic correlation times. Accordingly, FCS has been used to monitor sample flow, active transport, chemical reactions, and, of course, photophysical processes such as triplet blinking or reversible quenching of the chromophore due to environmental conditions (e.g., protonation). Although some of these processes are touched on later in the text, we first derive an expression for diffusion-driven number fluctuations because they are most widely studied in biochemistry and cell biology.

The fluorescence signal at the detector is composed of instrumental features and molecular properties of the fluorophore,

$$F(t) = \int W(\mathbf{r}) q c(\mathbf{r}) dV \tag{8.3}$$

where $W(\mathbf{r}) = S(\mathbf{r})I(\mathbf{r})/I_0$ describes the effective shape of the combined spatial sensitivity $S(\mathbf{r})$ and dimensionless excitation intensity profile $I(\mathbf{r})/I_0$ with the maximum intensity I_0 at the center.[4] Here $c(\mathbf{r})$ is the concentration and q the molecular brightness of the fluorophores. The brightness $q \propto I_0 \sigma_{\text{exc}} \phi$ depends linearly on the peak intensity I_0, the absorption cross section of the molecules σ_{exc}, and their quantum yield ϕ.

Let us consider a single fluorophore at position $\mathbf{r_1}$ traveling to some other position $\mathbf{r_2}$ during a lag-time τ. The conditional probability for the molecule to be found in an arbitrary volume element $d^3\mathbf{r}_2$ can be calculated with the so-called diffusion propagator $P_D(\mathbf{r}_2|\mathbf{r}_1,\tau)$, a solution to Fick's diffusion equation

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = D_t \nabla^2 c(\mathbf{r},t) \tag{8.4}$$

with the boundary conditions $c(\mathbf{r}_1,0) = \delta(\mathbf{r}_1)$:

$$P_D(\mathbf{r}_2|\mathbf{r}_1,\tau) = \frac{1}{(4\pi D_t \tau)^{1/2}} \exp\left[-\frac{|\mathbf{r}_1 - \mathbf{r}_2|^2}{4D_t\tau}\right] \tag{8.5}$$

Here, D_t is the translational diffusion coefficient and $\delta(\mathbf{r}_1)$ denotes the Dirac delta function. The probability of finding the fluorophore in volume element $d^3\mathbf{r}_2$ is then given by $P_D(\mathbf{r}_2|\mathbf{r}_1,\tau)\, d^3\mathbf{r}_2$, whereas the probability of finding it initially in $d^3\mathbf{r}_1$ is simply $d^3\mathbf{r}_1/V$, with V the total sample volume. The average signal collected from a single fluorophore is proportional to the integral of the molecular detection efficiency

$$\langle f(t) \rangle \propto \frac{1}{V} \int W(\mathbf{r}) dV \tag{8.6}$$

and the average product of two signals separated by a discrete diffusion step within the lag-time τ is

$$\langle f(t) f(t+\tau) \rangle \propto \frac{1}{V} \iint W(\mathbf{r}_1) P_D(\mathbf{r}_2|\mathbf{r}_1,\tau) W(\mathbf{r}_2) dV dV' \tag{8.7}$$

[4] Vectors are written in boldface type.

The overall fluorescence signal is the sum of the total number N of particles at their various positions:

$$F(t) = \sum_{n=1}^{N} f(\mathbf{r}_n(t)) \tag{8.8}$$

Substituting Eq. (8.8) into the first part of Eq. (8.1) and evaluating the sums shows that the ACF can be ascribed to the autocorrelation functions of single individual molecules, $g(\tau)$:

$$\begin{aligned} G(\tau) &= \frac{\langle F(t) \cdot F(t+\tau)\rangle}{\langle F\rangle^2} - 1 \\ &= \frac{N\langle f(t) f(t+\tau)\rangle + \left(N^2 - N\right)\langle f\rangle^2}{N^2\langle f\rangle^2} - 1 \\ &= \frac{\langle f(t) f(t+\tau)\rangle}{N\langle f\rangle^2} + 1 + \underbrace{\frac{1}{N}}_{\approx 0} - 1 \\ &= \frac{1}{N} g(\tau) \end{aligned} \tag{8.9}$$

The fact that the ACF is a sum from individual molecular contributions is the basis for the linear concentration dependence, as well as for analyzing the measured ACF with linear combinations of different models, for example, subspecies with different diffusion times [see Eq. (8.22)]. Introducing Eqs. (8.6) and (8.7) into Eq. (8.9) relates the measured ACF of a single type (q constant for all particles) of N fluorophores in a sample volume V to actual experimental conditions:

$$G(\tau) = \frac{V}{N} \frac{\iint W(\mathbf{r}_1) P_D(\mathbf{r}_2|\mathbf{r}_1,\tau) W(\mathbf{r}_2) dV dV'}{\left(\int W(\mathbf{r}_2)\, dV\right)^2} \tag{8.10}$$

Because dual-color applications are of utmost importance for bioapplications, we extend the formalism to cover the use of a second detection channel. This is simply done by discriminating the two correlated signals $F_i(t)$ and $F_j(\mathrm{t}+\tau)$, as well as the corresponding detection functions $W_i(\mathbf{r}_1)$ and $W_j(\mathbf{r}_2)$, with an index for the simultaneous treatment of autocorrelation ($i = j \leftrightarrow k$) and cross-correlation ($i \neq j \leftrightarrow \mathrm{x}$).

For conventional confocal microscopes with overfilled back apertures of the objective, a three-dimensional (3D) Gaussian distribution of light intensity is considered to be a sufficiently accurate approximation [9,10]:

$$W_i(\mathbf{r}) = \exp\left[-2\frac{(x-p_{xi})^2 + \left(y-p_{yi}\right)^2}{\omega_{0i}^2}\right] \exp\left[-2\frac{(z-p_{zi})^2}{z_{0i}^2}\right] \tag{8.11}$$

with the vector $\boldsymbol{p}$ pointing from an arbitrary origin to the center of the detection volume of each channel (Figure 8.3). The exponentials in Eq. (8.11) can be factorized and sorted according to x, y, and z, reducing the problem to one dimension:

$$G_{ij}^{(x)}(\tau) = \frac{1}{\sqrt{4\pi D_t \tau}} \frac{\iint \exp\left[-2\frac{(x_1-p_{xi})^2}{\omega_{0i}^2} - 2\frac{\left(x_2-p_{xj}\right)^2}{\omega_{0j}^2}\frac{(x_1-x_2)^2}{4D_t\tau}\right] dx_1 dx_2}{\int \exp\left[-2\frac{(x-p_{xi})^2}{\omega_{0i}^2}\right] dx \int \exp\left[-2\frac{\left(x-p_{xj}\right)^2}{\omega_{0j}^2}\right] dx} \tag{8.12}$$

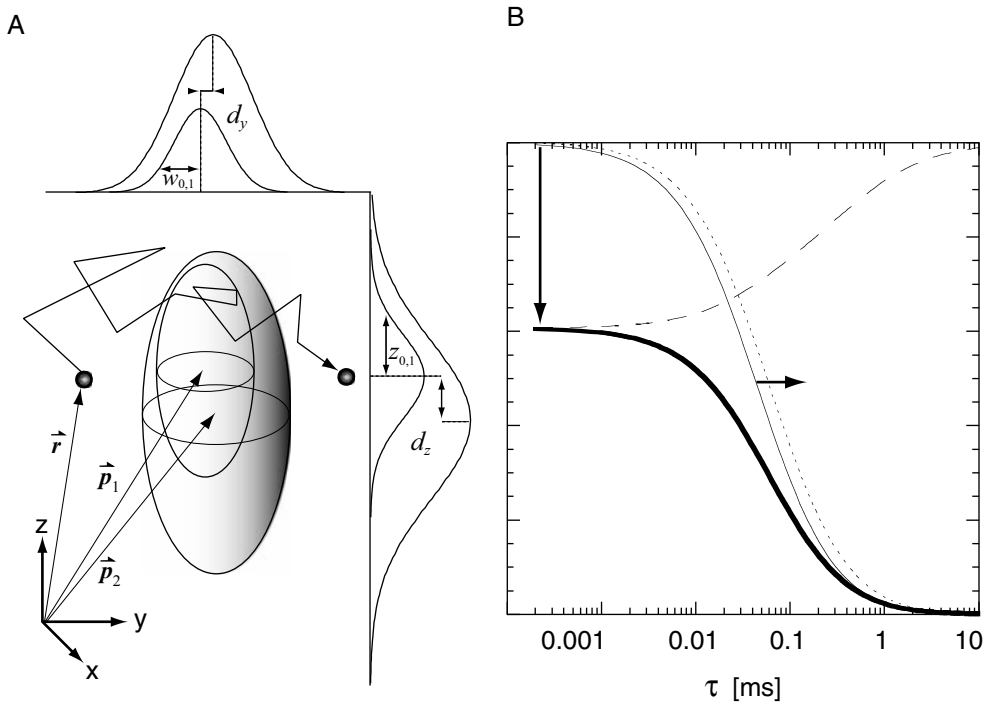


Figure 8.3. Modeling the detection volume [7]. **A.** Approximation by three-dimensional Gaussian intensity distributions with $1/e^2$-radius w_0 transverse and z_0 parallel to the optical axis z. The foci of a two-color setup centered at positions $\boldsymbol{p}_1$ for the green and $\boldsymbol{p}_2$ for the red excitation wavelengths differ in size and can show chromatic displacement $\boldsymbol{d}$. **B.** Cross-correlation functions in a two-color setup. A typical correlation curve (faint) corresponding to a ~10 nM solution of two-colored standard was computed for $w_0 = 220$ nm and $z_0 = 660$ nm. Chromatic displacement of $d_{yx} = 100$ nm, $d_z = 330$ nm leads to a time-dependent correction factor (*broken line*). The amplitude of the biased cross-correlation function (*bold line*) appears reduced (*vertical arrows*), and the function is also shown normalized (*dotted line*) to visualize an increase of an apparent correlation time (*horizontal arrow*).

Exploiting the infinite limit of the Gauss integral

$$\int_{-\infty}^{+\infty} \exp\left[-a^2x^2\right] = \sqrt{\frac{\pi}{a}} \tag{8.13}$$

allows us to reduce the denominator and to extract a constant factor

$$G_{ij}^{(x)}(\tau) = \frac{1}{\sqrt{4\pi D_t \tau}} \frac{2}{\pi \omega_{0i}\omega_{0j}} \iint \exp\left[-2\frac{(x_1 - p_{xi})^2}{\omega_{0i}^2} - 2\frac{\left(x_2 - p_{xj}\right)^2}{\omega_{0j}^2} \frac{(x_1 - x_2)^2}{4D_t\tau}\right] dx_1 dx_2 \tag{8.14}$$

Equation (8.14) is a crest-shaped surface integral over the two independent variables x_1 and x_2, which can be solved analytically by (1) translation of the peak at (p_{xi}, p_{xj}) to the point

of origin, followed (2) by rescaling and (3) rotation of the axes by $\pi/4$ with the following substitutions:

$$\left\{\begin{array}{l} u_1 = x_1 - p_{xi} \\ u_2 = x_2 - p_{xj} \\ d_x = p_{xi} - p_{xj} \end{array}\right\}, \left\{\begin{array}{l} v_1 = \dfrac{u_1}{\omega_{0i}} \\ v_2 = \dfrac{u_2}{\omega_{0j}} \end{array}\right\}, \text{ and } \left\{\begin{array}{l} \beta = \dfrac{1}{\sqrt{2}\pi x_{0eff}} \left(x_{0i} v_1 - x_{0j} v_2\right) \\ \alpha = \dfrac{1}{\sqrt{2}\pi x_{0eff}} \left(x_{0j} v_1 - x_{0i} v_2\right) \\ \alpha^2 + \beta^2 = v_1^2 + v_2^2 \end{array}\right\} \tag{8.15}$$

with $x_{0eff} = (x_{0i} + x_{0j})/2$ the $1/e^2$ radius of the Gaussian profile in x. Equation (8.14) then transforms into

$$G_{ij}^{(x)}(\tau) = \frac{1}{\sqrt{4\pi D_t \tau}} \frac{2}{\pi} \iint \exp\left[-2(\alpha^2 + \beta^2) - \frac{\left(\sqrt{2} x_{0,eff}\beta + d_x\right)^2}{4 D_t \tau}\right] d\alpha d\beta \tag{8.16}$$

and with use of Eq. (8.13) further into

$$G_{ij}^{(x)}(\tau) = \frac{1}{\sqrt{4\pi D_t \tau}} \frac{2}{\pi} \sqrt{\frac{\pi}{2}} \int \exp\left[-2\beta^2 - \frac{\left(\sqrt{2} x_{0,eff}\beta + d_x\right)^2}{4 D_t \tau}\right] d\beta \tag{8.17}$$

Recourse to a quadratic extension

$$\int_{-\infty}^{+\infty} \exp\left[-2\frac{x^2}{a^2} - 2\frac{(x - x_0)^2}{b^2}\right] dx = \exp\left[-2\frac{x_0^2}{a^2 + b^2}\sqrt{\frac{\pi}{2}}\frac{ab}{\sqrt{a^2 + b^2}}\right] \tag{8.18}$$

finally returns an explicit solution for Eq. (8.12):

$$G_{ij}^{(x)}(\tau) = \frac{1}{\sqrt{\pi} x_{0,eff}} \left(1 + \frac{4 D_t \tau}{x_{0,eff}^2}\right)^{-\frac{1}{2}} \exp\left[-2\frac{d_x^2}{4 D_t \tau + x_{0,eff}^2}\right] \tag{8.19}$$

where $\mathbf{d} = \mathbf{p}_i - \mathbf{p}_j$ is the displacement vector connecting the two centers of the foci of the two detection volumes. Analogous treatment of the axes y and z allows us to write the complete solution for an analytical ACF assuming diffusion-driven fluctuations through a three-dimensional Gaussian detection volume in two different detection channels (i, j):

$$G_{ij}^{(D)}(\tau) = \frac{1}{c V_{eff}} \underbrace{\left(1 + \frac{4 D_t \tau}{\omega_{0,eff}^2}\right)^{-1}}_{\text{2D}} \underbrace{\left(1 + \frac{4 D_t \tau}{z_{0,eff}^2}\right)^{-\frac{1}{2}}}_{\text{3D}} \underbrace{\exp\left[-\frac{d_x^2 + d_y^2}{4 D_t \tau + \omega_{0,eff}^2} - \frac{d_z^2}{4 D_t \tau + z_{0,eff}^2}\right]}_{\text{overlap}} \tag{8.20}$$

In the last step, $V_{eff} = \pi^{3/2} \omega_{0i} \omega_{0j} \sqrt{z_{0i} z_{0j}}$ combines all of the constant prefactors. Because the number fluctuations obey Poissonian statistics, it follows that the intercept of the ACF is the inverse average occupation number in the detection volume, $G(0) = 1/\langle N_{ob} \rangle$; consequently, V_{eff} is regarded as the effective detection volume. Apart from a factor, this is the volume of a 3D ellipsoid, $V_{\text{ellipsoid}} = 4/3\pi\omega_0^2 z_0$, with half-axes ω_0 and z_0, as usually sketched in

illustrations (Figure 8.3). The multiplied terms in Eq. (8.20) describe different aspects of the geometry. The "2D" (two-dimensional) factor accounts for diffusion across the finite borders in *y*-*x*, and the "3D" factor accounts for the confined focus along the optical axis in *z*. Measuring fluorescent particles trapped in horizontal 2D systems, like membranes, will not give rise to fluctuations along *z*. In those cases only, the 2D factor is used for fitting (infinite z_0 sets the 3D factor to 1). The exponential factor in Eq. (8.20) accounts for the imperfect overlap of the two detection volumes, which vanishes for autocorrelation, where $\mathbf{d} = \mathbf{0}$. Recently, two-foci setups were intentionally used to intrinsically calibrate measured diffusion times [11]. Imperfect overlap is a significant issue when combining two different laser sources into one objective, as for dual-color fluorescence cross-correlation (FCCS) (Figure 8.3B). It is interesting to note that cross-correlation between two displaced detection volumes is formally equivalent to autocorrelation in solutions with directed flow. In the latter context, the exponential factor was already presented in the classic papers of Magde, Elson, and Webb [12–14].

Introducing the parameters "diffusion time" $\tau_D = \omega_0^2 / 4D$ and the "axis ratio" $\kappa = z_0 / \omega_0$, which for well-aligned conventional confocal setups is a number between 3 and 6, we obtain the widely used ACF formula for one diffusion species in one channel:

$$G_k^{(D)}(\tau) = \frac{1}{cV_{eff}} \left(1 + \frac{\tau}{\tau_D}\right)^{-1} \left(1 + \kappa^2 \frac{\tau}{\tau_D}\right)^{-\frac{1}{2}} \qquad *(8.21)$$

Equation (8.21) shows that the axis ratio is returned by the fit, although usually with a weak significance. Moreover, the diffusion time and the axis ratio are tightly coupled. The common procedure is therefore to determine the axis ratio with a standard—for example, an organic dye—and to keep it constant during the fitting routines. Up to three different diffusion species can be reliably discriminated [15] by fitting

$$G_k^{(D)}(\tau) = \sum\nolimits_s G_k^{(D,s)}(\tau) = \frac{1}{cV_{eff}} \sum_{s=1}^{3} \rho_s \left(1 + \frac{\tau}{\tau_{D,s}}\right)^{-1} \left(1 + \kappa^2 \frac{\tau}{\tau_{D,s}}\right)^{-\frac{1}{2}} \qquad *(8.22)$$

and $\rho_s = c_s / \sum_s c_s$ denotes their relative contributions to the amplitude $G(0)$.

8.3.2. Chromophore Dynamics

Fluorescence emission is governed by dynamic processes at the molecular level that contribute to correlation curves [3]. Organic dyes in the first excited state are prone to intersystem crossing from the singlet into the triplet system of quantum states (energy levels occupied by two electrons with parallel spins). The lifetime of the triplet state is in the lower microsecond range, and it depopulates in solution in almost a radiationless way. The effect of the triplet on the shape of the ACF was reported by Widengren et al. [16]:

$$G_k(\tau) = G_k^T(\tau)\, G_k^D(\tau) = \left(1 + \frac{T}{(1-T)} \exp\left[-\frac{\tau}{\tau_T}\right]\right) G_k^D(\tau) \qquad *(8.23)$$

with the triplet fraction $T = 1 - G_k^D(0) / G_k(0)$ and a correlation time for triplet blinking τ_T. The more fluorophores are trapped in the triplet state, the greater is the increase of the amplitude compared to pure number fluctuations. Because the triplet is populated from the first excited state, it depends critically on the excitation power. Inappropriate photon statistics

in the lower-microsecond regime, combined with a trailing tail of detector afterpulsing, can add uncertainty in fitting the triplet amplitude and may affect concentration measurements.

Another important blinking phenomenon occurs in fluorescent proteins. Fluorescent proteins have had an enormous effect in cell biology and also play a crucial role in intracellular FCS applications. The photophysics of green-fluorescent protein (GFP) mutants was found to be surprisingly complex in ways that can be both helpful and problematic in biological applications. A number of detailed FCS studies were reported for GFP mutants from *Aequorea victoria*, as well as for red-fluorescent proteins (RFPs) from other organisms and demonstrate impressively the analytic power of this method [17–23]. The blinking behavior is mainly linked to pH dependence of the fluorescence quantum yield. The crystal structure of GFP is shown in Figure 8.4. The polypeptide forms a β-sheet barrel that mantles a covalent mesomeric π-electron system [24]. The chromophore of wild-type (wt) GFP is formed autocatalytically in the center of the protein by the side chains of Ser65, Tyr66, and Gly67 [25]. The substitutions F64L and S65T turn wtGFP into enhanced GFP (eGFP), by far most the widely used and spectrally superior variant [26]. For eGFP it was shown that Tyr66 acts as a weak acid (pKa = 5.8) [18]. Low pH and subsequent protonation of Tyr66 mostly abolishes

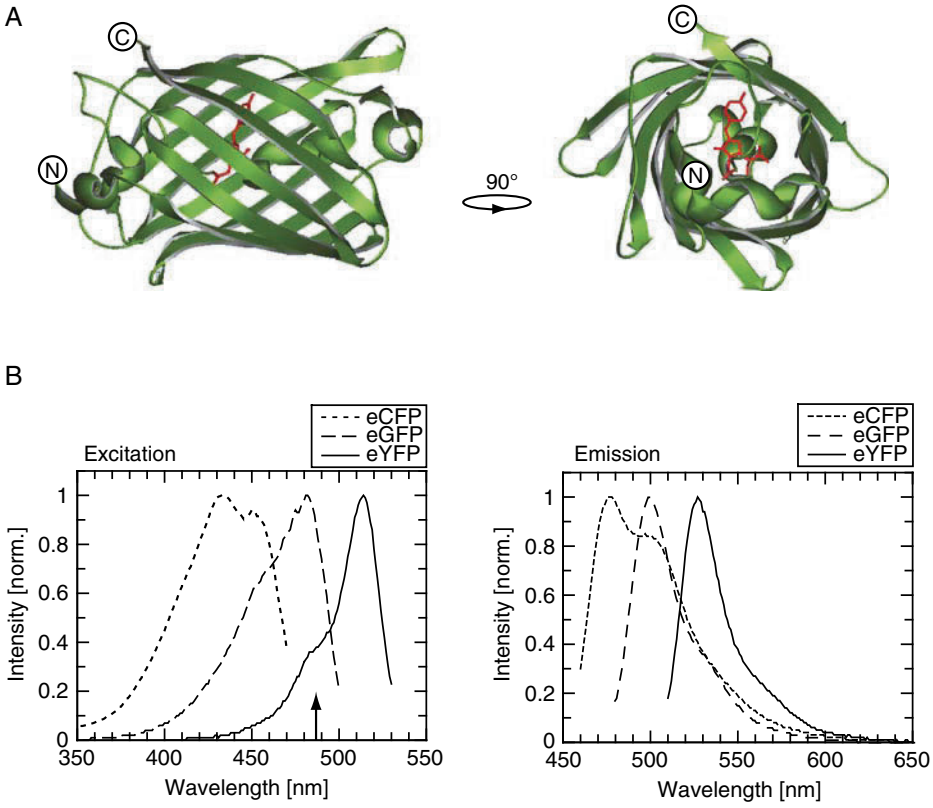


Figure 8.4. Green-fluorescent protein (GFP). **A.** Crystal structure of wild-type GFP (Protein Data Base code 1ema [24]) displayed in ribbon style. Atoms and bonds forming the chromophore are shown in red; the N- and C-termini are highlighted. **B.** Excitation and emission spectra of enhanced cyan-, green-, and yellow-fluorescent proteins (eCFP, eGFP, and eYFP, respectively) measured in lysed HEK293 cells diluted in phosphate-buffered saline (measured on a Spex Fluorolog). The common argon laser excitation line 488 nm is indicated by an arrow.

fluorescence emission. Processes on two different time scales were differentiated: dynamic exchange of protons with the surrounding solution, and proton rearrangements within the bonding network of the protein core. The latter was found to be excitation driven, especially in red-shifted variants like enhanced yellow-fluorescent proteins (YFPs) [23]. While the contribution of external protonation for eGFP can make up a significant fraction of the amplitude (e.g., ~80% at pH 5), internal protonation is constant at about 13%. The correlation time for external protonation is not coupled to diffusion and is quite stable with respect to excitation power and ionic strength. Besides their pH dependence, however, the correlation times (50–400 μs) decrease with increasing molarity of the phosphate buffers, suggesting that subtle conformational changes modulate the accessibility of the chromophore [17,18,27]. In analogy to the case for the triplet, Eq. (8.22) is frequently used to evaluate intracellular correlation curves of GFP constructs, thus averaging out the protonation dynamics with a single dark state. However, for biochemical in-solution studies, it may be necessary to include additional correlation times [23] by fitting with

$$G_k(\tau) = \left(1 + \frac{F_1}{1-F_1}\exp\left[-\frac{\tau}{\tau_1}\right]\right)\left(1 + \frac{F_2}{1-F_2}\exp\left[-\frac{\tau}{\tau_2}\right]\right) G_k^D(\tau) \tag{8.24}$$

Here F_1, τ_1 and F_2, τ_2 denote the correlation times and fractions of particles in the two respective dark states. Because GFP flickering due on external protonation is independent of the instrumental conditions and the size of the GFP construct, it was suggested to use it as a readout to calibrate pH and temperatures from cells to microfabricated systems [18,28].

8.3.3. Concentrations

FCS was quite often promoted for its potential for determining concentrations in the pico- to micromolar concentration range via Eq. (8.23) for $\tau = 0$:

$$c = \frac{1}{(1-T)}\frac{1}{G_k^D(0)\,V_{eff}} \tag{*8.25}$$

However, this ability seems to have been less enthusiastically followed up in practice. Three reasons may account for this: First, the open detection volume is weakly defined. Second, it is not always feasible to perform reliable dilution rows from stock solutions down to the nanomolar range. Dyes and labeled proteins tend to stick away, and even small losses at vial surfaces and pipette tips can result in large errors. Thus, the amplitude is hard to calibrate experimentally by a traditionally calibrated standard. The third and most severe limitation is that the linear reciprocal concentration dependency is only valid if the fluorescent particles are homogeneously fluorescent. Random multiple labeling of macromolecules, bimolecular associations with quenching effects, trace amounts of aggregates, or nonconjugated dye may lead to mixtures of particles having differences in their molecular brightness q. Such mixtures of different species s lead to an increased fluctuation amplitude $G(0)$ compared to the same particle concentration of a homogeneously fluorescent ensemble. This can be quantitatively expressed from

$$G_k^{(D)}(\tau) = \frac{1}{cV_{eff}}\frac{\sum_s c_s q_s^2 G_k^{(D,s)}}{\left(\sum_s c_s q_s\right)^2} \tag{8.26}$$

by summing up the weighted contributions of the individual subensembles [29]. Note that the total number of diffusing complexes in the detection volume

$$\langle N_{ob}\rangle = cV_{eff} = V_{eff}\sum\nolimits_s c_s = \frac{1}{G_k^{(D)}(0)}\frac{\left(\sum_s c_s q_s\right)^2}{\sum_s c_s q_s} \tag{8.27}$$

may significantly deviate from $1/G_k^{(D)}(0)$. Moreover, the amplitude fractions ρ_s as derived by fitting Eq. (8.22) will be biased in favor of the brighter particles and do not reflect the true molar fractions

$$\rho_s^{cor} = \frac{c_s}{\sum_s c_s} = \frac{\rho_s/q_s^2}{\sum_s \rho_s/q_s^2} \tag{8.28}^*$$

Nevertheless, concentration measurements still have buried potential for FCS in single cells, where it can provide valuable information about the concentration range and distribution of expressed GFP-tagged proteins. The beta-can of fluorescent proteins provides effective chemical shielding of the chromophore, which might make it possible to study oligomerization via the molecular brightness q. In addition, the chaperone machinery inside the cell keeps the constructs in their native state, thereby preventing unspecific aggregations. Furthermore, interactions with compartment boundaries like membranes are highly regulated or minimized. On the other hand, cells can contain high concentrations of weakly fluorescent metabolic intermediates such as, for example, the flavin derivatives. These compounds build up a background of uncorrelated signal, which affects the linear reciprocal relationship. Considering the total signal F_{tot} to be composed of correlated fluctuations F and uncorrelated background B, we can rewrite Eq. (8.1) as

$$\begin{aligned} G(\tau) &= \frac{\langle \delta(F(t)+B(t))\,\delta(F(t+\tau)+B(t+\tau))\rangle}{\langle F+B\rangle^2} \\ &= \frac{\langle \delta F(t)\delta F(t+\tau)\rangle + \overbrace{\langle \delta F(t)\delta B(t+\tau)\rangle + \langle \delta B(t)\delta F(t+\tau)\rangle + \langle \delta B(t)\delta B(t+\tau)\rangle}^{=0}}{\langle F\rangle^2\left(1+2\frac{\langle B\rangle}{\langle F\rangle}+\frac{\langle B\rangle^2}{\langle F\rangle^2}\right)} \\ &= \frac{\langle \delta F(t)\delta F(t+\tau)\rangle}{\langle F\rangle^2}\left(1+\frac{\langle B\rangle}{\langle F\rangle}\right)^{-2} = G_k^D(\tau)\left(1+\frac{\langle B\rangle}{\langle F\rangle}\right)^{-2} = G_k^D(\tau)\left(\frac{\langle F\rangle}{\langle F\rangle+\langle B\rangle}\right)^2 \end{aligned} \tag{8.29}^*$$

which is the ACF for number fluctuations multiplied by a correction factor. The background issue was discussed in an early FCS paper by Koppel [30]. Note that in contrast to the previous situation, $F = F_{tot} - B$ now denotes the background-corrected intensity leading to number fluctuations. Evaluating the correction factor reveals that uncorrelated background at the detector with $F/B > 10$ reduces the amplitude, and hence increases the concentration values, less than 10%.

8.3.4. Interactions

The discovery and characterization of interactions are of major interest in cell biology, biochemistry, and biophysics. With FCS both the amplitude and the diffusion times

can be used to detect binding events of fluorescently labeled biomolecules. For example, binding of a small fluorescent ligand to a large nonfluorescent receptor may be detected by its larger diffusion time. Evaluating the molar fractions of bound and free ligand from Eq. (8.22) with a two-component model and plotting the fractions against varying receptor concentrations yields a binding isotherm, which can be analyzed with regard to affinity, cooperativity, and stoichiometry. Assuming the ligand and the complex to be globular shaped spheres, the increase in diffusion time can be estimated using the Stokes–Einstein relation

$$D_t = \frac{kT}{6\pi\eta R} \tag{8.30}$$

where k is the Boltzmann constant, T the absolute temperature, η the viscosity, and R the hydrodynamic radius of the particles. With homogeneous mass distribution within the spheres, $MW \propto \sqrt[3]{R}$, this implies a weak dependence between diffusion time and the size of the particles. Thus, a simple one-to-one binding reaction can hardly be resolved. In contrast, the amplitude is doubled on saturated one-to-one binding, and therefore there is a more sensitive readout. However, a large error due to scattering triplet amplitudes causes some difficulty in analyzing binding with autocorrelation amplitudes. An elegant version for the detection of interactions is dual-color fluorescence cross-correlation analysis,) in which the signal of one channel (e.g., $i \rightarrow g$, green) is correlated with the signal of second, spectrally well-separated color channel (e.g., $j \rightarrow r$, red) [31,32]. Here, the triplet dynamics, uncorrelated between different dyes, disappears and the amplitude builds up from the number fluctuations of particles simultaneously detected in both color channels. A convenient way of calibrating cross-correlation amplitudes is to divide them with the corresponding autocorrelation amplitude:

$$\frac{G_\times(0)}{G_i(0)} = \frac{V_{eff,i}}{V_{eff,\times}} \frac{c_0^{(g)}}{c_0^{(r)}} \frac{\sum_s c^{(s)} n_g^{(s)} n_r^{(s)}}{\sum_s c^{(s)} n_g^{(s)2}} \tag{8.31}$$

with n_g and n_r counting the number of green and red fluorophores, respectively, coupled to the diffusing complex of a certain species s. For a simple one-to-one binding (n_g and $n_r = 1$) the ratio of amplitudes is directly proportional to the fraction of bound ligands:

$$\frac{G_\times(0)}{G_i(0)} = \frac{V_{eff,i}}{V_{eff,\times}} \frac{c^{(gr)}}{c_0^{(r)}} \tag{8.32}$$

Note that calibration with the green channel delivers the fraction of complexes with respect to the total concentration of red fluorophores and vice versa. The ratio is further biased by the color-specific detection volumes $V_{eff,k}$ and the effective volume for cross-correlation, $V_{eff,\times}$, which is very sensitive to the chromatic shift [see Eq. (8.20)]. In a first approximation, the volumes of the two color channels scale with the cubic ratio of the excitation wavelength, $V_{eff,j} = (\lambda_{ex,j}/\lambda_{ex,i})^3 V_{eff,i}$; however, the effective cross-correlation detection volume requires experimental controls. For quantitative FCCS, it is advisable to determine both the upper and lower limits with well-defined standards, such as a mixture of noninteracting reactants for minimum and double-labeled particles for a maximum value, and to fit the binding data progressively in between. Various cases of different binding modes and stoichiometries can be found in the literature [7,33].

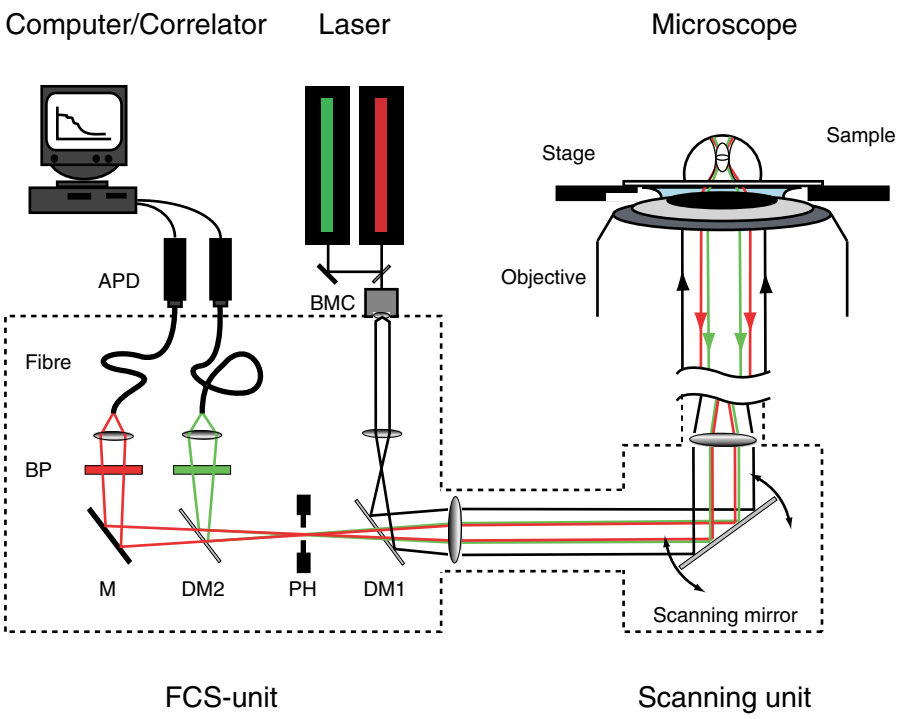


Figure 8.5. Typical setup of a confocal laser scanning microscope upgraded by a fluorescence correlation spectroscopy (FCS) module. Laser light of different color is supplied by a beam combiner (BMC) and reflected by a dichroic mirror (DM1) onto the back aperture of the objective. The fluorescence light is collected by the same optics, passes DM1 and the pinhole (PH), and is separated by DM2 into two different color channels terminated by bandpass filters (BP) and fibre-coupled avalanche photon-counting detectors (APDs). The hardware correlator card (not shown) is usually part of the computer. Note that in this configuration a scanning mirror allows both confocal imaging and FCS using the same optical path.

The application range for binding studies is defined by the accessible concentration range in FCS. Especially for FCCS the range is limited because both of the binding partners are labeled. Assuming that titrations should span at least one order of magnitude in concentration values centered on the dissociation constant K_d, one can derive convenient affinities between 1 and 100 nM [34]. With one-color applications affinities up to 10 μM can be reliably determined.[5]

8.4. Instrumentation

FCS developed its full potential through the use of (fiber-coupled) avalanche photodiodes (APDs) as single-photon counting detectors in combination with high–numerical aperture objectives in a confocal setup. The main features are sketched in Figure 8.5. Laser light is reflected by a dichroic mirror (DM1) onto the back aperture of the objective and focused in solution. Usually water immersion objectives are used. The red-shifted fluorescence is collected by the same optics, passes the first dichroic mirror, and is fed by a second dichroic

[5] M. Hintersteiner personal communication.

mirror (DM2) into two subsequent color channels. Between DM2 and the detectors, projecting lenses direct the beam through confocally positioned pinholes. Thus, the pinholes do not compress the illuminated volume but reject out-of-focus fluorescence along the optical axis z, which is crucial to reaching single-molecule sensitivity. APDs have a photon detection efficiency of ~50% for green light (500–600 nm) and are usually more sensitive at longer wavelengths. The time resolution of APDs is limited compared to that of photomultipliers, although resolution down to 50 ps is now available (e.g., the MPD series from PicoQuant, Berlin, Germany). The stream of photon counts is usually binned with 200-ns time intervals and correlated in real time[6] by a multiple-tau correlator card (e.g., ALV, Langen, Germany). A number of commercial FCS setups are on the market integrated into state-of-the-art confocal microscope systems. These include the Confocor 2 and 3 (Zeiss, Jena, Germany), FCS 2 (Leica Microsystems, Wetzlar, Germany), Micro Time 200 (PicoQuant), and Alba FCS (ISS, Urbana-Champaign, Illinois). With Clarina II, Evotec Technologies integrated the FCS$^+$ system into a fully automated liquid handling system for homogeneous high-throughput assays for drug discovery (Perkin Elmer, Hamburg, Germany).

8.4.1. Position-Sensitive Single-Point FCS

Because a confocal spot is about 10^3–10^4 times smaller than a eukaryotic cellular volume, the idea of performing FCS with subcellular resolution was straightforward (Figure 8.6) [35,36]. An important technical detail was to combine confocal imaging and FCS in the same optical path, thus providing a user-friendly platform for park-and-probe operation [37,38]. Combining spectroscopy and imaging in one beam line (Figure 8.5) not only increased the positioning accuracy in a complicated cellular environment, it also opened up the convenient possibility of transforming intensities from confocal cross sections into absolute particle numbers [37,39,40]. The latter procedure requires two steps: first, to determine the molecular brightness of the fluorescent tag inside the cell, for example, freely diffusing eGFP, and, second, to image the spatial distribution of the tagged protein with the APDs under the same instrumental conditions. Scanned images containing the true photon counts for a certain residence time of the laser (pixel time [PT]) can then easily be calibrated into a concentration map via

$$c_s(x,y) = \frac{1}{V_{eff}} \frac{\text{counts}(x,y)}{q_s PT} \tag{8.33}$$

where x, y are the pixel coordinates in the cross section and q_s is the molecular brightness of a fluorescent species s.

8.4.2. Scanning FCS

In cells, the molecules of interest often bind to membranes and other large-scale cellular structures that reduce their mobility. Large macromolecules such as DNA or molecular assemblies such as membrane vesicles may diffuse very slowly in free solution. Because time averaging in Eq. (8.1) has to be performed over a sufficient number of independent "fluctuation events," long measurement times are needed for the ACF to build up. However, in

[6] This avoids accumulation of large amounts of data and allows on-line inspection of the measured curves. However, with the constant advances in computer performance, software-based correlators also have come into use.

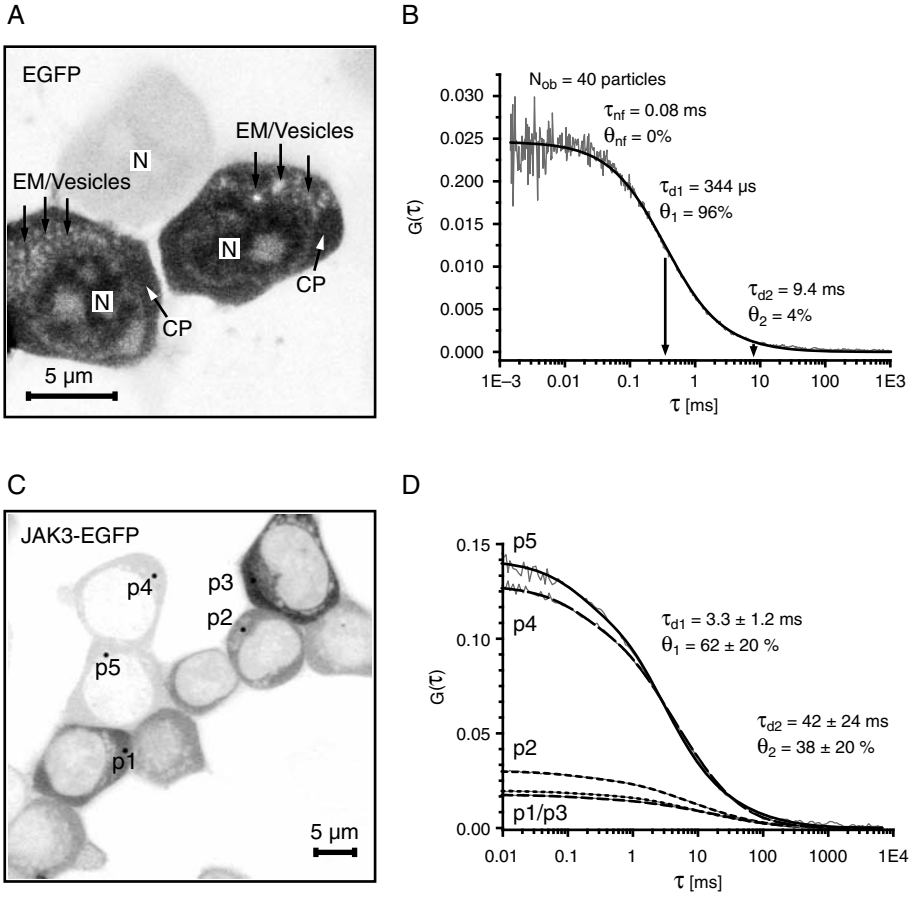


Figure 8.6. Examples for autocorrelation in living HEK 293T cells. (**A**) Confocal image of cells transiently expressing eGFP (239 amino acids) from the vector pEGFP-N1 (Clontech). Dark areas represent higher fluorescence intensities. The grey scale was chosen non-linear to illustrate sub-cellular structures. (**B**) Example ACF as measured in A at indicated cytoplasmic position CP. (**C**) Cells transiently expressing Janus kinase 3 (JAK3) fused to eGFP (1373 amino acids). (**D**) Corresponding correlation curves measured in C at indicated positions p1–p5. Note the reciprocal dependence between fluorescence intensity in the image and the amplitude of the ACF. The given diffusion times were averaged over p1–p5.

most of these cases, photobleaching of the fluorophores imposes tight limits for extending the measurement times. If diffusion is too slow or concentrations are too high, complementary techniques such as fluorescence recovery after photobleaching (FRAP) [41,42], continuous photobleaching (CP) [38,40], single-particle tracking (SPT) [43,44], or image correlation spectroscopy (ICS) [45,46] might be valuable alternatives to FCS.

An interesting hybrid approach for measuring slow fluctuations is scanning FCS, a family of techniques in which the laser focus is moved relative to the sample [43]. Although S-FCS correlates the signal from a moving detection volume, the high temporal resolution is preserved, and the technique could be classified between purely stationary temporal (FCS) and spatial image correlation (ICS). The motivations of S-FCS were to improve statistical accuracy, study binding to immobile samples, avoid photobleaching, and perform

measurements at many locations quasi-simultaneously. The method was invented [47] soon after FCS and later pioneered by Petersen et al. to study ligand-induced aggregation of receptors on the cell surface [48,49]. Subsequent studies employed linear [50–53], circular [54–58], or even random trajectories [59] of the laser focus and explored the application range for determining concentrations, aggregation states, and diffusion behavior of fluorescent particles in various systems.

A major drawback of S-FCS is that the scanned region should be homogeneous on a large scale. Although this is easily achieved in solution, measurements in membranes are prone to instabilities. In membranes the fluorescence originates from a thin, 5-nm layer placed in a micron-scale detection volume. The membrane, although a large object, has inherently slow thermal motions. Thus, repositioning of the membrane within different areas of the detection volume during data acquisition can lead to slow signal fluctuations, which dominate the measured ACF. A number of advanced concepts for membrane measurements have been recently described [60]. For example, linear perpendicular line scanning through the equatorial plane of giant lamellar vesicles was corrected for membrane undulations postmeasurement. The method requires saving the spatiotemporal photon stream and realigning the membrane-originating signal in time according to the maximum intensity of each line. The aligned intensity trace can then be correlated [61].

8.4.3. Alternative Excitation Modes

All of the aforementioned applications can be implemented with conventional continuous wave (CW) laser equipment providing a few milliwatts of power. In general, any excitation scheme for fluorescence that can be confined to a microscopic volume can be used to resolve number fluctuations and thus is suited to combining high-resolution imaging with FCS. Here we briefly introduce two of these, two-photon (TP) excitation and total internal reflection (TIR), which are widely applied for the fluorescence microscopy of cells.

TP excitation of fluorescent molecules occurs by simultaneous absorption of two photons of half the energy required for a transition to the excited state. Consequently, the wavelength used for illumination is much larger, about twice as large as the respective absorption maximum for conventional excitation. Simultaneous ($\sim 10^{-15}$ sec) absorption within the cross section of the dye molecule ($\sim 10^{-16}$ cm^2) requires high instantaneous photon-flux densities on the order of 10^{31} photons/(cm^2 sec), which are usually obtained from mode-locked lasers providing short, preferably $\sim$100-fs pulses with a high repetition rate of 40–80 MHz. The joint excitation probability of two photons per process is proportional to the mean square of the intensity and results in inherent depth discrimination: Only the immediate vicinity of the objective's focal spot receives sufficient intensity for absorption. In contrast to conventional CW excitation, where excitation still occurs all along the double cone–shaped profile of a focused laser beam (Figure 8.1B), in TP-FCS, photodestruction is confined to a smaller volume. The intrinsic compression of the illumination volume along the optical axis, as well as the higher penetration depth of infrared light in cells or tissues, has made two-photon laser scanning microscopy an attractive alternative to confocal imaging techniques [62]. Another potential advantage comes from the different photophysics underlying the excitation process. Two-photon excitation spectra of common fluorophores differ considerably from their one-photon counterparts without any change in emission, which opens up the possibility of accomplishing simultaneous excitation of spectrally distinct dyes [63]. This feature was successfully applied in a dual-color cross-correlation scheme, where a minimal spectral overlap in their emission properties is required [64]. It can be seen that the optical alignment for

TP-FCS is simplified compared to that sketched in Figure 8.5: Only one excitation line is needed, and, due to intrinsic volume discrimination, no pinholes are necessary. However, the achievable molecular brightness and hence the photon statistics are somewhat reduced when compared to CW applications.

Whereas all of the aforementioned model functions can be also used to evaluated TP-FCS data, different excitation geometries as in TIR microscopy require new expressions for the ACF. Although combination of TIR fluorescence microscopy and FCS was introduced some time ago [65], it has just recently been presented to a wider audience for its potential in studying surface-associated processes [66]. In TIR-FCS, light propagates parallel to a planar solid/liquid interface, that is, between a medium with higher refractive index n_2 (e.g., glass coverslip) and one with lower refractive index n_1 (e.g., a solution containing fluorescent molecules). If the incidence angle is larger than a critical value $\theta_{crit} = \arcsin(n_1/n_2)$, all of the light is reflected; however, a so-called evanescent wave penetrates the interface and leads to excitation of only a very thin layer ($\sim$100 nm) in solution. The intensity decay length perpendicular to the interface $d = \lambda \big/ 4\pi \left(n_2^2 \sin^2\theta - n_1^2\right)^{-1/2}$ depends on the vacuum wavelength λ, the refractive indices, and the angle of incidence θ. The lateral extension of the TIR-excitation profile is, at several microns, rather large, while only a small area is observed. As in CW excitation, out-off-focus fluorescence is rejected by a pinhole. Although this may cause some bleaching problems for very long measurements, the photon-flux densities are much lower than in conventional confocal microscopy. It can be quite complicated to derive analytical functions for fitting and to judge different application scenarios [67,68]. Nevertheless, TIR-FCS may become a promising method for investigating membrane-related events on solid supports such as lipid bilayers or even basal membrane areas where cells contact their substrate. TIR-FCS has been successfully used to study lateral translational diffusion of labeled immunoglobulin G (IgG) in planar model membranes [69,70], as well as the kinetics of IgG binding and dissociation to membrane-bound Fc receptors [71].

In general, alternative excitation schemes provide different focal geometries and length scales that are applicable to different topologies of the sample. Recently FCS was employed in combination with stimulated emission depletion (STED) microscopy, a new method that seems to have a huge potential for resolving dynamics and concentrations in hitherto inaccessible biological systems [72].

8.5. Biological Implications

Figure 8.7A illustrates the remarkable growth of FCS-related literature starting from the mid 1990s, when most of the components for FCS instrumentation, such as APDs and hardware correlators, became commercially available. The application of FCS at the single-cell level was somewhat delayed but has shown a dynamic development. By 2005, the annual number of publications on single-cell FCS measurements increased to more than 50. From this point of view, it is clear that a comprehensive discussion is beyond the scope of the chapter. We found 337 stringent hits for "FCS in living cells" with several keywords associated with different fields of cell biology. The results are summarized in Figure 8.7B. A clear preference is evident when looking at the top candidates: "membrane," "receptors," and "signaling"; these terms appear each in about 40% of the abstracts. Therefore, in the following, we put particular emphasis on the analysis of signaling by FCS.

A canonical signaling event is composed of a coordinated sequence of molecular rearrangements. This may include (1) binding of the ligand to corresponding cell surface

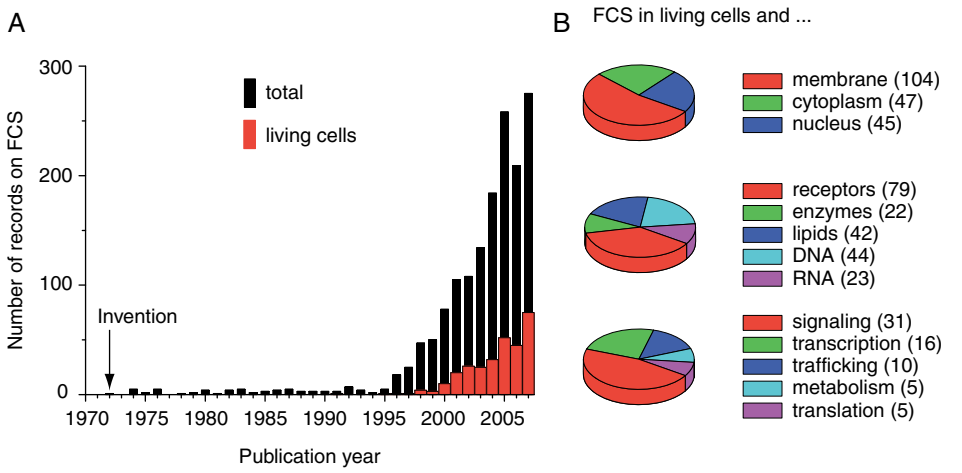


Figure 8.7. Literature analysis using the Web of Science (Thomson Reuters) search of the relevant databases. **A.** Publications per year on fluorescence correlation spectroscopy (FCS) ($n = 1{,}730$) after its invention in 1972. The subset of publications addressing living cells is shown in red ($n = 337$). **B.** The relative frequency of keywords for different cellular compartments, classes of biomolecules, and basic biological processes mentioned within this subset.

receptors, (2) ligand-induced conformational changes, (3) recruitment of the co-receptors, (4) transient binding of cytoplasmic proteins, and (5) activation and translocation of a freely diffusing signal transducers to their site of action. The initial step of receptor binding at the cell membrane leads to a dramatic reduction in mobility of the ligand and has been qualitatively probed by stationary FCS at the top of the plasma membrane. This includes labeled compounds [73,74] and labeled protein ligands [75]. Although conformational changes (step 2) are inaccessible for FCS, the oligomerization of receptors at the cell surface (step 3) was addressed by a number of studies. Baird and co-workers determined the mobility state of a fluorescently labeled IgE bound to FcεRI at the surface of mast cells [76,77]. The process is of major clinical importance because IgE-mediated clustering of FcεRI induces degranulation of the cells by which allergic mediators like histamines are released. A one-color autocorrelation analysis [77] corroborated FRAP data, and Larson et al. showed, by FCCS, the interaction between cross-linked IgE receptors and the downstream kinase Lyn, an inner leaflet–anchored signaling factor that is believed to partition into IgE-FcεRI forming membrane microdomains. IgE is the natural ligand for FcεRI; however, antibodies can be used to label other classes of cell surface receptors in different colors. Using FCCS with Alexa488- and Cy5-labeled antibodies specific for extracellular epitopes of interleukin receptor 2α (IL2Rα) and ILR15α, it was shown that the these receptor chains diffuse jointly in the surface membrane of lymphocytes together with members of the major histocompatibility complex (MHC) class I [78]. IL2Rα and ILR15α signal via the JAK/STAT pathway. STATs are signal transducers representing step 5, since they bind transiently to the activated receptors, become phosphorylated, dimerize, and translocate into the nucleus to modulate STAT-responsive genes. It was shown that the STAT-GFP constructs are functional [79], and the intracellular mobility of STAT3 was investigated by FCS [80]. Although the ligand-induced changes in particle numbers confirm a translocation process into the nucleus, the reported diffusion times of STAT3-GFP are difficult to interpret in terms of molecular mass.

As often for intracellular FCS, the ACF decays more slowly and with more than one correlation time as expected from pure diffusion (see, e.g., Figure 8.6). This phenomenon was attributed to hindered, so-called anomalous diffusion, that is, diffusion behavior in which the mean square displacement of the molecules increases more slowly and nonlinearly as compared to normal 3D Brownian motion. Anomalous diffusion has been observed in basically every cellular compartment [39,81–83]. The underlying reasons range from topological constraints due to endogenous membranes or macromolecular barriers such as chromatin to transient unspecific interactions. Unfortunately, FCS is poorly suited to discriminate between different molecular mechanisms. Recently, it was proposed to probe anomalous diffusion in membranes by plotting the correlation times for varying detection volumes [84]. Extrapolating the linear dependency of τ_d versus V_{eff} for the limit $V_{eff} = 0$ crosses the origin in the case of free diffusion but leads to a constant offset in the case of confinements. Using this method, it was shown for the death receptor Fas (CD95) that a palmitoylation site—a cysteine close to the membrane—is responsible for recruiting Fas into signaling relevant lipid rafts [85], whereas a mutant lacking this cysteine shows no offset for $V_{eff} = 0$. In a similar way it was shown that the Fas ligand (CD95 ligand) and the raft marker GM_1 diffuse in confined way, whereas a fluorescent phosphatidyl choline does not [86].

Clustering and activation of the Fas receptor by the Fas ligand represent one pathway leading to apoptosis [87]. A key step of apoptosis is the activation of caspases, a family of proteases that linger in the cytoplasm as inactive proenzymes. The precise mechanism for activating these enzymes is not known. The decision of the cells to live or die is of utmost importance in cell biology and believed to depend critically on subtle changes in concentration ratios of pro- and antiapoptotic factors [88]. An interesting single-molecule assay for monitoring protease activity makes use of enzyme-specific peptidic cleavage substrates flanked at both ends with orthogonal colored fluorescent proteins [89]. Two such examples were reported for caspase-3 [89 90] and used to screen the effect of low–molecular weight inhibitors [90].

Divalent calcium ions function as ubiquitous intracellular mediators in almost every signaling pathway. One of the most abundant Ca^{2+}-binding proteins is calmodulin (CaM), which can make up about 1% of the total protein mass of a eukaryotic cell. On binding of four Ca^{2+} ions, CaM undergoes a switch-like conformational change, which in turn enables the recognition of CaM-dependent factors such as CaM kinases. The interaction between Alexa633-labeled CaM and a CaM kinase II-eGFP was investigated by TP-FCCS [33,91]. The latter study is of particular interest because it is one of the rare examples in which FCCS amplitudes were quantitatively evaluated to derive the stoichiometry of the complex (CaM:CaM kinase II = 12:1).

The largest family of receptors transmitting signals from the external environment into the cell are G protein–coupled receptors. The seven-pass transmembrane receptors act in a ligand-induced way as a guanosine triphosphate (GTP)–exchanging factor for the trimeric G protein ($G_{\alpha\beta\gamma}$). G proteins are anchored via two alkyl chains (G_α, G_γ) to the inner leaflet. The question of whether the trimeric G protein is recruited on stimulation or associated with the receptor in the quiescent state is under debate. Equally interesting is the integrity of the trimeric G protein after GTP binding. Although it was anticipated for a long time that the activated GTP-binding α-subunit always dissociates from $G_{\beta\gamma}$, it was reported, at least in some cases, that activation is merely a result of exposing hitherto buried binding sites of the intact trimeric complex. Both topics were recently addressed by FCS in culture cells [92,93], as well as in *Dictyostelium discoideum* [94]. It was found that the diffusion coefficient of the $G_{\beta\gamma}$ was significantly increased on stimulation and, of interest, that this induced mobility

change depended critically on the presence of caveolin-1, the main component of certain membrane domains termed caveolae [95].

Although G protein–coupled receptor (GPCR) activation and phosphorylation cascades at the surface membrane are often rather short termed, the signaling pathways governing cell cycle control must convert the response into a long-lasting change. A prominent family of kinases responsible for this task is made up of the mitogen-activated (MAP) protein kinases. MAP kinases activate each other in successive order (MAPK → MAPKK → MAPKKK). Two groups independently investigated GFP- and mCherry-fused MAP kinases in living yeast cells [96,97]. The studies demonstrated the huge advantage of combining genetic manipulation with high-sensitivity detection. Combinations of protein–protein interaction were assayed at physiological concentrations in yeast strains. In yeast, the fluorescent constructs were introduced by means of homologous recombination and thus were expressed under the control of natural promoters. Maeder et al. [96] used three fluorescent protein domains in a row to improve the molecular brightness. Both studies quantified dissociation constants for pairwise interactions, albeit with discrepancies spanning two orders of magnitude. Recently, cell cycle control was monitored by quantifying concentrations and diffusion of the cyclin-dependent protein factor Cdc20 fused to eCFP [98].

Acknowledgments

TW is indebted to Konstantin Klenin and Malte Wachsmuth for their contributions in deriving FCS theory. TW and PS thank Eugene Petrov for helpful discussions.

References

1. Bacia, K., Kim, S. A., & Schwille, P. (2006). Fluorescence cross-correlation spectroscopy in living cells. *Nat Methods* **3**, 83–9.
2. Kim, S. A., Heinze, K. G., & Schwille, P. (2007). Fluorescence correlation spectroscopy in living cells. *Nat Methods* **4**, 963–73.
3. Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy*. 3rd ed. New York: Springer.
4. Bacia, K., & Schwille, P. (2007). Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. *Nat Protoc* **2**, 2842–56.
5. Haustein, E., & Schwille, P. (2003). Ultrasensitive investigations of biological systems by fluorescence correlation spectroscopy. *Methods* **29**, 153–66.
6. Hess, S. T., Huang, S., Heikal, A. A., & Webb, W. W. (2002). Biological and chemical applications of fluorescence correlation spectroscopy: a review. *Biochemistry* **41**, 697–705.
7. Weidemann, T., Wachsmuth, M., Tewes, M., Rippe, K., & Langowski, J. (2002). Analysis of ligand binding by two-colour fluorescence cross-correlation spectroscopy. *Single Mol* **3**, 49–61.
8. Widengren, J., & Rigler, R. (1998). Fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces. *Cell Mol Biol* (Noisy-le-grand) **44**, 857–79.
9. Enderlein, J., Gregor, I., Patra, D., Dertinger, T., & Kaupp, U. B. (2005). Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. *Chemphyschem* **6**, 2324–36.
10. Hess, S. T., & Webb, W. W. (2002). Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophys J* **83**, 2300–17.
11. Dertinger, T., Pacheco, V., von der Hocht, I., Hartmann, R., Gregor, I., & Enderlein, J. (2007). Two-focus fluorescence correlation spectroscopy: a new tool for accurate and absolute diffusion measurements. *Chemphyschem* **8**, 433–43.
12. Magde, D., Elson, E. L., & Webb, W. W. (1972). Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlations spectroscopy. *Phys Rev Let* **29**, 705–708.

13. Magde, D., Elson, E. L., & Webb, W. W. (1974). Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers* **13**, 29–61.
14. Magde, D., Webb, W. W., & Elson, E. L. (1978). Fluorescence correlation spectroscopy. III. Uniform translation and laminar flow. *Biopolymers* **17**, 361–76.
15. Meseth, U., Wohland, T., Rigler, R., & Vogel, H. (1999). Resolution of fluorescence correlation measurements. *Biophys J* **76**, 1619–31.
16. Widengren, J., Mets, Ü., & Rigler, R. (1995). Fluorescence correlation spectroscopy of triplet states in solution: a theoretical and experimental study. *J Phys Chem* **99**, 13368–79.
17. Bosisio, C., Quercioli, V., Collini, M., D'Alfonso, L., Baldini, G., Bettati, S., Campanini, B., Raboni, S., & Chirico, G. (2008). Protonation and conformational dynamics of GFP mutants by two-photon excitation fluorescence correlation spectroscopy. *J Phys Chem B* **112**, 8806–14.
18. Haupts, U., Maiti, S., Schwille, P., & Webb, W. W. (1998). Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proc Natl Acad Sci USA* **95**, 13573–8.
19. Heikal, A. A., Hess, S. T., Baird, G. S., Tsien, R. Y., & Webb, W. W. (2000). Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: coral red (dsRed) and yellow (Citrine). *Proc Natl Acad Sci USA* **97**, 11996–2001.
20. Hendrix, J., Flors, C., Dedecker, P., Hofkens, J., & Engelborghs, Y. (2008). Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. *Biophys J* **94**, 4103–13.
21. Malvezzi-Campeggi, F., Jahnz, M., Heinze, K. G., Dittrich, P., & Schwille, P. (2001). Light-induced flickering of DsRed provides evidence for distinct and interconvertible fluorescent states. *Biophys J* **81**, 1776–85.
22. Schenk, A., Ivanchenko, S., Rocker, C., Wiedenmann, J., & Nienhaus, G. U. (2004). Photodynamics of red fluorescent proteins studied by fluorescence correlation spectroscopy. *Biophys J* **86**, 384–94.
23. Schwille, P., Kummer, S., Heikal, A. A., Moerner, W. E., & Webb, W. W. (2000). Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins. *Proc Natl Acad Sci USA* **97**, 151–6.
24. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., & Remington, S. J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392–5.
25. Heim, R., Prasher, D. C., & Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci USA* **91**, 12501–4.
26. Cormack, B. P., Valdivia, R. H., & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–8.
27. Liu, Y., Kim, H. R., & Heikal, A. A. (2006). Structural basis of fluorescence fluctuation dynamics of green fluorescent proteins in acidic environments. *J Phys Chem B* **110**, 24138–46.
28. Wong, F. H., Banks, D. S., Abu-Arish, A., & Fradin, C. (2007). A molecular thermometer based on fluorescent protein blinking. *J Am Chem Soc* **129**, 10302–3.
29. Thompson, N. L. (1991). Fluorescence correlation spectroscopy. In *Topics in Fluorescence Spectroscopy. Vol. 1: Techniques* (Lankowicz, J. R., ed.), pp. 337–378. New York: Plenum Press.
30. Koppel, D. (1974). Statistical accuracy in fluorescence correlation spectroscopy. *Phys Rev A* **10**, 1938–45.
31. Eigen, M., & Rigler, R. (1994). Sorting single molecules: application to diagnostics and evolutionary biotechnology. *Proc Natl Acad Sci USA* **91**, 5740–7.
32. Schwille, P., Meyer-Almes, F. J., & Rigler, R. (1997). Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys J* **72**, 1878–86.
33. Kim, S. A., Heinze, K. G., Bacia, K., Waxham, M. N., & Schwille, P. (2005). Two-photon cross-correlation analysis of intracellular reactions with variable stoichiometry. *Biophys J* **88**, 4319–36.
34. Stoevesandt, O., & Brock, R. (2006). One-step analysis of protein complexes in microliters of cell lysate using indirect immunolabeling & fluorescence cross-correlation spectroscopy. *Nat Protoc* **1**, 223–9.
35. Berland, K. M., So, P. T., & Gratton, E. (1995). Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment. *Biophys J* **68**, 694–701.
36. Brock, R., Hink, M. A., & Jovin, T. M. (1998). Fluorescence correlation microscopy of cells in the presence of autofluorescence. *Biophys J* **75**, 2547–57.
37. Gennerich, A., & Schild, D. (2000). Fluorescence correlation spectroscopy in small cytosolic compartments depends critically on the diffusion model used. *Biophys J* **79**, 3294–306.
38. Wachsmuth, M., Weidemann, T., Muller, G., Hoffmann-Rohrer, U. W., Knoch, T. A., Waldeck, W., & Langowski, J. (2003). Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophys J* **84**, 3353–63.
39. Wachsmuth, M., Waldeck, W., & Langowski, J. (2000). Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J Mol Biol* **298**, 677–89.

40. Weidemann, T., Wachsmuth, M., Knoch, T. A., Muller, G., Waldeck, W., & Langowski, J. (2003). Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J Mol Biol* **334**, 229–40.
41. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E., & Webb, W. W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys J* **16**, 1055–69.
42. Sprague, B. L., & McNally, J. G. (2005). FRAP analysis of binding: proper and fitting. *Trends Cell Biol* **15**, 84–91.
43. Chen, Y., Lagerholm, B. C., Yang, B., & Jacobson, K. (2006). Methods to measure the lateral diffusion of membrane lipids and proteins. *Methods* **39**, 147–53.
44. Guo, L., Har, J. Y., Sankaran, J., Hong, Y., Kannan, B., & Wohland, T. (2008). Molecular diffusion measurement in lipid bilayers over wide concentration ranges: a comparative study. *Chemphyschem* **9**, 721–8.
45. Kolin, D. L., & Wiseman, P. W. (2007). Advances in image correlation spectroscopy: measuring number densities, aggregation states, and dynamics of fluorescently labeled macromolecules in cells. *Cell Biochem Biophys* **49**, 141–64.
46. Petersen, N. O., Hoddelius, P. L., Wiseman, P. W., Seger, O., & Magnusson, K. E. (1993). Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. *Biophys J* **65**, 1135–46.
47. Weissman, M., Schindler, H., & Feher, G. (1976). Determination of molecular weights by fluctuation spectroscopy: application to DNA. *Proc Natl Acad Sci USA* **73**, 2776–80.
48. Petersen, N. O. (1986). Scanning fluorescence correlation spectroscopy. I. Theory and simulation of aggregation measurements. *Biophys J* **49**, 809–15.
49. Petersen, N. O., Johnson, D. C., & Schlesinger, M. J. (1986). Scanning fluorescence correlation spectroscopy. II. Application to virus glycoprotein aggregation. *Biophys J* **49**, 817–20.
50. Digman, M. A., Brown, C. M., Sengupta, P., Wiseman, P. W., Horwitz, A. R., & Gratton, E. (2005). Measuring fast dynamics in solutions and cells with a laser scanning microscope. *Biophys J* **89**, 1317–27.
51. Digman, M. A., Sengupta, P., Wiseman, P. W., Brown, C. M., Horwitz, A. R., & Gratton, E. (2005). Fluctuation correlation spectroscopy with a laser-scanning microscope: exploiting the hidden time structure. *Biophys J* **88**, L33–6.
52. Koppel, D. E., Morgan, F., Cowan, A. E., & Carson, J. H. (1994). Scanning concentration correlation spectroscopy using the confocal laser microscope. *Biophys J* **66**, 502–7.
53. Xiao, Y., Buschmann, V., & Weston, K. D. (2005). Scanning fluorescence correlation spectroscopy: a tool for probing microsecond dynamics of surface-bound fluorescent species. *Anal Chem* **77**, 36–46.
54. Berland, K. M., So, P. T., Chen, Y., Mantulin, W. W., & Gratton, E. (1996). Scanning two-photon fluctuation correlation spectroscopy: particle counting measurements for detection of molecular aggregation. *Biophys J* **71**, 410–20.
55. Meyer, T., & Schindler, H. (1988). Particle counting by fluorescence correlation spectroscopy. Simultaneous measurement of aggregation and diffusion of molecules in solutions and in membranes. *Biophys J* **54**, 983–93.
56. Petrasek, Z., & Schwille, P. (2008). Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys J* **94**, 1437–48.
57. Ruan, Q., Cheng, M. A., Levi, M., Gratton, E., & Mantulin, W. W. (2004). Spatial-temporal studies of membrane dynamics: scanning fluorescence correlation spectroscopy (SFCS). *Biophys J* **87**, 1260–7.
58. Skinner, J. P., Chen, Y., & Muller, J. D. (2005). Position-sensitive scanning fluorescence correlation spectroscopy. *Biophys J* **89**, 1288–301.
59. Amediek, A., Haustein, E., Scherfeld, D., & Schwille, P. (2002). Scanning dual-color cross-correlation analysis for dynamic co-localization studies of immobile molecules. *Single Mol* **3**, 201–210.
60. Ries, J., & Schwille, P. (2008). New concepts for fluorescence correlation spectroscopy on membranes. *Phys Chem Chem Phys* **10**, 3487–97.
61. Ries, J., & Schwille, P. (2006). Studying slow membrane dynamics with continuous wave scanning fluorescence correlation spectroscopy. *Biophys J* **91**, 1915–24.
62. Denk, W., Strickler, J. H., & Webb, W. W. (1990). 2-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76.
63. Xu, C., Zipfel, W., Shear, J. B., Williams, R. M., & Webb, W. W. (1996). Multiphoton fluorescence excitation: new spectral windows for biological nonlinear microscopy. *Proc Natl Acad Sci USA* **93**, 10763–68.
64. Heinze, K. G., Koltermann, A., & Schwille, P. (2000). Simultaneous two-photon excitation of distinct labels for dual-color fluorescence crosscorrelation analysis. *Proc Natl Acad Sci USA* **97**, 10377–82.
65. Thompson, N. L., & Axelrod, D. (1983). Immunoglobulin surface-binding kinetics studied by total internal reflection with fluorescence correlation spectroscopy. *Biophys J* **43**, 103–14.

66. Thompson, N. L., & Steele, B. L. (2007). Total internal reflection with fluorescence correlation spectroscopy. *Nat Protoc* **2**, 878–90.
67. Lieto, A. M., & Thompson, N. L. (2004). Total internal reflection with fluorescence correlation spectroscopy: nonfluorescent competitors. *Biophys J* **87**, 1268–78.
68. Starr, T. E., & Thompson, N. L. (2001). Total internal reflection with fluorescence correlation spectroscopy: combined surface reaction and solution diffusion. *Biophys J* **80**, 1575–84.
69. Pero, J. K., Haas, E. M., & Thompson, N. L. (2006). Size dependence of protein diffusion very close to membrane surfaces: measurement by total internal reflection with fluorescence correlation spectroscopy. *J Phys Chem B* **110**, 10910–18.
70. Starr, T. E., & Thompson, N. L. (2002). Local diffusion and concentration of IgG near planar membranes: measurement by total internal reflection with fluorescence correlation spectroscopy. *J Phys Chem B* **106**, 2365–71.
71. Lieto, A. M., Cush, R. C., & Thompson, N. L. (2003). Ligand-receptor kinetics measured by total internal reflection with fluorescence correlation spectroscopy. *Biophys J* **85**, 3294–302.
72. Kastrup, L., Blom, H., Eggeling, C., & Hell, S. W. (2005). Fluorescence fluctuation spectroscopy in subdiffraction focal volumes. *Phys Rev Lett* **94**(17): 178104.
73. Briddon, S. J., Middleton, R. J., Yates, A. S., George, M. W., Kellam, B., & Hill, S. J. (2004). Application of fluorescence correlation spectroscopy to the measurement of agonist binding to a G-protein coupled receptor at the single cell level. *Faraday Discuss* **126**, 197–207; discussion 245–54.
74. Schwartz, J. W., Novarino, G., Piston, D. W., & DeFelice, L. J. (2005). Substrate binding stoichiometry and kinetics of the norepinephrine transporter. *J Biol Chem* **280**, 19177–84.
75. Abouzied, M. M., El-Tahir, H. M., Prenner, L., Haberlein, H., Gieselmann, V., & Franken, S. (2005). Hepatoma-derived growth factor. Significance of amino acid residues 81–100 in cell surface interaction and proliferative activity. *J Biol Chem* **280**, 10945–54.
76. Larson, D. R., Gosse, J. A., Holowka, D. A., Baird, B. A., & Webb, W. W. (2005). Temporally resolved interactions between antigen-stimulated IgE receptors and Lyn kinase on living cells. *J Cell Biol* **171**, 527–36.
77. Pyenta, P. S., Schwille, P., Webb, W. W., Holowka, D., & Baird, B. (2003). Lateral diffusion of membrane lipid-anchored probes before and after aggregation of cell surface IgE-receptors. *J Phys Chem A* **107**, 8310–18.
78. Vamosi, G., Bodnar, A., Vereb, G., Jenei, A., Goldman, C. K., Langowski, J., Toth, K., Matyus, L., Szollosi, J., Waldmann, T. A., & Damjanovich, S. (2004). IL-2 and IL-15 receptor alpha-subunits are coexpressed in a supramolecular receptor cluster in lipid rafts of T cells. *Proc Natl Acad Sci USA* **101**, 11082–87.
79. Kretzschmar, A. K., Dinger, M. C., Henze, C., Brocke-Heidrich, K., & Horn, F. (2004). Analysis of Stat3 (signal transducer and activator of transcription 3) dimerization by fluorescence resonance energy transfer in living cells. *Biochem J* **377**, 289–97.
80. Watanabe, K., Saito, K., Kinjo, M., Matsuda, T., Tamura, M., Kon, S., Miyazaki, T., & Uede, T. (2004). Molecular dynamics of STAT3 on IL-6 signaling pathway in living cells. *Biochem Biophys Res Commun* **324**, 1264–73.
81. Schwille, P., Haupts, U., Maiti, S., & Webb, W. W. (1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophys J* **77**, 2251–65.
82. Weiss, M., Elsner, M., Kartberg, F., & Nilsson, T. (2004). Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophys J* **87**, 3518–24.
83. Weiss, M., Hashimoto, H., & Nilsson, T. (2003). Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys J* **84**, 4043–52.
84. Wawrezinieck, L., Rigneault, H., Marguet, D., & Lenne, P. F. (2005). Fluorescence correlation spectroscopy diffusion laws to probe the submicron cell membrane organization. *Biophys J* **89**, 4029–42.
85. Chakrabandhu, K., Herincs, Z., Huault, S., Dost, B., Peng, L., Conchonaud, F., Marguet, D., He, H. T., & Hueber, A. O. (2007). Palmitoylation is required for efficient Fas cell death signaling. *EMBO J* **26**, 209–20.
86. Cahuzac, N., Baum, W., Kirkin, V., Conchonaud, F., Wawrezinieck, L., Marguet, D., Janssen, O., Zornig, M., & Hueber, A. O. (2006). Fas ligand is localized to membrane rafts, where it displays increased cell death-inducing activity. *Blood* **107**, 2384–91.
87. Nagata, S. (1997). Apoptosis by death factor. *Cell* **88**, 355–65.
88. Krammer, P. H., Arnold, R., & Lavrik, I. N. (2007). Life and death in peripheral T cells. *Nat Rev Immunol* **7**, 532–42.
89. Kohl, T., Haustein, E., & Schwille, P. (2005). Determining protease activity in vivo by fluorescence cross-correlation analysis. *Biophys J* **89**, 2770–82.
90. Keese, M., Offterdinger, M., Tischer, C., Girod, A., Lommerse, P. H., Yagublu, V., Magdeburg, R., & Bastiaens, P. I. (2007). Quantitative imaging of apoptosis commitment in colorectal tumor cells. *Differentiation* **75**, 809–18.

91. Kim, S. A., Heinze, K. G., Waxham, M. N., & Schwille, P. (2004). Intracellular calmodulin availability accessed with two-photon cross-correlation. *Proc Natl Acad Sci USA* **101**, 105–10.
92. Philip, F., Sengupta, P., & Scarlata, S. (2007). Signaling through a G Protein-coupled receptor and its corresponding G protein follows a stoichiometrically limited model. *J Biol Chem* **282**, 19203–16.
93. Sengupta, P., Philip, F., & Scarlata, S. (2008). Caveolin-1 alters Ca(2+) signal duration through specific interaction with the G alpha q family of G proteins. *J Cell Sci* **121**, 1363–72.
94. Knol, J. C., Engel, R., Blaauw, M., Visser, A. J., & van Haastert, P. J. (2005). The phosducin-like protein PhLP1 is essential for G{beta}{gamma} dimer formation in Dictyostelium discoideum. *Mol Cell Biol* **25**, 8393–400.
95. Simons, K., & Vaz, W. L. (2004). Model systems, lipid rafts, and cell membranes. *Annu Rev Biophys Biomol Struct* **33**, 269–95.
96. Maeder, C. I., Hink, M. A., Kinkhabwala, A., Mayr, R., Bastiaens, P. I., & Knop, M. (2007). Spatial regulation of Fus3 MAP kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling. *Nat Cell Biol* **9**, 1319–26.
97. Slaughter, B. D., Schwartz, J. W., & Li, R. (2007). Mapping dynamic protein interactions in MAP kinase signaling using live-cell fluorescence fluctuation spectroscopy and imaging. *Proc Natl Acad Sci USA* **104**, 20320–5.
98. Wang, Z., Shah, J. V., Berns, M. W., & Cleveland, D. W. (2006). In vivo quantitative studies of dynamic intracellular processes using fluorescence correlation spectroscopy. *Biophys J* **91**, 343–51.

9

Precise Measurements of Diffusion in Solution by Fluorescence Correlations Spectroscopy

Jörg Enderlein

Abstract The chapter presents a general introduction to the method of fluorescence correlation spectroscopy (FCS) and its recently developed modification, dual-focus FCS. It explains the general optical setup of an FCS system, data acquisition, and data analysis. It discusses numerous potential optical and photophysical artifacts of FCS measurements and explains how these problems are circumvented by dual-focus FCS.

9.1. Introduction

Fluorescence correlation spectroscopy (FCS) is the general name for a set of spectroscopic methods that are based on the measurement and correlation analysis of fluorescence intensity fluctuations originating from a small number of fluorescing molecules, usually contained within a sufficiently small detection region. Any process that influences the fluorescence intensity of these molecules (such as changes of their position within the measurement system, their photophysics, chemical reactions, conformational changes, etc.) will lead to a temporally changing fluorescence signal, most often in a stochastic way. For example, molecules that are free to diffuse in an out of the detection region will generate a stochastically changing fluorescence intensity signal. Similarly, molecules that, besides cycling through the first excited singlet and ground states, can switch from time to time into a nonfluorescent triplet state will generate a fluctuating fluorescence intensity signal. The important point is that the character of these fluorescence signal fluctuations is connected with the underlying physical processes and their parameters (such as the diffusion coefficient or photophysical rate constants). The core idea of FCS is to evaluate the observed intensity fluctuation in such a way that one can determine these parameters. The standard approach is to perform a second-order correlation analysis on the measured fluorescence intensity signal by calculating

$$g(\tau) = \langle I(t)I(t+\tau)\rangle\,, \tag{9.1}$$

J. Enderlein • Drittes Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_9,

where $I(t)$ is the fluorescence intensity at time t, and the angular brackets denote averaging over all time values t. The physical meaning of the autocorrelation is that it is directly proportional to the probability of detecting a photon at time τ if there was a photon detection event at time zero. This probability is composed of two basically different terms. First, the two photons detected at time zero and at time τ can originate from uncorrelated background or from different fluorescing molecules and therefore do not have any physical correlation (provided there is no interaction of the different fluorescing molecules). These events will contribute to a constant offset of $g(\tau)$ that is completely independent of τ (the joint probability to detect two physically uncorrelated photons is completely independent of the time between their detection). Second, the two photons can originate from the same molecule and are then physically correlated.

Let us start with some very simple qualitative considerations concerning the lag-time dependence of $g(\tau)$. Suppose a molecule is close to the center of the detection volume. Then there will be a high probability of detecting a large number of consecutive fluorescence photons from this molecule, that is, the fluorescence signal will be highly correlated in time. When the molecule, due to diffusion, starts to exit the detection volume, this correlation will continually decrease, that is, the probability of seeing further fluorescence photons will decrease in time until the molecule has completely diffused away and the correlation is completely lost. Of course, the temporal decay of the correlation, that is, the temporal decay of $g(\tau)$ with increasing lag-time τ, will be proportional to the diffusion speed of the molecule: the larger the diffusion coefficient, the faster are the fluorescence correlation decays.

A second important property of the autocorrelation function (ACF) is its dependence on the concentration of fluorescing molecules. It is rather obvious that the fluorescence intensity fluctuations will be larger for lower molecule concentrations. Indeed, if one has, on average, only a signal molecule within the detection volume, than the diffusion of this molecule out of this volume or the diffusion of another molecule into this volume will cause a big change in measured fluorescence intensity. On the contrary, if the average number of fluorescing molecules within the detection volume is rather large (e.g., several hundreds), then the leaving or entering of a molecule causes only small signal variations. Intuitively, one may expect a direct connection between the average number of molecules within the detection volume (i.e., concentration) and the amplitude of the fluorescence intensity fluctuations. Indeed, there is a direct connection between the inverse concentration of fluorescing molecules and the amplitude of the ACF.

Thus, FCS measurements can provide information about the diffusion and concentration of fluorescing molecules. Any process that alters one (or both) of these quantities can also be measured by FCS. For example, consider the binding of two proteins in solution: By labelling one of the binding partners with a fluorescence label and monitoring with FCS the changing value of the diffusion coefficient of the labelled molecules on binding of their binding partner, one can directly measure binding affinities and kinetics.

On different time scales, the temporal behaviour of the autocorrelation function is determined by different properties of the fluorescing molecules: On a nanosecond time scale, photon antibunching can be observed, reflecting the fact that directly after the emission of a photon the molecule needs to get reexcited to be able to emit the next photon, leading to a steep decrease of $g(\tau)$ towards short times. On a microsecond time scale, $g(\tau)$ is dominated by triplet-state dynamics. If excitation and/or detection is performed with polarization filters, the autocorrelation will also show contributions from rotational diffusion dynamics of the molecules. On a millisecond-to-second level, the autocorrelation

function shows a typical decay due to the lateral diffusion of the molecules out of the detection region. The diffusion coefficient is also the parameter most frequently addressed by FCS measurements.

FCS was originally introduced by Elson, Magde, and Webb in the early 1970s [1–3]. In its original form it was invented for measuring diffusion, concentration, and chemical or biochemical interactions or reactions of fluorescent or fluorescently labelled molecules at nanomolar concentrations in solution. It took nearly two decades until the technique saw a renaissance with the development of new lasers with high beam quality and temporal stability, low-noise, single-photon detectors, and high-quality microscope objectives with nearly perfect imaging quality at high numerical aperture. Achieving values of the detection volume within the range of a few cubic micrometers made the technique applicable for samples at reasonably high concentrations and short measurement times.

The advantage of FCS is its relative simplicity. Its drawback is that it works only within a very limited concentration range: If the concentration of fluorescing molecules becomes too large (typically $>>10^{-8}$ M), then the contribution from correlated photons from individual molecules, scaling with the number N of molecules within the detection volume, becomes very small compared with the contribution by uncorrelated photons from different molecules, scaling with N^2. If the concentration is too low (typically $<10^{-13}$ M), then the probability of finding a molecule within the detection region becomes extremely low. In both cases, the measurement time for obtaining a high-quality autocorrelation function gets prohibitively large, although a remedy for that problem is to rapidly scan the laser focus through the solution [4,5].

There are numerous excellent reviews and overviews of FCS [6–8], as well as a complete book devoted to it [9]. This chapter gives a very general introduction to the philosophy of FCS and attempts to be self-contained, developing the fundamental principles of FCS and also describing recent methodological advances that are not well covered by previous reviews. In what follows, the focus will be mainly on the application of FCS to precisely measure diffusion coefficients. This also allows a thorough discussion of the experimental setup, potential optical problems, and the data evaluation.

9.2. Optical Setup

A typical FCS measurement setup is shown in Figure 9.1 [10]. Fluorescent molecules are dissolved in an aqueous solution that is placed on top of a chambered cover-slide. A collimated laser beam with perfect Gaussian TEM_{00} mode [11] is coupled via a dichroic mirror into an objective with high numerical aperture (NA) that focuses the laser into a diffraction-limited spot in the sample. The dichroic mirror is reflective at the laser's wavelength and transmissive at the wavelengths of the fluorescence emission. The use of a Gaussian TEM_{00} mode assures diffraction-limited focusing of light, thus achieving minimum focus diameter in the sample. Fluorescence light generated in the sample is collected by the same objective (so-called epifluorescence setup), transmitted through the dichroic mirror, and focused onto a circular confocal aperture. Behind the aperture, the fluorescence light is refocused onto two sensitive light detectors, usually single-photon avalanche diodes (SPADs). The confocal aperture effectively rejects fluorescence light that is generated outside the focal plane. In combination, fluorescence generation (by diffraction-limited focusing of excitation light) and fluorescence detection (by confocal detection) generate an effective detection volume

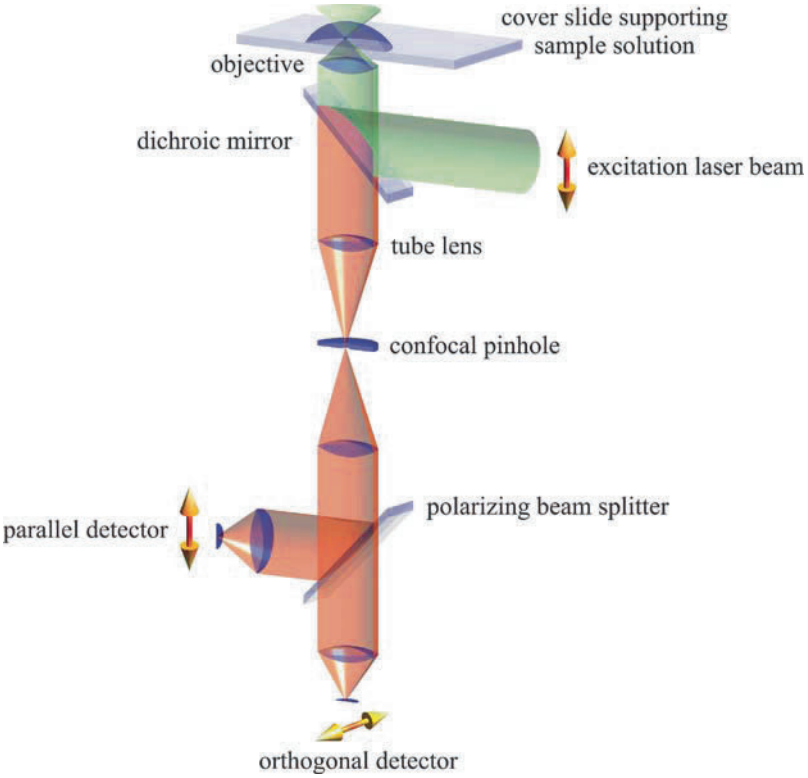


Figure 9.1. Principal scheme of a confocal epifluorescence microscope as used in fluorescence correlation spectroscopy . Shown is a setup with a linearly polarized excitation laser. Detection is done within two detection channels after splitting the light with a polarizing beam splitter. The detectors are usually single-photon avalanche diodes or photomultiplier tubes. The vertical position of the objective can be accurately adjusted using a piezo actuator (not shown).

of about 0.5 μm in diameter in the focal plane and a few micrometers along the optical axis.

Use of two detectors is important for efficient elimination of the effects of SPAD dead time and afterpulsing on an ACF. Usually, detector dead times are in the range of several tens to hundreds of nanoseconds. They cause the measured ACF to drop towards zero at lag times that are comparable with the detector's dead time. Detector afterpulsing is the effect that a genuine photon detection pulse is followed by a so-called afterpulse with a delay between microseconds to seconds. The origin of afterpulsing in SPADs is as follows: A primary photoelectron initiates an avalanche of ionizations that causes a breakdown pulse at the detector output. Some of the generated charge carriers may become temporarily trapped and afterwards released by thermal excitation, so that new charge carriers are created that lead to afterpulses that are correlated with the initial event. The probability of afterpulsing depends on many different parameters, such as material defects, temperature, and operating conditions of the detector. Afterpulsing becomes typically visible as a fast decay of the ACF at microsecond lag times. By using two detectors and correlating only photons from different detectors, the effects of both dead time and afterpulsing are successfully eliminated when calculating an ACF.

The exact shape and size of that detection volume determine the shape and temporal decay of the autocorrelation function. For example, the smaller the detection volume, the faster will molecules diffuse out of it and the faster will the ACF decays, and vice versa. The actual quantity that defines the autocorrelation function is the so-called molecule detection function (MDF). The MDF describes the chance of seeing a fluorescence photon from a molecule at a given position **r** in the sample. Thus, the MDF, which we will denote by $U(\mathbf{r})$, is a function of position **r** and rapidly falls off to zero if one moves away from the optical axis and/or the focal plane. As we will see, knowing the exact shape of the MDF allows one to calculate exactly the shape of the ACF, which can then be used to fit experimental data for obtaining, for example, diffusion and/or concentration values of the fluorescent molecules. However, this is also the principal problem of FCS: A precise quantitative evaluation of an ACF critically depends on the exact knowledge of the MDF. This is discussed in detail in Section 9.4.

9.3. Data Acquisition and Evaluation

Conventional FCS setups used to employ hardware autocorrelators that calculated the ACF onboard on the basis of the signal from the photodetectors. Recently, most setups have been using fast photon counting electronics for asynchronously recording and storing the arrival times of the detected photons, and subsequently use software algorithms for calculating the ACF from the recorded photon data. This permits much more flexibility in data handling and evaluation, as will be seen, for example, in the case of fluorescence lifetime correlation spectroscopy, and we will describe this approach here in more detail.

Asynchronously measured single-photon counting data consist of a linear file of detection times $\{t_1, t_2, \ldots, t_N\}$ of the detected photons, where N is the total number of detected photons during a given measurement. A special feature of these detection times is that they are integer multiples of some minimal time δt, determined by the temporal resolution of the detection electronics. Without restriction of generality, it can be assumed that all times are measured in units of δt, so that all the numbers t_j take integer values. The value $g(\tau)$ of the autocorrelation function for a given lag time τ is defined in Eq. (9.1). For a photon detection measurement with temporal resolution δt, the intensity values $I(t)$ within consecutive time intervals can only take the value $1/\delta t$ or 0, depending on whether there was a photon detection event during a time interval of width δt or not. The average in Eq. (9.1) is then calculated as the sum over all consecutive time intervals of width δt, divided by the total number of summed intervals. In practice, one does not compute the autocorrelation function for all possible values of lag time τ, but at increasingly spaced lag-time values. If the temporal resolution of the photon detection is, say, 100 ns, and one desires to follow correlation processes up to 1 min, possible values of lag time τ are any value between 100 ns and 60 sec in intervals of 100 ns, resulting in 6×10^8 possible lag-time values. Calculation of $g(\tau)$ for all of these values would be an enormously time-consuming numerical effort. Instead, the autocorrelation is calculated for only few, approximately logarithmically spaced values of τ.

A straightforward way of calculating the autocorrelation function is to divide the total measurement time, $t_N - t_1$, into intervals of unit length δt and to sort the detected photons into these intervals corresponding to their arrival times t_j. The result is a synchronous photon detection intensity file I_j with j running from 1 through $t_N - t_1$, where the I_j can only adopt the value 1 or 0. The fluorescence autocorrelation can then be calculated as given by Eq. (9.1). In practice, such an approach is prohibitively demanding of memory and computationally

expensive. An alternative, and much more efficient, FCS algorithm works directly on the arrival times $\{t_1, t_2, \ldots, t_N\}$, without converting them into time-binned data. For a given lag time τ, the algorithm searches for all photon pairs in the data stream that are a temporal distance τ apart from each other. The number of photon pairs with a distance τ is directly proportional to the autocorrelation value at lag time τ. The technical details of the algorithm are given Ref. 12.

9.4. Measuring Diffusion and Concentration

Thermally induced translational diffusion is one of the fundamental properties exhibited by molecules within a solution. Via the Stokes–Einstein relation it is directly coupled with the hydrodynamic radius of the molecules [13]. Any change in that radius will change the associated diffusion coefficient of the molecules. Such changes occur to most biomolecules—in particular proteins, RNA, and DNA—when interacting with their environment (e.g., binding of ions or other biomolecules), performing biologically important functions (such as enzymatic catalysis), or reacting to changes in environmental parameters such as pH, temperature, or chemical composition (like protein unfolding). Therefore, the ability to precisely measure diffusion coefficients has a large range of potential applications, such as monitoring conformational changes in proteins on ion binding or unfolding. However, many biologically relevant conformational changes are connected with rather small changes in hydrodynamic radius on the order of 0.1 nm (see, e.g., Ref. 14). To monitor these small changes, it is necessary to measure the diffusion coefficient with an accuracy of better than a few percent. Standard methods for diffusion coefficient measurements achieving this accuracy are dynamic light scattering (DLS) [15], pulsed-field-gradient nuclear magnetic resonance (NMR) [16], size exclusion electrophoresis [17], and analytical ultracentrifugation [18]. However, all of these methods operate at rather high sample concentrations, far away from the limit of infinite dilution. To obtain the correct infinite-dilution limit and thus a correct estimate of the hydrodynamic radius, one has often to measure at different concentrations and to extrapolate the concentration/diffusion coefficient curve towards zero concentration (see, e.g., Ref. 19). Another problem is that proteins are often prone to aggregation [20] at the concentrations needed for obtaining sufficient data quality. Thus, FCS is a relatively simple and attractive alternative for measuring diffusion coefficients, and the next sections will explain in detail how this is done.

9.4.1. One-Focus FCS

Following Eq. (9.1), the ACF is the correlation of the fluorescence intensity with a time-shifted replica of itself, calculated for all possible lag times τ. The measured signal $I(t)$ stems from the fluorescence of all the molecules within the sample plus uncorrelated background I_{bg} (light scattering, electronic noise, etc.):

$$I(t) = I_{bg} + \sum_j I_j(t) \tag{9.2}$$

where the index j refers to the jth molecule, and the summation runs over all molecules in the sample. Thus, the ACF $g(\tau)$ is given by

$$\begin{aligned} g(\tau) &= \left\langle \left(I_{bg}(t) + \sum_j I_j(t) \right) \left(I_{bg}(t+\tau) + \sum_k I_k(t+\tau) \right) \right\rangle \\ &= \sum_j \left\langle I_j(t) I_j(t+\tau) \right\rangle + \sum_{j \neq k} \bar{I}_j \bar{I}_k + + \sum_j \bar{I}_j \bar{I}_{bg} + \bar{I}_{bg}^2 \end{aligned} \tag{9.3}$$

where the angular brackets and bars denote averaging over all possible time values t. In the last line it was taken into account that fluorescence photons coming from different molecules are completely uncorrelated (no intermolecular interaction provided). Because all molecules in solution are indistinguishable, the last equation can be simplified further to

$$g(\tau) = N \langle i(t) i(t+\tau) \rangle + N(N-1) \langle i(t) \rangle^2 \tag{9.4}$$

where i is the measured fluorescence intensity of *any* molecule, and N is the total number of molecules present in the sample. Thus, the task of calculating the function $g(t)$ reduces to calculating $\langle i(t) i(t+\tau) \rangle$, the correlation of the fluorescence signal from one and the same molecule, and $\langle i(t) \rangle$, the average detected fluorescence intensity of one molecule.

The correlation $\langle i(t) i(t+\tau) \rangle$ of the fluorescence signal from a given molecule can be easily derived when remembering its physical meaning: It is proportional to the chance of seeing, from the same molecule, a photon at time $t + \tau$ if there was a photon detected at time t. The probability of finding a molecule within an infinitely small volume dV anywhere in the sample is equal to dV/V, where V is the total sample volume. Next, the probability of detecting a photon from a molecule at a given position $\mathbf{r}_0$ is directly proportional to the value of the MDF at this position, that is, to $U(\mathbf{r}_0)$. Furthermore, the chance that the molecule diffuses from position $\mathbf{r}_0$ to position $\mathbf{r}_1$ within time τ is given by the solution of the diffusion equation

$$\frac{\partial G}{\partial \tau} = D \Delta G \tag{9.5}$$

where Δ is the three-dimensional Laplace operator in coordinate $\mathbf{r}_1$, D is the diffusion coefficient of the molecule, and G approaches a three-dimensional Dirac function for $\tau \to 0$, $G(\mathbf{r}_1, \mathbf{r}_0, \tau = 0) = \delta(\mathbf{r}_1 - \mathbf{r}_0)$, that is, the molecule is exactly at position $\mathbf{r}_0$ at time zero. For a sample with far-removed boundaries, this solution is explicitly given by

$$G(\mathbf{r}_1, \mathbf{r}_0, \tau) \equiv G(\mathbf{r}_1 - \mathbf{r}_0, \tau) = \frac{1}{(4\pi D \tau)^{3/2}} \exp\left(-\frac{|\mathbf{r}_1 - \mathbf{r}_0|^2}{4D\tau} \right) \tag{9.6}$$

Finally, the chance of detecting a photon from the molecule at the new position $\mathbf{r}_1$ is again proportional to the value of the MDF at this position, that is, to $U(\mathbf{r}_1)$. Thus, the autocorrelation $\langle i(t) i(t+\tau) \rangle$ is calculated as the product of all of these individual contributions and averaging over all possible initial and final positions of the molecule, that is, integrating over $\mathbf{r}_0$ and $\mathbf{r}_1$:

$$\langle i(t)\, i(t+\tau) \rangle = \frac{1}{V} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_0 U(\mathbf{r}_1)\, G(\mathbf{r}_1, \mathbf{r}_0, t)\, U(\mathbf{r}_0) \tag{9.7}$$

Similarly, the average fluorescence intensity from a single molecule in the sample is given by

$$\langle i\,(t)\rangle = \frac{1}{V}\int\limits_V d\mathbf{r}U\,(\mathbf{r}) \tag{9.8}$$

so that the full ACF, in its most general form, reads

$$g\,(\tau) = c\int\limits_V d\mathbf{r}_1 \int\limits_V d\mathbf{r}_0 U\,(\mathbf{r}_1)\,G\,(\mathbf{r}_1,\mathbf{r}_0,t)\,U\,(\mathbf{r}_0) + \left[c\int\limits_V d\mathbf{r}U\,(\mathbf{r})\right]^2 \tag{9.9}$$

where c denotes the concentration of fluorescent molecules (numbers per volume) and one has used the fact that in the limit of large sample volume, $N/V \to c$ and $N(N-1)/V^2 \to c^2$

Equations (9.7) and (9.8) are of general validity, but before being able to apply them to the evaluation of real FCS experiments, one has to specify the MDF $U(\mathbf{r})$. The majority of publications on FCS adopt a very simple approximation of the MDF, assuming that it is well described by a three-dimensional Gaussian distribution, that is,

$$U(\mathbf{r}) = \kappa \exp\left[-\frac{2}{a^2}(x^2+y^2) - \frac{2}{b^2}z^2\right] \tag{9.10}$$

where κ is some overall constant, (x, y, z) are Cartesian coordinates centered at the intersection of focal plane and optical axis, with $x = y = 0$ being the optical axis, and a and b are the characteristic half-axes of the cylindrically symmetric, Gaussian-shaped detection volume. This corresponds to the lowest-order polynomial expansion of $\ln U(\mathbf{r})$ (due to axial and mirror symmetry, terms linear in x, y, z are absent). The characteristic parameters a and b are not known *a priori* and are usually determined by reference measurements on a sample with known diffusion coefficient. Using Eq. (9.10), we can now be explicitly calculate the single-molecule autocorrelation $\langle i(t)i(t+\tau)\rangle$ as

$$\begin{aligned}\langle i(t)i(t+\tau)\rangle &= \frac{c\varepsilon^2}{(4\pi Dt)^{3/2}}\int\limits_V d\mathbf{r}\int\limits_V d\tilde{\mathbf{n}}\,U(\mathbf{r}+\tilde{\mathbf{n}})\exp\left(-\frac{|\tilde{\mathbf{n}}|^2}{4Dt}\right)U(\mathbf{r}) \\ &= \frac{\pi^{3/2}}{8}\frac{c\varepsilon^2 a^2 b}{(1+4DT/a^2)\sqrt{1+4DT/b^2}}\end{aligned} \tag{9.11}$$

where ε is a constant factor taking into account overall detection efficiency of the measurement system, the absolute fluorescence brightness of the molecules (defined by absorption cross section and fluorescence quantum yield), and so on. In a similar way, the average fluorescence signal coming from one molecule is given by

$$\langle i(t)\rangle = c\varepsilon\int\limits_V d\mathbf{r}\,U(\mathbf{r}) = \left(\frac{\pi^3}{8}\right)^{1/2} c\varepsilon a^2 b \tag{9.12}$$

Thus, the final result for the total autocorrelation reads

$$g\,(\tau) = \left(\frac{\pi^{3/2}}{8}\right)\frac{c\varepsilon^2 a^2 b}{\left(1+4Dt/a^2\right)\sqrt{1+4Dt/b^2}} + \left[I_{bg} + \left(\frac{\pi^3}{8}\right)^{1/2} c\varepsilon a^2 b\right]^2 \tag{9.13}$$

where an additional background intensity I_{bg} has been included. An important property of the ACF is that the concentration of the fluorescent species can be derived from Eq. (9.13) via

$$\frac{g(\infty)}{g(0)-g(\infty)} = c\frac{\left[\int d\mathbf{r}U(\mathbf{r}) + I_{bg}\right]^2}{\int d\mathbf{r}U^2(\mathbf{r})} \tag{9.14}$$

where we have taken into account that $G(\boldsymbol{\rho},\tau)$ in Eq. (9.6) approaches a Dirac function for $\tau \to 0$. Using Eq. (9.14), one can define the *effective detection volume* V_{eff} as

$$V_{eff} = \frac{\left[\int d\mathbf{r}U(\mathbf{r})\right]^2}{\int d\mathbf{r}U^2(\mathbf{r})} \tag{9.15}$$

so that, for negligible background, the left-hand side of Eq. (9.14) equals cV_{eff}, that is, the mean particle number within V_{eff}. Thus, the ACF is often used for estimating concentrations of fluorescing molecules.

Although Eq. (9.13) is remarkably successful in fitting measured autocorrelation curves, the physical meaning of the parameters a and b is rather obscure because the actual MDF is usually much more complicated than that given by Eq. (9.10). The real shape of the MDF is only poorly described as a three-dimensional Gaussian (see Figure 9.2).

A more serious problem is that the exact form of the MDF is extremely sensitive to several optical and photophysical artifacts and can easily change from one measurement to another [21]. The most severe will be discussed here. The first common problem is that state-of-the-art water immersion objectives used in FCS setups are designed to image through a cover slide of a specific thickness. In this sense, the cover slide acts as the last optical element of an objective, and the optical quality of imaging (and laser focusing) critically depends on the exact matching between the cover slide thickness to which the objective is adjusted and its actual thickness. What happens when the cover slide thickness deviates from its design value by only a few micrometers is shown in Figure 9.3, where one can see the severe optical aberrations introduced by cover-slide-thickness mismatch and the resulting deformation of an ACF and the shift of its decay towards longer lag times. The enlargement of the MDF results in increased diffusion times, that is, apparently lower diffusion coefficients, and in an apparently increased concentration (there are more molecules present in the detection volume because the latter has become larger). In general, any aberration results in an increased detection volume and thus leads to the same trend of an apparently lower diffusion coefficient and higher concentration with increasing aberration. The effect on the apparent concentration is much stronger than on the apparent diffusion, resulting, for a cover-slide-thickness deviation of 10 μm, in an error of greater than 100% for the first and of roughly 30% for the second. It should be noted that the errors shown do not change significantly when changing the focus position in the solution.

This is in stark contrast to the effect of refractive index mismatch, which is considered next. An optical microscope using a water immersion objective is optimally corrected for imaging in water. However, in many biophysical applications, one has to work in buffer solutions with slightly different refractive indices. In addition, when measuring in cells or tissues, one faces similarly slight refractive index variations. Typical values of interest are between 1.333 and 1.360. Figure 9.4 shows the effect of refractive index mismatch on the MDF and ACF and subsequently on the apparent diffusion coefficient and concentration. The effect of even a slight refractive index mismatch is much more dramatic than that of cover

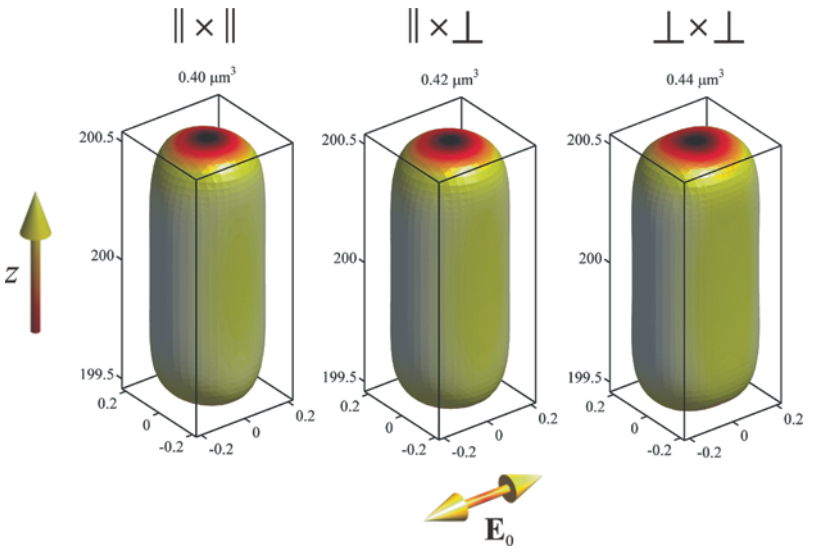


Figure 9.2. Shape of the molecule detection function (MDF) for three different detection and correlation schemes, visualized by displaying the isosurfaces where the MDF has fallen to $1/e^2$ ($\sim 13\%$) of its maximum value in the center. **Left.** The MDF of the autocorrelation of the detector signal monitoring fluorescence polarization parallel to the incident laser beam. **Right.** That of the detector signal monitoring fluorescence polarization perpendicular to the incident laser beam. **Middle.** The MDF for the cross-correlation between both detectors. The extreme case of completely anisotropic molecules was studied (where the maximum impact of polarization effects on the autocorrelation function [ACF] is expected). The vertical axis (z) is the optical axis; all units are given in micrometers; $z = 0$ is at the surface of the cover slide. Shown is also the polarization of the incident laser beam ($\mathbf{E}_0$). Shading of the isosurfaces indicates distance from the optical axis. The differences among the three parts of the figure are subtle, resulting in nearly identical ACF curves. Thus, in practice, one can safely neglect the effects of polarized detection when one is concerned with the diffusional part of an ACF [21].

slide thickness. This is mostly due to the large assumed distance of the focus position from the cover slide surface (200 μm, the default value of commercial instruments such as the Zeiss Confocor). In contrast to cover slide thickness, the aberrations introduced by refractive index mismatch accumulate with increasing distance of the focus from the cover slide surface because an increasingly thicker layer of solution with mismatched refractive index lies between the optics and the detection volume. The effect of refractive index mismatch can be much reduced by positioning the detection volume closer to the surface.

Another purely optical effect is laser beam astigmatism, that is, different focus positions within different axial planes of the laser beam. Astigmatism is easily introduced by slight curvatures of reflective elements in the optical setup (such as the dichroic mirror) or by slight axial asymmetry of the optical fiber that is often used for guiding the excitation light towards the objective.

The effect of astigmatism on the shape of the MDF and ACF, as well as on the apparent diffusion coefficient and chemical concentration, is shown in Figure 9.5. As can be seen, the effect of astigmatism on measured diffusion and concentration is of similar magnitude to that of cover-slide-thickness deviation. As for cover slide thickness, the effect of astigmatism is rather independent of focus position in the sample.

A particularly intriguing effect in FCS measurements is the dependence of the ACF on excitation intensity due to optical saturation. Optical saturation occurs when the excitation

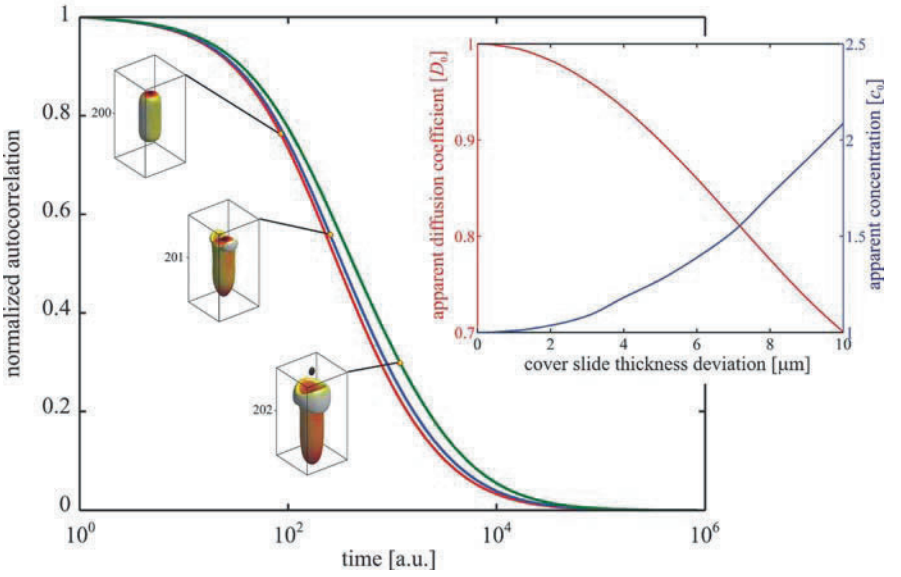


Figure 9.3. Left to right. The molecule detection function (MDF) and autocorrelation function (ACF) for three increasing values of cover-slide-thickness deviation, $\delta = 0$, $\delta = 5$ μm, and $\delta = 10$ μm, respectively. The box size of the MDF displays is $1 \times 1 \times 2\ \mu m^3$; the number next to the box gives the center position along the optical axis in micrometers. Note the shift of the center of the MDF along the optical axis for increasing values of δ. The inset shows the dependence of the apparent diffusion coefficient and the chemical concentration on the thickness deviation value. These values would be obtained when performing a comparative fluorescence correlation spectroscopy measurement using an ideal ACF ($\delta = 0$) as reference (same for the following figures).

intensity becomes so large that a molecule spends more and more time in a nonexcitable state, so that increasing the excitation intensity does not lead to a proportional increase in emitted fluorescence intensity [22]. The most common sources of optical saturation are (1) excited-state saturation, that is, the molecule is still in the excited state when the next photon arrives; (2) triplet-state saturation, that is, the molecule undergoes intersystem crossing from the excited to the triplet state so that it can no longer become excited until it returns back to the ground state; and (3) other photo-induced transitions into a nonfluorescing state, such as the photo-induced *cis–trans* isomerization in cyanine dyes or the optically induced dark states in quantum dots. The exact relation between fluorescence emission intensity and excitation intensity can be very complex [22] and even dependent on the excitation mode (pulsed or continuous wave), but a sufficiently good approximation of the dependence of fluorescence intensity on excitation intensity is given by the simple relation

$$I_{fluo} \propto \frac{I_{exc}}{1 + I_{exc}/I_{sat}} \tag{9.16}$$

where I_{fluo} and I_{exc} are the fluorescence and excitation intensity, respectively, and I_{sat} is a parameter called the saturation intensity, which describes the saturation behaviour of a given dye. Figure 9.6 shows how optical saturation changes the shape of the MDF and ACF and the apparent diffusion coefficient and concentration. An important feature is the behaviour of the curves of apparent diffusion and concentration in the limit of vanishing excitation intensity:

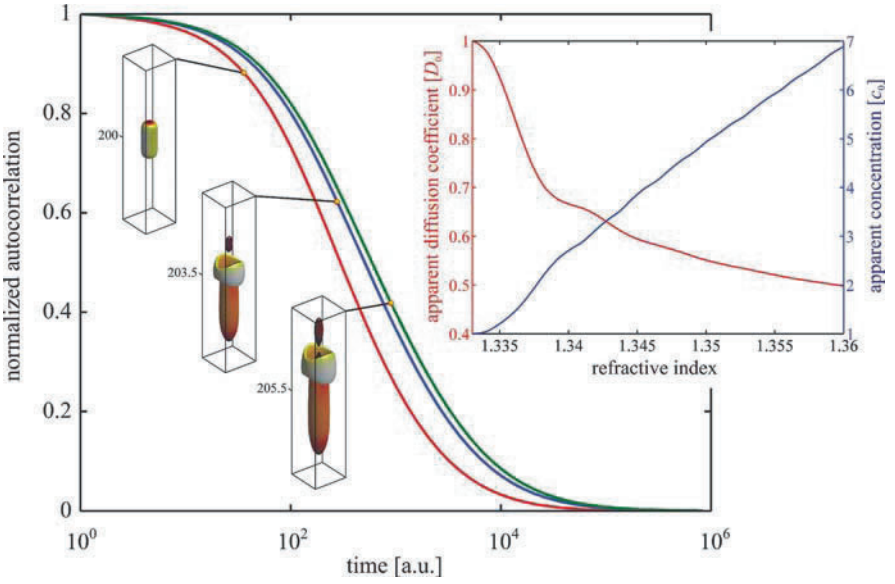


Figure 9.4. Left to right. The molecule detection function (MDF) and autocorrelation function (ACF) for three increasing values of refractive index of the sample solution, $n_m = 1.333$, $n_m = 1.346$, and $n_m = 1.360$, respectively. The box size of the MDF displays is $1 \times 1 \times 5\ \mu m^3$. Note again the shift of the center of the MDF along the optical axis for increasing values of n_m. The inset shows the dependence of the apparent diffusion coefficient and the concentration on refractive index.

Whereas for all optical effects studied previously the slope of these curves tended to zero for vanishing aberration (or astigmatism), its absolute value now is largest at zero intensity.

To better understand the reason for that behaviour, consider an ideal Gaussian excitation profile $I_0 \exp(-x^2/2\sigma^2)$ with mean square deviation of 1. Figure 9.7 shows the widening of such a profile when transformed by a saturation to $I_0 \exp(-x^2/2\sigma^2)/[1 + I_0 \exp(-x^2/2\sigma^2)]$. As can be seen, relative change in profile width is fastest in the limit of zero intensity $I_0 \to 0$, explaining why one sees most of the changes in FCS at low saturation levels.

All of these effects make a quantitative evaluation of standard FCS measurements quantitatively unreliable. As pointed out before, the core problem is the absence of an extrinsic and fixed length scale in the experiment. Even referential measurements, that is, using a dye with known diffusion coefficient for determining the parameters a and b and then using them to measure the diffusion of a sample, can be problematic due to the strong dependence of an FCS result on optical saturation, which is itself determined, in a complex manner, by the photophysics of a particular dye. Even worse, the photophysical parameters of the same dye can change on binding it to a protein or other target molecule. The next section describes a recent modification of the standard FCS measurement that seems to solve this long-standing problem and that allows for reproducible, quantitative, and absolute measurements of diffusion coefficients.

9.4.2. Dual-Focus FCS

Recently, a new and straightforward modification of FCS was developed [23], namely dual-focus FCS (2fFCS), that fulfils two requirements: (1) it introduces an external ruler

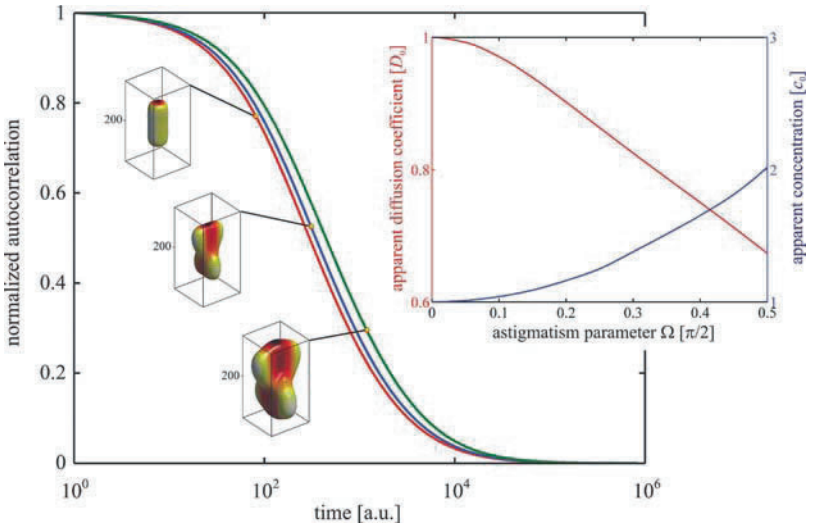


Figure 9.5. Left to right. The molecule detection function (MDF) and autocorrelation function (ACF) for three increasing values of laser beam astigmatism (for definition of astigmatism see Ref. 21). The box size of the MDF-displays is $1 \times 1 \times 2\ \mu m^3$. There is no shift of the center of the MDF along the optical axis for increasing values of astigmatism. The inset shows the dependence of the apparent diffusion coefficient and the concentration on beam astigmatism.

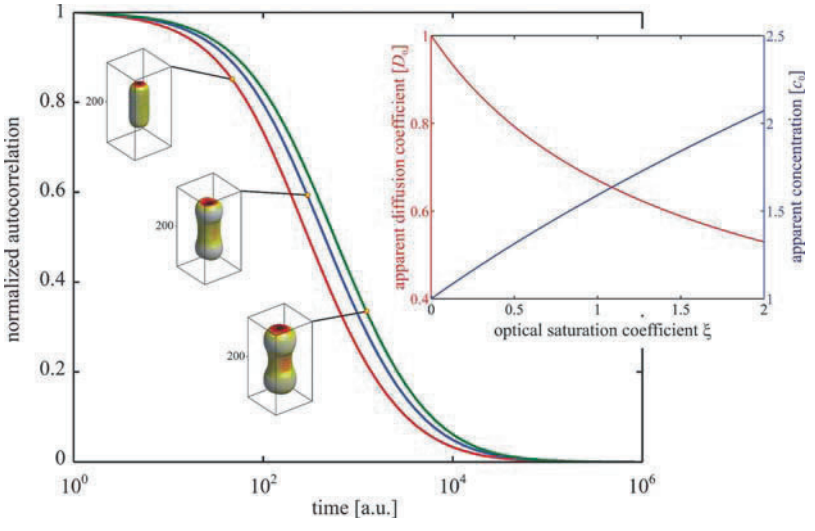


Figure 9.6. Left to right. The molecule detection function (MDF) and autocorrelation function (ACF) for three increasing values of optical saturation, $\zeta = 0$, $\zeta = 1$, and $\zeta = 2$, respectively. The box size of the MDF displays is $1 \times 1 \times 2\ \mu m^3$. The inset shows the dependence of apparent diffusion coefficient and concentration on optical saturation, that is, excitation intensity.

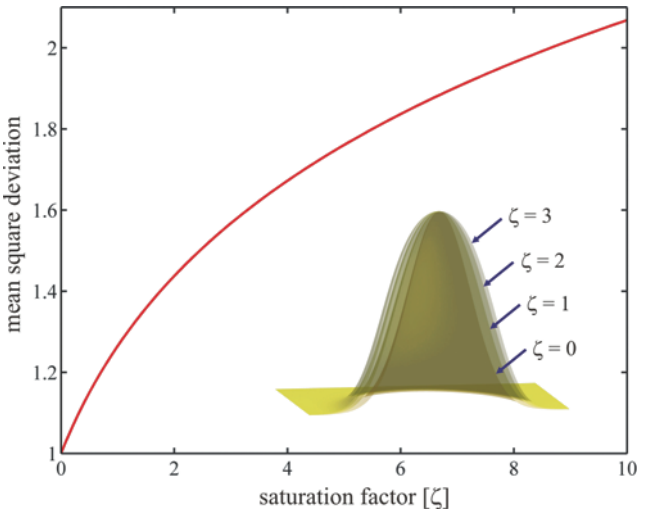


Figure 9.7. Change of the mean square deviation of the distribution $\exp(-x^2/2)/[1 + \zeta \exp(-x^2/2)]$ with increasing value of ζ.

into the measurement by generating two overlapping laser foci of precisely known and fixed distance, and (2) it generates the two foci and corresponding detection regions in such a way that the corresponding MDFs are sufficiently well described by a simple two-parameter model yielding accurate diffusion coefficients when applied to 2fFCS data analysis. Both of these properties allow for measuring absolute values of the diffusion coefficient with an accuracy of a few percent. Moreover, the new technique is robust against refractive index mismatch and optical saturation effects, which are troubling to standard FCS measurements.

The 2fFCS setup, as shown in Figure 9.8, is based on a standard confocal epifluorescence microscope as was shown in Figure 9.1. However, instead of using a single excitation laser, the light of two identical, linearly polarized pulsed diode lasers is combined by a polarizing beam splitter. Both lasers are pulsed *alternately* with a high repetition rate (ca. 40–80 MHz), an excitation scheme called pulsed interleaved excitation (PIE) [24]. Both beams are then coupled into a polarization-maintaining single-mode fiber. At the output, the light is again collimated. Thus, the combined light consists of a train of laser pulses with alternating orthogonal polarization. The beam is then reflected by a dichroic mirror towards the microscope's water-immersion objective, but before entering the objective, the light beam is passed through a Nomarski prism that is normally exploited for differential interference contrast (DIC) microscopy. The Nomarski prism is an optical element that deflects the laser pulses into two different directions according to their corresponding polarization, as shown in Figure 9.9. Thus, after focusing the light through the objective, two overlapping excitation foci are generated, with a small lateral shift between them. The distance between the beams is uniquely defined by the chosen Nomarski prism and is independent of the sample's refractive index, cover slide thickness, and laser beam astigmatism because all of these properties may introduce severe aberrations, although they will not change the main distance between the axes of propagation of both focused laser beams.

As in one-focus FCS, the generated fluorescence is collected by the same objective, passed through the Nomarski prism and the dichroic mirror, and focused onto a single circular

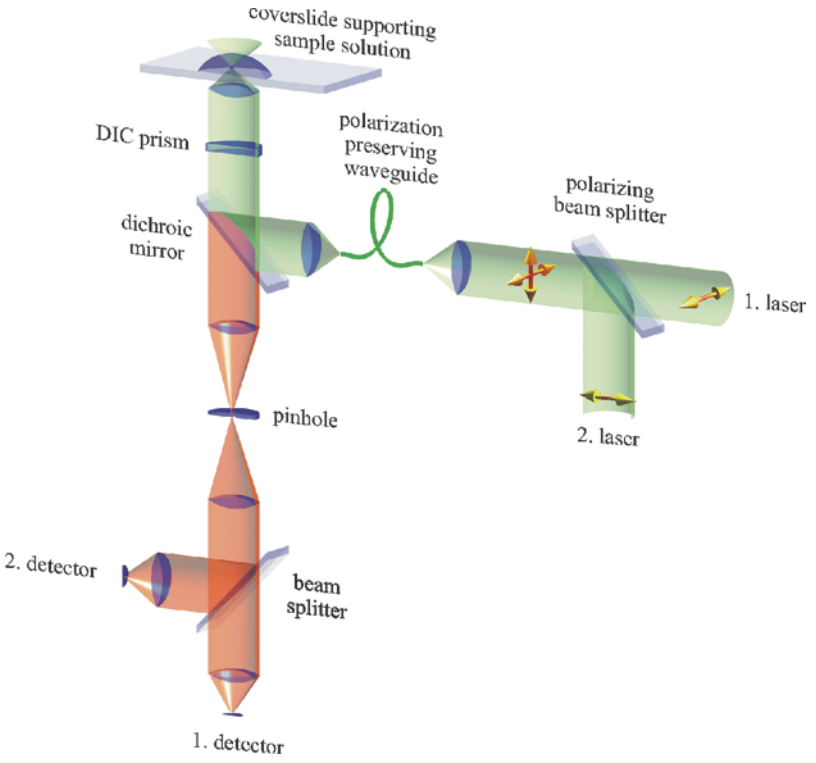


Figure 9.8. Schematic of the dual-focus fluorescence correlation spectroscopy (2fFCS) setup. Excitation is done by two interleaved pulsed lasers of the same wavelength. The polarization of each laser is linear, but they are orthogonal to each other. Light is then combined by a polarizing beam splitter and coupled into a polarization-maintaining optical single-mode fiber. After exiting the fiber, the laser light is collimated by an appropriate lens and reflected by a dichroic beam splitter through a differential interference contrast (DIC) prism. The DIC prism separates the laser light into two beams according to the polarization of the incoming laser pulses. The microscope objective focuses the two beams into two laterally shifted foci. Fluorescence is collected by the same objective. The tube lens focuses the detected fluorescence from both excitation foci on a single pinhole. Subsequently, the fluorescence light is split by a 50/50 beam splitter and detected by two single-photon avalanche diodes.

aperture (diameter 200 μm), which is positioned symmetrically with respect to both focus positions and chosen to be large enough to easily pass the light from both foci. After the pinhole, the light is collimated, split by a nonpolarizing beam splitter cube, and focused onto two SPADs. A single-photon counting electronics unit is used to record the detected photons from both SPADs with picosecond temporal resolution. The picosecond temporal resolution is used to decide which laser has excited which fluorescence photon, that is, within which laser focus/detection volume the light was generated. This is done by correlating the detection time of each photon with time of the last preceding laser pulse. In the data evaluation, all photons that fall into the first time window are associated with the first laser, and all photons that fall into the second time window are associated with the second laser. For a successful working of that method it is, of course, necessary that the time between laser pulses be significantly larger than the fluorescence lifetime of the fluorescent molecules. By knowing which photon was generated in which detection volume, one can calculate the ACFs for each detection volume, as well as the cross-correlation function (CCF) between the two detection volumes. The CCF

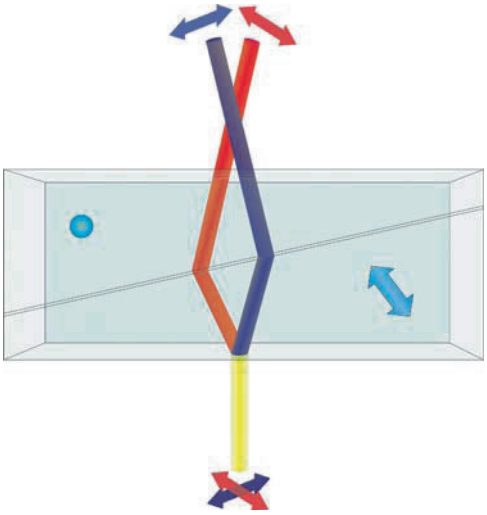


Figure 9.9. Working principle of a Nomarski prism: The prism consists of two optically birefringent crystal wedges cemented together. The bottom wedge is modified by cutting the crystal in such a manner that the optical axis is oriented obliquely with respect to the flat surface, whereas the top wedge has the optical axis oriented parallel to the surface of the prism (light blue double arrows within the prism body). When a light beam with mixed polarization enters the prism from the bottom, it is split into two orthogonal polarizations, which are deflected into two different directions. Focusing the two deflected beams with the microscope's objective results in two laterally shifted but overlapping foci within the sample.

is calculated in a similar way as the ACF but correlating photons only from different detection volumes. The CCF at lag time τ is thus proportional to the chance of seeing a photon from the second detection volume at any time $t + \tau$ if there was a detection event from the first detection volume at time t and vice versa.

A crucial point for a successful 2fFCS data analysis is to have a sufficiently appropriate model function for the MDF. It was found [23] that a suitable expression is given by

$$U(\mathbf{r}) = \frac{\kappa(z)}{w^2(z)} \exp\left[-2\frac{x^2+y^2}{w^2(z)}\right] \tag{9.17}$$

where $\kappa(z)$ and $w(z)$ are functions of the axial coordinate z (optical axis) defined by

$$w(z) = w_0 \left[1 + \left(\frac{\lambda_{ex} z}{\pi w_0^2 n}\right)^2\right]^{1/2} \tag{9.18}$$

and

$$\kappa(z) = 2\int_0^a \frac{d\rho\rho}{R^2(z)} \exp\left[-\frac{2\rho^2}{R^2(z)}\right] = 1 - \exp\left[-\frac{2a^2}{R^2(z)}\right] \tag{9.19}$$

where $R(z)$ is defined by an expression similar to Eq. (9.18):

$$R(z) = R_0 \left[1 + \left(\frac{\lambda_{em} z}{\pi R_0^2 n} \right)^2 \right]^{1/2} \tag{9.20}$$

In these equations, λ_{ex} is the excitation wavelength and λ_{em} the center emission wavelength, n is the refractive index of the immersion medium (water), a is the radius of the confocal aperture divided by magnification, and w_0 and R_0 are two (generally unknown) model parameters. Equation (9.18) is nothing else than the scalar approximation for the radius of a diverging laser beam with beam waist radius w_0. Equation (9.17) is a modification of the three-dimensional Gaussian we have already met when discussing one-focus FCS and says that in each plane perpendicular to the optical axis, the MDF is approximated by a Gaussian distribution having width $w(z)$ and amplitude $\kappa(z)/w^2(z)$.

It remains to calculate the auto- and cross-correlation curves of the two-focus setup. One derives these expressions following a similar philosophy of calculating the photon detection and diffusion probabilities as in the previous section. For example, the cross-correlation function between the fluorescence signal coming from the two different detection volumes is given by a similar integral as that in Eq. (9.13)

$$\begin{aligned} g(\tau,\delta) = {} & \frac{c\varepsilon_1\varepsilon_2}{(4\pi D t)^{3/2}} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_0 U(\mathbf{r}_1) \exp\left(-\frac{|\mathbf{r}_1 - \mathbf{r}_0 + \delta\hat{\mathbf{x}}|^2}{4Dt} \right) U(\mathbf{r}_0) \\ & + \varepsilon_1\varepsilon_2 \left[c \int_V d\mathbf{r} U(\mathbf{r}) \right]^2 \end{aligned} \tag{9.21}$$

Here, we have taken into account that the MDFs of both detection volumes are identical but shifted by a distance δ along the x axis (along unit vector $\hat{\mathbf{x}}$) and having potentially two different overall detection efficiencies ε_1 and ε_2. Inserting Eqs. (9.17)–(9.20) into Eq. (9.21) yields </YIΠ>

$$\begin{aligned} g(t,\delta) = g_\infty(\delta) + \frac{\varepsilon_1\varepsilon_2 c}{4} \sqrt{\frac{\pi}{Dt}} \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2 \\ \frac{\kappa(z_1)\kappa(z_2)}{8Dt + w^2(z_1) + w^2(z_2)} \exp\left[-\frac{(z_2 - z_1)^2}{4Dt} - \frac{2\delta^2}{8Dt + w^2(z_1) + w^2(z_2)} \right] \end{aligned} \tag{9.22}$$

which is certainly more complicated than the simple expression of the second line in Eq. (9.13) but not much harder to handle numerically in the age of powerful personal computers. For numerical purposes, it is useful to slightly modify this result by changing the variables to

$$\xi = \frac{z_2 - z_1}{2\sqrt{Dt}} \text{ and } \eta = \frac{z_2 + z_1}{2} \tag{9.23}$$

leading to the expression

$$g(t,\delta) = g_\infty(\delta) + 2\varepsilon_1\varepsilon_2 c\sqrt{\pi}\int_0^\infty d\xi \int_0^\infty d\eta \frac{\kappa(\eta - \sqrt{Dt}\xi)\kappa(\eta + \sqrt{Dt}\xi)}{8Dt + w^2(\eta - \sqrt{Dt}\xi) + w^2(\eta + \sqrt{Dt}\xi)} \times \exp\left[-\xi^2 - \frac{2\delta^2}{8Dt + w^2(\eta - \sqrt{Dt}\xi) + w^2(\eta + \sqrt{Dt}\xi)}\right] \tag{9.24}$$

Because w and κ are rapidly decaying functions for large argument, the infinite integrations over η and ζ can be approximated by numerically evaluating the integrals within a finite two-dimensional strip defined by $|\eta \pm (Dt)^{1/2}\zeta| < M$, where M is a truncation value chosen in such a way that the numerical integration result does not change when increasing M further. Numerical integration can be done by a simple finite-element scheme, and convergence is checked by testing whether the numerical result remains the same on refining the finite element size and when increasing the threshold value M.

Data fitting is usually performed with least-squares fitting of the model curve, Eq. (9.24), against the measured ACF ($\delta = 0$, with $\varepsilon_1\varepsilon_2$ replaced by either ${\varepsilon_1}^2$ or ${\varepsilon_2}^2$) and cross-correlation (CCF) *simultaneously* in a global fit. As fit parameters one has $\varepsilon_1 c^{1/2}$, $\varepsilon_2 c^{1/2}$, D, w_0, and R_0, as well as three offset values g_∞. The distance δ between the detection regions is determined by the properties of the Nomarski prism and has to be exactly known *a priori*, thus introducing an external length scale into data evaluation. An elegant and effective way of determining this distance is to perform a comparative measurement of the diffusion of fluorescently labelled beads with dynamic light scattering and with 2fFCS [25]. Because both methods yield the same value of the diffusion coefficient, one can use the comparison to retrieve the correct interfocal distance. However, to avoid the introduction of systematic errors due to the finite size of the beads, one should make sure that bead size is less than ~100 nm in diameter [26].

A typical measurement result or 2fFCS experiment together with the fitted curves using Eq. (9.24) is shown in Figure 9.10. It is important to notice that a crucial criterion of fit quality is not only simultaneously to reproduce the temporal shape of both ACFs and the cross-correlation function, but also to reproduce their three amplitudes $g_{t\to 0} - g_\infty$ using only the two parameters $\varepsilon_1 c^{1/2}$ and $\varepsilon_2 c^{1/2}$. The relation between the amplitudes of the CCF and the amplitudes of the ACFs is determined by the overlap between the two MDFs, and thus by the shape parameters w_0 and R_0. Thus, achieving a good fit quality for the relative amplitudes of the ACFs and the CCF strongly helps to find the correct values of these parameters.

Due to the presence of an external length scale determined by the distance δ between the detection volumes and a reasonably accurate model of the MDF, 2fFCS is indeed a method of superior accuracy and stability for measuring diffusion. An optimal distance between foci is equal to their radius in the focal plane, giving a sufficiently large overlap between detection volumes that the amplitude of the cross-correlation function between both detection volumes is roughly one half of the amplitude of each autocorrelation function. Larger distances will lead to significantly longer measurement times for accumulating a sufficiently good cross-correlation and smaller distances will lead to a cross-correlation function too similar to the autocorrelation functions, so that data fitting becomes unreliable.

As was shown in Ref. 21, the achievable accuracy of 2fFCS is better than 5% in absolute value for diffusion measurements. The method has been used to measure and remeasure the diffusion of several dyes in water throughout the visible spectrum [27,28]. The values are reported in Table 9.1. Remarkably, it has been found that the value of the diffusion coefficient of rhodamine 6G, which has served for many years as the "gold standard" for calibrating conventional FCS measurements, is larger by 37% than reported in the literature. However,

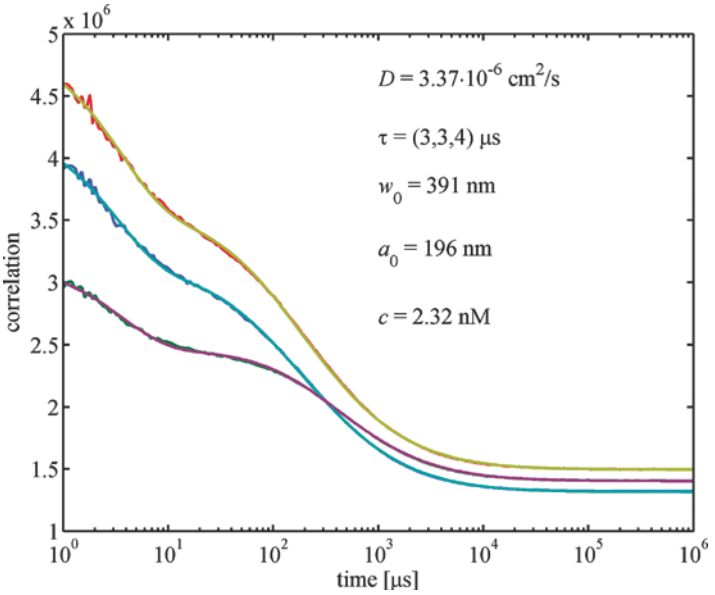


Figure 9.10. Typical result of a dual-focus fluorescence correlation spectroscopy (2fFCS) measurement for a nanomolar concentration of the dye Alexa633 in water. Shown are the two autocorrelation functions (ACFs) for each focus (*red/yellow*, *cyan/blue*), as well as the cross-correlation function (CCF) (*green/magenta*). The slightly different values of the ACF amplitudes are due to slightly different excitation intensities of the two lasers. However, the shape of both ACFs is identical. One clearly sees also the fast exponential decays ($\tau = 3$, 3, and 4 μs, respectively) of both ACFs and the CCF on the microsecond time scale due to the triplet-state dynamics of the dye. The fit returns, besides an absolute value of the diffusion coefficient D, also values for the beam waist parameter w_0 and the confocal pinhole parameter a_0 [R_0 in Eq. (9.20)], which can the be used to estimate the detection volume and thus the concentration c of the dye.

Table 9.1. Reference diffusion coefficients of various dyes in aqueous solution

Dye	λ_{abs} (nm)	λ_{abs} (nm)	$D_{25°C}$ (cm^2/sec)
Cy5	650	670	$(3.7 \pm 0.15) \times 10^{-6}$
Atto655-COOH	665	690	$(4.26 \pm 0.08) \times 10^{-6}$
Atto655-maleimide	665	690	$(4.07 \pm 0.1) \times 10^{-6}$
Rhodamine 6G	530	560	$(4.14 \pm 0.05) \times 10^{-6}$
Oregon green	488	540	$(4.11 \pm 0.06) \times 10^{-6}$

it should be noted that recently several groups have found similarly large values by using alternative measurement techniques [29,30].

9.5. Conclusion and Outlook

This chapter has mainly considered the application of FCS to diffusion and concentration measurements, which are the most difficult applications in terms of accuracy, reproducibility, and robustness. After discussing the most relevant optical and photophysical

artifacts that are problems for standard one-focus FCS, the chapter described the generalized method of 2fFCS, which, for the first time since the invention of FCS, yields accurate and precise results for diffusion coefficients in absolute terms at the infinite-dilution limit. However, it should be noted that even that method, using the improved representation of the MDF as given by Eq. (9.17), yields an estimation of the detection volume (and thus concentration) that still can be inaccurate by more than 100% in absolute terms. Thus, we still lack a simple and efficient correlation-based method for precise *absolute* concentration measurements in solutions at the pico- to nanomolar concentration level.

However, for *absolute* measurements of diffusion coefficients, 2fFCS is certainly much superior and more accurate than conventional FCS, whereas the technical cost of modifying an existing FCS into a 2fFCS system is rather moderate. The achievable accuracy of 2fFCS in determining a diffusion coefficient is better than 4% in absolute numbers. The first commercial version of a 2fFCS system is already available.

This chapter has completely omitted the application of FCS and related techniques to measure intramolecular dynamics, intermolecular interactions, stoichiometry, or rotational diffusion, a thorough discussion of which is beyond its scope. Moreover, in all of these other applications, the optical problems and artifacts discussed here are not of great concern because the time scale of, for example, intramolecular conformations or rotational diffusion is much faster than the time scale of the diffusive motion within the detection volume, so that the size and shape of the detection volume usually have no any influence on the correlation curve at these short times. The reader interested in these other applications of FCS is referred to the excellent reviews and books cited in the introduction.

References

1. D. Magde, E. Elson, W. W. Webb, "Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy." *Phys. Rev. Lett.* **29** (1972) 705–8.
2. E. L. Elson, D. Magde, "Fluorescence correlation spectroscopy. I. Conceptual basis and theory." *Biopolymers* **13** (1974) 1–27.
3. D. Magde, E. Elson, W. W. Webb, "Fluorescence correlation spectroscopy. II. An experimental realization." *Biopolymers* **13** (1974) 29–61.
4. N. O. Petersen, "Scanning fluorescence correlation spectroscopy. I. Theory and simulation of aggregation measurements." *Biophys. J.* **49** (1986) 809–15.
5. N. O. Petersen, D. C. Johnson, M. J. Schlesinger, "Scanning fluorescence correlation spectroscopy. II. Application to virus glycoprotein aggregation." *Biophys. J.* **49** (1986) 817–20.
6. J. Widengren, Ü. Mets, "Conceptual basis of fluorescence correlation spectroscopy and related techniques as tools in bioscience." In: *Single-Molecule Detection in Solution—Methods and Applications*. C. Zander, J. Enderlein, R. A. Keller, eds. (Wiley-VCH, Berlin, 2002), pp. 69–95.
7. P. Schwille, "Fluorescence correlation spectroscopy and its potential for intracellular applications." *Cell. Biochem. Biophys.* **34** (2001) 383–408.
8. S. T. Hess, S. Huang, A. A. Heikal, W. W. Webb, "Biological and chemical applications of fluorescence correlation spectroscopy: A Review." *Biochem.* **41** (2002) 697–705.
9. R. Rigler, E. Elson (eds.), *Fluorescence Correlation Spectroscopy* (Springer, Berlin, 2001).
10. M. Böhmer, J. Enderlein, "Single molecule detection on surfaces with the confocal laser scanning microscope." In: *Single-Molecule Detection in Solution— Methods and Applications*, C. Zander, J. Enderlein, R. A. Keller, eds. (Wiley-VCH, Berlin, 2002), pp. 145–183.
11. J. Enderlein, F. Pampaloni, "Unified operator approach for deriving Hermite-Gaussian and Laguerre-Gaussian laser modes." *J. Opt. Soc. Am. A* **21** (2004) 1553–1558.
12. M. Wahl I. Gregor, M. Patting, J. Enderlein, "Fast calculation of fluorescence correlation data with asynchronous time-correlated single-photon counting." *Opt. Expr.* **11** (2003) 3583–3591.

13. A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover, New York, 1985).
14. A. M. Weljie, A. P. Yamniuk, H. Yoshino, Y. Izumi, H. J. Vogel, "Protein conformational changes studied by diffusion NMR spectroscopy: Application to helix-loop-helix calcium binding proteins." *Protein Sci.* **12** (2003) 228–236.
15. B. J. Berne, R. Pecora, *Dynamic Light Scattering* (Dover, New York, 2000).
16. P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy* (Clarendon Press, Oxford, 1991).
17. D. Harvey, *Modern Analytical Chemistry* (McGraw-Hill, Boston, 2000), pp. 593–595.
18. J. L. Cole, and Hansen J. C., "Analytical ultracentrifugation as a contemporary biomolecular research tool." *J. Biomol. Tech.* **10** (1999) 163–176.
19. W. Liu, T. Cellmer, D. Keerl, J. M. Prausnitz, H. W. Blanch, "Interactions of lysozyme in guanidinium chloride solutions from static and dynamic light-scattering measurements." *Biotechnol. Bioeng.* **90** (2005) 482–490.
20. T. Kiefhaber, R. Rudolph, H. H. Kohler, J. Buchner, "Protein aggregation *in vitro* and *in vivo*: A quantitative model of the kinetic competition between folding and aggregation." *Nature Biotechnol.* **9** (1991) 825–829.
21. J. Enderlein, I. Gregor, D. Patra, T. Dertinger, B. Kaupp, "Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration." *ChemPhysChem* **6** (2005) 2324–2336.
22. J. Enderlein, "Dependence of the optical saturation of fluorescence on rotational diffusion." *Chem. Phys. Lett.* **410** (2005) 452–456.
23. T. Dertinger, V. Pacheco, I. von der Hocht, R. Hartmann, I. Gregor, J. Enderlein, "Two-focus fluorescence correlation spectroscopy: A new tool for accurate and absolute diffusion measurements." *ChemPhysChem* **8** (2007) 433–443.
24. B. K. Müller, E. Zaychikov, C. Bräuchle, D. C. Lamb, "Cross talk free fluorescence cross correlation spectroscopy in live cells." *Biophys. J.* **89** (2005) 3508–3522.
25. C. B. Müller, K. Weiss, W. Richtering, A. Loman, J. Enderlein, "Calibrating differential interference contrast microscopy with dual-focus fluorescence correlation spectroscopy." *Opt. Expr.* **16** (2008) 4322–4329.
26. C. B. Müller, A. Loman, W. Richtering, J. Enderlein, "Dual-focus fluorescence correlation spectroscopy of colloidal solutions: influence of particle size." *J. Phys. Chem. B* **112** (2008) 8236–8240.
27. A. Loman, T. Dertinger, F. Koberling, J. Enderlein, "Comparison of optical saturation effects in conventional and dual-focus fluorescence correlation spectroscopy." *Chem. Phys. Lett.* **459** (2008) 18–21.
28. C. B. Müller, A. Loman, V. Pacheco, F. Koberling, D. Willbold, W. Richtering, J. Enderlein, "Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy." *Eur. Phys. Lett.* **83** (2008) 46001.
29. C. T. Culbertson, S. C. Jacobson, J. M. Ramsey, "Diffusion coefficient measurements in microfluidic devices." *Talanta* **56** (2002) 365–373.
30. Z. Petrasek, P. Schwille, "Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy." *Biophys. J.* **94** (2008) 1437–1448.

10

Single-Molecule Studies of Nucleic Acid Interactions Using Nanopores

Meni Wanunu, Gautam V. Soni, and Amit Meller

Abstract This chapter presents biophysical studies of single biopolymers using nanopores. Starting from the fundamental process of voltage-driven biopolymer translocation, the understanding of which is a prerequisite for virtually all nanopore applications, the chapter describes recent experiments that resolve nucleic acid structure and its interaction with enzymes, such as exonucleases and polymerases. It then outlines progress made with solid-state nanopores fabricated in ultrathin membranes and discusses experiments describing biopolymer dynamics in synthetic pores. The chapter concludes with a discussion on some of the main challenges facing nanopore technology, as well as on some of the future prospects associated with nanopore-based tools.

10.1. Introduction

Nanopores are an emerging class of single-molecule (SM) sensors optimized for probing the structure of biopolymers and are undoubtedly among the most attractive tools in the SM field. The principle behind the technique is simple and is primarily based on monitoring ion current flowing through the nanopore during entry and exit of analyte molecules. When the nanopore diameter is similar to the biopolymer cross section and the nanopore thickness is much smaller than the biopolymer length, nanopores can be used to probe fine structural properties of extremely long biopolymers. This technique offers a unique combination of high-throughput detection, no requirement for analyte labeling/immobilization, and nondestructive detection (in most cases). A look at recent nanopore literature reveals a broad range of current and prospective nanopore studies, including molecular identification, DNA genotyping, DNA sequencing, nanopore force spectroscopy, protein characterization, and the probing of DNA–protein interactions. Moreover, recent breakthroughs with synthetic nanopores have catalyzed efforts for commercial realization of some of these applications, as well as a much broader scope of applications, due to the size, physical robustness, and surface tunability of synthetic materials.

M. Wanunu, G.V. Soni, and A. Meller • Department of Biomedical Engineering and Department of Physics, Boston University, Boston, MA 02215, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_10,

The process of biopolymer transport through small pores is not uncommon in biology. Numerous natural processes involve biopolymer transport across large channels embedded in the plasma or nuclear membranes (Schatz and Dobberstein, 1996). Examples include gene transfer by transduction (Driselkelmann, 1994; Kokjohn, 1989), gene swapping through bacterial pili (Ippen-Ihler, 1989; Miller, 1998), RNA and transcription factor transport through nuclear pore complexes (Bustamante et al., 1995; Kasamatsu and Nakanishi, 1998; Salman et al., 2001; Whittaker and Helenius, 1998), the uptake of oligonucleotides by some membrane proteins (Hanss et al., 1998), and ejection of viral DNA (Grayson et al., 2007). These biological processes have inspired *in vitro* nanopore experiments using voltage-driven transport of biomolecules, as discussed in this review. From an experimental standpoint, the analysis of fast-moving biopolymers presents a major challenge with nanopores, as addressed by several groups. The voltage-driven dynamics of nucleic acids through nanopores has been a subject of numerous experimental and theoretical studies and has been the focus of some reviews (Meller, 2003). Nanopores has also been extensively used as biochemical sensors (Bayley and Cremer, 2001). This chapter focuses on one specific area in this field, namely the applicability of nanopores for biophysical studies, beginning with a brief historical perspective, then a discussion of some key experimental works, and concluding with some of the main current challenges and future prospects.

10.2. Nanopore Basics

Resistive sensing is based on particle counting and sizing in a conducting fluid, a method invented by Coulter in the early 1950s (Coulter, 1953) and later refined by DeBlois and Bean (1970). In the Coulter counter, small particles are pressured through a micron-scale aperture made in an insulating wall (e.g., glass). A voltage is applied across the aperture to generate ionic current in the electrolyte. As particles pass through the aperture, they temporarily displace electrolyte in the aperture volume, causing transient blockades in the ionic current. Since the magnitude of these current blockades is roughly proportional to the volume of each particle, particles (e.g., living cells) can be nondestructively counted and sized using this method.

Nanopores are single-molecule Coulter counters, consisting of a several-nanometer aperture through an insulating membrane. Application of voltage across the membrane results in an electrochemical gradient that drives ions through the nanopore (Figure 10.1A), giving rise to a measurable, steady current. Entry of molecules into the nanopore results in fluctuations in the ion current, which correlate with molecular properties (e.g., size and charge) and reflects its dynamics as it translocates from the *cis* side to the *trans* side of the membrane (Figure 10.1B). While detection of small molecules (e.g., trinitrotoluene) (Guan et al., 2005) and polymers (such as polyethylene glycol [PEG]) (Bezrukov et al., 1994) has been demonstrated, this review focuses on the detection of charged biopolymers, such as nucleic acids and proteins, which are actively driven by the applied voltage. The far-reaching impact of nanopores was stimulated in the 1990s by the realization that serial passage of a sequence-encoded biopolymer (e.g., DNA) through nanopores can be decoded by analyzing its characteristic current signal (Deamer and Branton, 2002). Indeed, sequential reading of a biopolymer as it translocates through nanopores is a highly attractive and novel concept that, to a first approximation, mimics natural biological machinery. This goal is quite ambitious, and the attractive potential of fast DNA (and for that matter protein) sequencing using nanopores

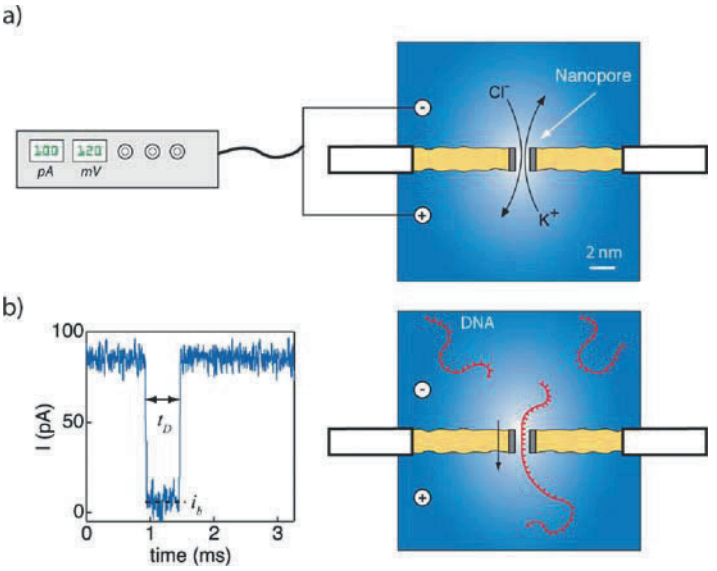


Figure 10.1. Schematic description of the nanopore method. **A.** Application of voltage across an insulating membrane in which a single nanopore is punctured results in a steady ion current measured by a sensitive electrometer. **B.** Charged biopolymers are threaded through the nanopore from *cis* ("–" pole in this example) to *trans* ("+" pole), giving rise to a blocked current level, as shown on the left. The duration of the blocked level and its average level are indicated.

has invited many further nanopore studies aimed at manipulating and probing biopolymer properties.

The most important advantage of nanopore sensing over other single-molecule techniques is that molecules can be detected without labeling or surface immobilization. Covalently labeling molecules with probes (e.g., fluorescent molecules, nanoparticles, quantum dots) or with surface-anchoring groups (e.g., biotin, His-tag, amines, thiols, etc.) is often used to detect and localize single biomolecules. However, labeling methods suffer from analytical and practical drawbacks, in that the function of labeled molecules may be altered during the labeling/immobilization process, the cost of labeling/purification can be high, and the throughput of studying immobilized molecules is usually low. With the use of nanopores, only minutes are required to nondestructively analyze thousands of single molecules, making this technique a unique high-throughput method.

Ion current detection in nanopores relies on a fairly high concentration of ions in solution, and typical salt concentrations in most of the early nanopore experiments (typically 1 M) were higher than physiological ionic strengths. However, recent studies demonstrated the ability to use physiological ionic strength, as discussed in Section 10.3. The relative nanopore/analyte dimensions also affect the contrast of the measurement. For example, purine and pyrimidine polynucleotides threaded through a lipid-embedded α-hemolysin (α-HL) channel produce different levels of blocked ion currents owing to differences in the size of the bases (Akeson et al., 1999; Meller et al., 2000). Other membrane channels are known to have appropriate dimensions for sensing single biopolymers, although channel stability and self-gating may prevent their widespread use. Single 2-nm alamethicin channels have been used for investigating the partitioning of PEG in and out of the channel (Bezrukov et al.,

1994). Translocation of double-stranded DNA (dsDNA) through the reconstituted *Bacillus subtilis* ion channel (Szabò et al., 1997) was subsequently demonstrated. The fragility of the lipid bilayer, coupled to inflexibility in fine tuning protein channel dimensions to accommodate larger biopolymers (proteins, dsDNA), has spurred a challenging yet fruitful exploration into synthetic nanopore fabrication. Synthetic nanopores of molecular dimensions are at the forefront of materials science, and recent developments have allowed superb control over nanopore dimensions and properties, rendering them promising tools for future nanopore devices.

When an individual biopolymer enters a nanopore it occludes a fraction of the ion current, resulting in an abrupt drop in the current to a level defined as i_b (Figure 10.1B). The fractional current is defined as $I_B = i_b/i_o$, where i_o is the unperturbed ("open") pore current. The fractional current facilitates a comparison between experiments using different molecules having different cross sections, as well as between different nanopore of slightly different i_o values, since to a first approximation, it depends *only* on geometrical parameters of the system, and not on environmental parameters such as salt concentration, voltage, or temperature. Specifically, I_B scales with the geometric ratio $\sim 1 - (a/d)^2$, where a represents the average hydrodynamic cross section of the biopolymer and d is the nanopore diameter. This is a powerful relationship because in principle it allows one to estimate the absolute hydrodynamic cross section of biomolecules when d is known. To achieve accurate estimation of the molecule's size, the nanopore diameter should ideally be only slightly larger than the molecular cross section, yielding high-contrast signals (i.e., small I_B values). The "dwell time" (t_D) is defined as the time during which the ion current remains below a threshold level, reflecting the residence time of a molecule in the nanopore. It is important to note here that for events involving fast transients, proper selection of the measurement bandwidth is critical to avoid bias in t_D statistics.

10.3. Biophysical Studies Using Protein Pores

10.3.1. α-Hemolysin

α-Hemolysin is a 232.4-kDa transmembrane channel (Gouax et al., 1994) secreted by *Staphylococcus aureus*. Exposure to α-HL is known to induce apoptosis (cell death) in different cell types, including rabbit erythrocytes, human erythrocytes, monocytes, and lymphocytes (Bhakdi and Tranum-Jensen, 1991; Jonas et al., 1994), via two separate mechanisms: (a) K^+ efflux and Na^+ influx through the plasma membrane, and (b) cytochrome c release from the outer mitochondrial membrane, triggered by α-HL insertion (Bantel et al., 2001). The crystal structure of the α-HL, solved in 1996 by Song and coworkers (Song et al., 1996), reveals a mushroom-shaped heptamer containing a water-filled channel 10 nm in length and 1.4–1.8 nm wide in the stem region. The cross-sectional view of the pore complex reveals two parts: A ~10-nm-wide extramembranal cap and a transmembrane β-barrel stem (Figure 10.2). The internal cavity of the α-HL cap (the "vestibule") upholds a 2.6-nm entrance and has a maximum diameter of ~3.6 nm. The β-barrel stem interior is believed to be hydrophilic, with a 1.5-nm limiting aperture at its entrance. The internal volume of the channel is ~18 nm^3, and, on average, roughly 600 water molecules can be accommodated in the stem. In 1 M KCl solution, ~11 K^+ and ~11 Cl^- molecules occupy the channel (Aksimentiev and Schulten, 2005).

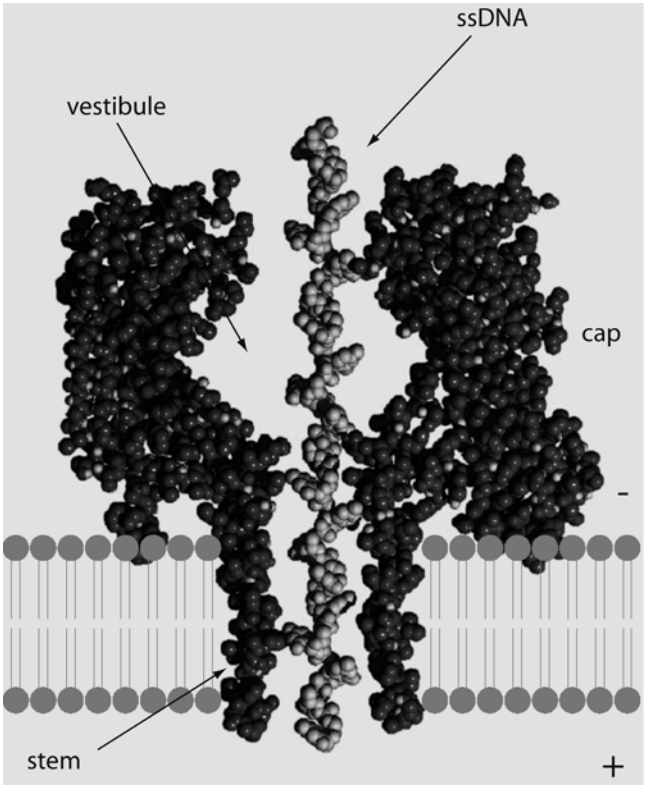


Figure 10.2. A molecular model of lipid-embedded α-hemolysin shown in cross section, based on the protein coordinates of the crystal structure determined by Song et al. (1996). A single-stranded DNA (ssDNA) molecule drawn through the pore is also shown. The stem of the complex spans the lipid bilayer, and the cap part resides on the *cis* side. The inner dimensions of the stem part are ~1.4–1.8 nm, allowing single-stranded but not double-stranded DNA to enter. [Modified from Deamer and Akeson (2000) with permission.]

α-HL can be expressed in *E. coli* and purified by standard sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) methods and is commercially available. It spontaneously self-assembles *in vitro* into lipid bilayers, forming a highly stable nongating channel at a pH range 7–9. The formed channel has a highly reproducible structure, as indicated by <2% pore-to-pore variation in the open-pore conductance (under temperature-stabilized conditions). The current–voltage curve of a single α-HL pore is slightly asymmetric, exhibiting a larger conductance when the stem part is positively biased (Figure 10.3A). These observations are explained by taking into account the asymmetrical inner surface charge distribution of the α-HL stem, resulting in differential rates of K^+ entry for opposite bias values (Aksimentiev and Schulten, 2005; Noskov et al., 2004). The temperature dependence on the open pore current of the channel can be explained by viscosity effects on the bulk ion mobility (Figure 10.3B). The observed robustness with temperature suggests a structural stability, ruling out the possibility of multiple large-scale channel conformations, although high-frequency fluctuations are not excluded. The narrowest portion of the α-HL channel (~1.4–1.8 nm) is ideal for translocation studies of single-stranded nucleic acids and short polypeptides.

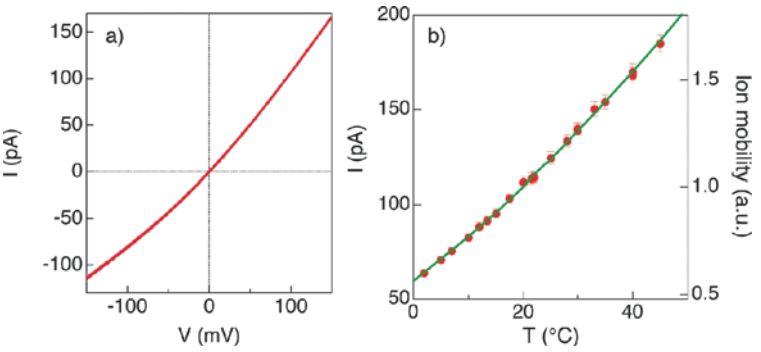


Figure 10.3. Current versus voltage (1 M KCl, 25°C) **(A)** and current versus temperature (1 M KCl, 120 mV) **(B)** curves measured for a single α-hemolysin pore embedded in lipid bilayer. The green line represents the bulk ion mobility of 1 M KCl as a function of temperature.

10.3.2. DNA Translocation Dynamics: The Role of Biopolymer–Pore Interactions and DNA Structure inside a Nanometer Confinement

The voltage-driven transport of nucleic acids and proteins through nanopores, in particular through α-HL, has constituted the basis for numerous biophysical studies in the last few years. We begin this section by reviewing some of the most important discoveries and features of DNA dynamics in nanopores. These studies have led to the development of nanopore force spectroscopy applied for DNA unzipping studies and for various DNA–enzyme interaction studies involving exonucleases and, more recently, polymerases.

Polynucleotide Translocation Dynamics

In 1996, Kasianowicz and coworkers showed that single-stranded RNA and DNA molecules can be driven through a single α-HL pore embedded in a phospholipid bilayer membrane (Kasianowicz et al., 1996). They found that the dwell-time distribution of poly(U) molecules with mean length of 210 nucleotides (nt) displays three apparent timescales (∼90, 290, and 1288 μs). Experiments with different poly(U) lengths showed that the second and third timescales grew linearly with polymer length, whereas the shortest timescale was length independent. The blockade events corresponding to the first timescale were therefore attributed to molecules colliding with and failing to traverse the pore, given that such collisions are expected to yield extremely fast, bandwidth-limited blockades. On the other hand, blockade events contained in the second and third timescales were attributed to molecules that traverse the pore (entering with either their 3′ or 5′ end to the pore), as confirmed by quantitative polymerase chain reaction (PCR) analysis of the molecules in the *trans* chamber after a fixed number of translocation events.

Akeson and coworkers studied the translocation properties of RNA and DNA homopolymers (Akeson et al., 1999). Remarkably, polycytidylic acid (single-stranded [ss] RNA) and polydeoxycytidylic acid (ssDNA) of comparable lengths were readily distinguished by three independent parameters: (1) The mean dwell time t_D of poly(dC) was found to be much shorter than that of poly(C) [0.9 μs/base for poly(dC) as compared to 5.8 μs/base for poly(C)]; (2) the blocked current level of poly(C) was found to be lower than that of poly(dC) (∼7 vs. ∼14 pA, respectively, measured at room temperature [RT] and 120 mV); and (3) the

blockade events of the poly(dC) polymers yield a single, tight cluster of events on the events diagram (blocked current level vs. dwell time), whereas poly(C) events are organized in two clusters spread over a large range of t_D (roughly 4–12 μs/base). These remarkable differences were rationalized by the idea that secondary structure existing in poly(C) molecules and not in poly(dC) might be responsible for the different behavior: poly(C) molecules have strong tendency to form single-stranded helices ~1.3 nm in diameter (Arnott et al., 1976) with narrow enough structure to transverse the pore, causing larger blockades than the less structured poly(dC). It was also found that the unstructured polyuridylic acid translates at rates similar to poly(dC), at roughly 1.4 μs/base.

Meller and coworkers investigated the translocation dynamics of a variety of ssDNA molecules at different temperatures (Meller et al., 2000). One of the striking results in this study was the large difference between characteristic translocation times of purines and pyrimidines [especially poly(dA) vs. poly(dC) or poly(dT)]. Figure 10.4 displays event diagrams of $poly(dC)_{100}$ (blue) and $poly(dA)_{100}$ (red) ssDNA and a mixed $poly(dCdA)_{50}$ molecule (green). Qualitatively, it can be seen that the dwell times for poly(dA) are significantly longer than for poly(dC) and that the blocked current for poly(dA) is lower. These features are shown in the two insets, where the dwell-time and blocked current distributions are displayed. The figure reveals a threefold ratio between the most probable translocation times t_P for poly(dC) and poly(dA), permitting a real-time identification of single homopolymers in a binary mixture of dA's and dC's. A temperature dependence study also showed that the threefold ratio persists even at high temperatures (~40°C), where self-stacking interactions in poly(dA) are known to be unstable (Saenger, 1988). This discovery revealed that base

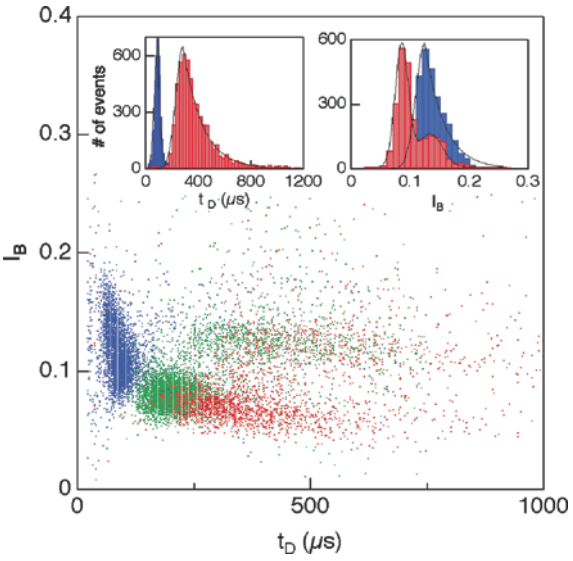


Figure 10.4. Event diagram measured for $poly(dA)_{100}$, $poly(dC)_{100}$, and $poly(dAdC)_{50}$ single-stranded DNA (*red*, *blue*, and *green*, respectively) measured using a single α-hemolysin nanopore at 120 mV and room temperature. Each dot represents a single translocation event with dwell time t_D and fractional blocked current I_B. The dwell-time and current distributions of the poly(dA) and poly(dC) are shown in the insets using the same color scheme. Adenines display much longer translocation times and smaller blocking currents than the cytosine homopolymer. The mixed $poly(dAdC)_{50}$ falls in between. [Modified from Meller et al. (2000).]

stacking (or secondary structure) alone could not account for the large difference in t_P values for purines versus pyrimidines, suggesting that DNA–pore interactions constitute a major barrier for polynucleotide translocation.

A later study confirmed this assertion by comparing alternating sequence polymers, poly(dAdC)$_n$, with "diblock" biopolymers, poly(dC)$_n$-poly(dA)$_n$ with the exact same length (Meller and Branton, 2002). Because purine-based stacking interactions are completely disrupted in the purine–pyrimidine alternating sequence, it was expected that the alternating sequence would translocate roughly twice as fast as the diblock biopolymer. However, roughly equal translocation times were found, indicating that base stacking (alone) is not the predominant factor determining the translocation dynamics of DNA polynucleotides.

Temperature dependence studies of polynucleotide translocation revealed a much stronger slowing down with decreasing temperature than that expected by viscous drag (e.g., a factor of ten in going from RT to 0°C) (Meller et al., 2001). These experiments again underlined the role of DNA–pore interactions, furthermore permitting the detection of extremely short polynucleotide, down to dinucleotides of deoxyadenines. It was found that molecules shorter than the α-HL channel (∼5 nm or 12 nt) translocate, on average, faster than longer polynucleotides (in terms of time per base). Similar to the linear length dependence of poly(U), it was found that t_P increases linearly with the number of bases N when $N > 12$ (Meller et al., 2001). The nearly threefold ratio between poly(dA) and poly(dC) was maintained for all DNA lengths greater than 12-mers. It indicates stronger interactions of the purine bases with α-HL as compared with pyrimidines (Meller et al., 2000). The strength of these interactions (characterized by t_p) grows roughly linearly with the number of purine bases embedded in a pyrimidine-rich DNA, saturating when the purine-to-pyrimidine ratio approaches ∼50% (Meller and Branton, 2002). To summarize, in the case of RNA homopolymers, the translocation dynamics through the α-HL pore is greatly affected by local structure of the polynucleotides (Akeson et al., 1999). For DNA, the dynamics is mostly dominated by base-specific nucleotide–pore interactions, in particular for poly(dC) and poly(dA).

Orientation Dependence of Polynucleotide Entry and Dynamics

While it was hypothesized that the existence of two "translocation" peaks in the dwell-time histogram is attributable to molecules entering with either their 3′ or their 5′ end to the channel, the physical basis for such a difference remained unknown (Kasianowicz et al., 1996). Nearly 9 years passed before this assertion was unambiguously proven and the two timescales were identified with the two respective polynucleotide ends. By plotting I_B histograms of thousands of translocation events, Mathé and coworkers observed that the ssDNA polynucleotides produced two overlapping current peaks (Figure 10.5A), but only one peak was observed when one of the polymers ends was blocked with a stable hairpin, which prevented one side from entering the pore (Mathé et al., 2005). Strikingly, the levels of the single peaks obtained when one of the polynucleotide ends was blocked matched perfectly the levels of the two peaks obtained with the unblocked polynucleotides, allowing assignment of the lower blocking peak with DNA molecules entering with their 3′ end and the higher, blocked level with molecules entering with their 5′ end (Figure 10.5b).

Furthermore, it was found that homopolymer dynamics in the pore is also orientation dependent. Specifically, those polynucleotides that were threaded with their 3′ end to the pore required a much longer time to escape from the pore than those with the 5′-threaded ssDNA. To decouple the effect of the voltage bias, Mathé and coworkers used dynamic voltage control

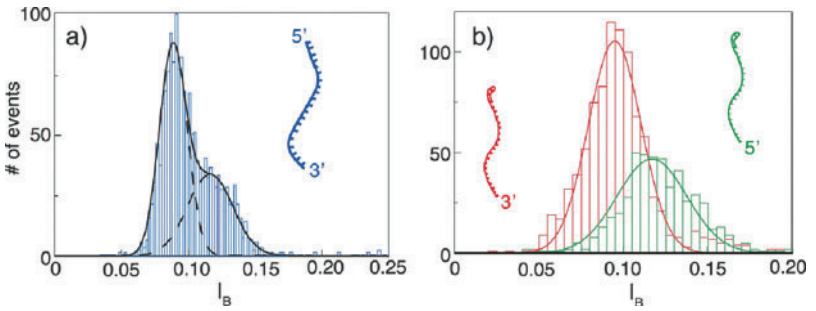


Figure 10.5. A. The fractional blocked current distribution of poly(dA) displays two peaks (centered at 0.09 and 0.12), each of which can be well fitted by a Gaussian function. **B.** Using single-stranded DNA molecules capped with stable DNA hairpins, the DNA is forced to enter the α-hemolysin pore with either its 3′ end (*red*) or the 5′ end (*green*), yielding two distinct single-peak distributions at the 0.09 and 0.12, respectively. [Modified from Mathé et al. (2005).]

to measure the force-free diffusion of 3′- and 5′-threaded poly(dA) molecules (with blocking terminal hairpins). They found a twofold increase in the escape dynamics of the 3′-threaded poly(dA) compared with the same-length, 5′-threaded poly(dA). Molecular dynamics simulations of the K^+ and Cl^- ions flowing through an α-HL pore in the presence of either orientation of ssDNA have independently reproduced these results and suggested a mechanism for this asymmetry: When the ssDNA is confined in the nanopore, the bases develop a uniform tilt toward their 5′ end. This tilt, in addition to the asymmetric nanopore structure, functions as a series of rectifying valves by hindering the flow of ions (mainly K^+ ions) when the nucleotides are oriented in one direction versus the other (Figure 10.6). A simple model that involves one-dimensional (1D) diffusion and drift was introduced to explain the escape time distributions of either 3′- or 5′-threaded DNA, indicating that the process is strongly

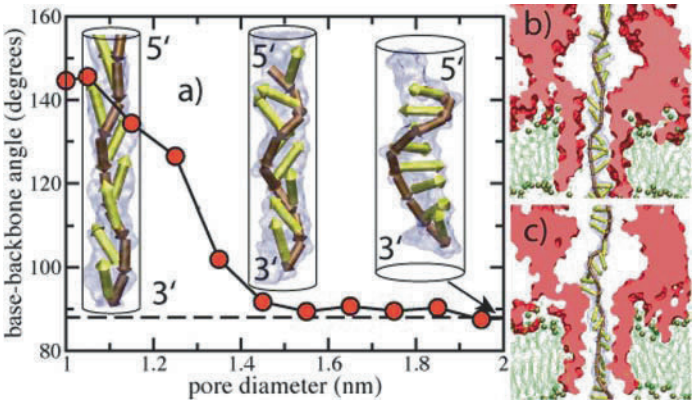


Figure 10.6. Results from molecular dynamics simulations of single-stranded DNA threaded into the α-hemolysin pore. **A.** The average angle between the base and the backbone of a nucleotide in a poly$(dA)_{11}$ strand confined inside a cylindrical pore as a function of the pore diameter. Snapshots at 1.0, 1.25, and 3.0 nm from left to right illustrate the nucleotide conformation. As the pore diameter is reduced. the bases develop a collective tilt, always toward the DNA's 5′ end. The right-hand panels are magnified images of a longer poly(dA) strand inserted either with its 3′ **(B)** or 5′ **(C)** end to the pore. [Modified from Mathé et al. (2005).]

voltage dependent and the large differences between the two strands at low (or vanishing) voltages disappear for assisting voltages >40 mV (Wanunu et al., 2008a).

10.3.3. Probing Secondary Structure: DNA End-Fraying and DNA Unzipping Kinetics

Double-stranded DNA cannot fit inside the ~1.5-nm α-HL stem. Consequently, blunt-ended molecules produce different current blockade signals when compared with ssDNA: The ion current is blocked to ~50% of the open-pore current (as opposed to the ~10% with ssDNA). The partial current blockade levels have been associated with temporary lodging of the DNA duplex in the pore vestibule (2.6 nm) that can admit double-stranded nucleic acids (2.2 nm), as shown in Figure 10.7 (Vercoutere et al., 2001). Vercoutere and coworkers found that for blunt-ended DNA hairpin lodged in the α-HL vestibule a mid-level blocked current (Figure 10.7B) is often followed by a brief full blockade level (C). This pattern was interpreted as dissociation of the hairpin duplex and eventual threading and rapid translocation of the ssDNA through the pore, causing a full but brief current blockade. The dependence of the mid-level blockade on the hairpin length revealed two features: (1) I_B systematically dropped from ~0.7 to 0.4 when the hairpin duplex region increased from 3 to 8 bp, respectively, and (2) the duration times of the mid-level blockades were broadly distributed, with their average value increasing exponentially with hairpin length. These are in line with the idea that the mid-level state is terminated by the spontaneous denaturation and eventual translocation of the ssDNA through the pore. Further experiments have revealed the existence of discrete mid-level fluctuations for some of the hairpins under study. These were shown to be a result of the kinetics of the fraying ends of the DNA hairpin (Winters-Hilt et al., 2003).

Nanopores can also be used to directly apply unzipping forces on DNA and RNA molecules, which allow probing the structure of highly stable nucleic acid conformations (e.g., long hairpins). While distinguishing unzipping energy barriers can provide an important tool for biophysics, it is also important for applications in genomics, by enabling the detection of subtle sequence variations in oligonucleotide analytes (e.g., single-nucleotide polymorphisms). Experiments with DNA hairpins or bimolecular DNA duplexes with single-stranded overhangs reveal longer dwell times as compared with ssDNA molecules of similar lengths (Mathé et al., 2004; Nakane et al., 2004; Sauer-Budge et al., 2003). Characteristic dwell times grow *exponentially* with the number of base pairs in the duplex region, also growing with the hairpin G-C content. A single base-pair mismatch in an otherwise perfect hairpin structure (e.g., of 10 bp) can be readily detected by a shift to shorter dwell times (Mathé et al., 2004). In addition, the characteristic dwell time depends exponentially (to a very good approximation) on the applied voltage $\sim e^{-V/V_\beta}$, where V_β is a characteristic voltage. These measurements suggest that the dwell times correspond to the hairpin unzipping time, τ_U. Moreover, the exponential dependence rules out possibilities that the DNA hairpins translocate through the pore without unzipping or escape from the pore at the entry side, in that such processes would have yielded weaker, nonexponential, dependencies.

Nanopore force spectroscopy (NFS) allows the application of a time-dependent voltage ramp $V(t) = vt$, where the slope v is called the *velocity* (Mathé et al., 2004). This method is an analog of force spectroscopy in other SM techniques (Evans, 2001; Evans et al., 1995; Hugel and Seitz, 2001). One of the advantages of NFS is that it permits measurements in inaccessible regimes using the fixed-voltage method, in particular at low voltages (forces) (Figure 10.8). Mathé and coworkers found that hairpin unzipping kinetics is characterized by two regimes: the high-force (high-velocity) regime, where hairpin unzipping is essentially

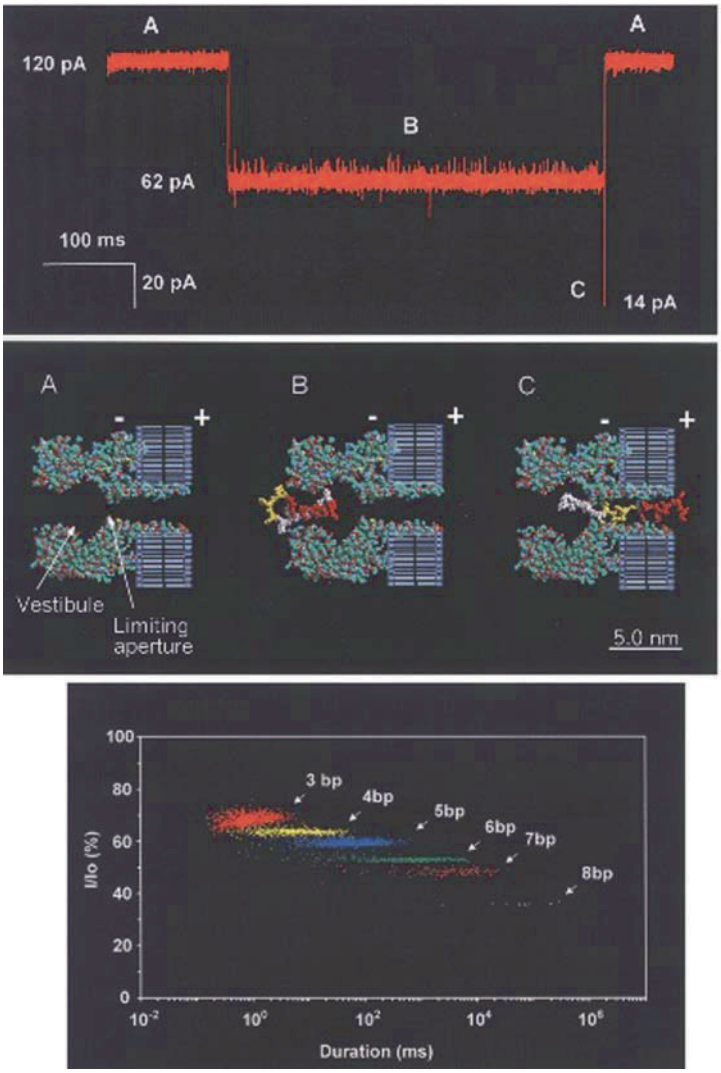


Figure 10.7. Current blockade of the α-hemolysin nanopore by a blunt-ended DNA hairpin. **Top.** A current trace caused by the capture and translocation of a 6-bp DNA hairpin. M**iddle.** The α-hemolysin and the hairpin during the three states of the process. **Bottom.** Semilog event diagram for DNA hairpins with 3- to 8-bp stems, as indicated. [Adapted from Vercoutere et al. (2001).]

an irreversible process, and the low-force (low-velocity) regime, where rezipping may occur several times before the hairpin is pushed outside the pore (Mathé et al., 2005, 2006). This interpretation is based on a simplified model of a single-energy-barrier crossing. But a more involved theoretical framework developed to explain the NFS data showed that re-zipping is not necessary (Dudko et al., 2007). Moreover, the new theory unifies the fixed-force and time-varying force in a single universal unzipping curve, spanning a broader range of time-scales (Dudko et al. 2008). NFS has been found to be extremely sensitive to the hairpin energy: A single base mismatch can be readily distinguished, enabling single-molecule-hybridization

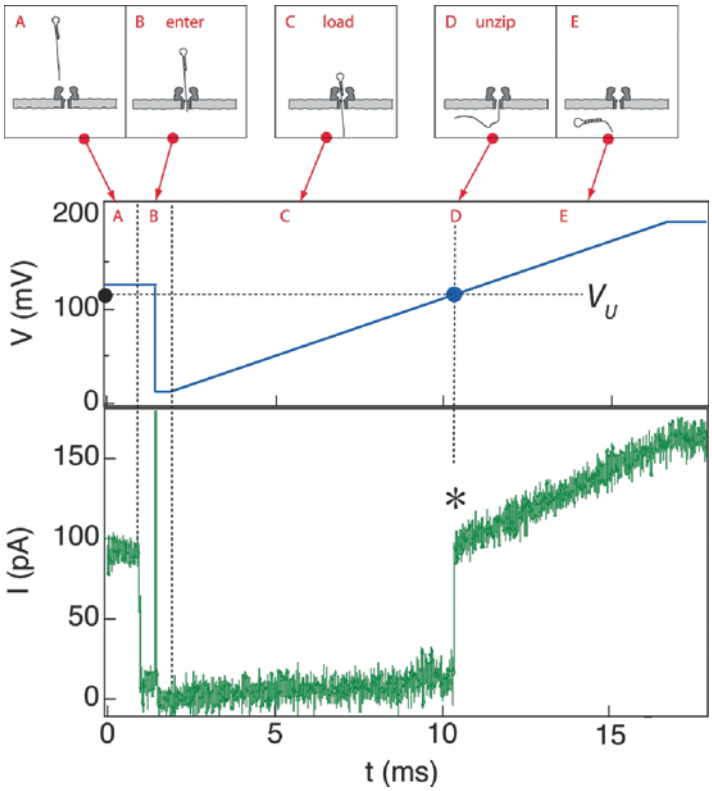


Figure 10.8. Nanopore force spectroscopy is used to measure the unzipping kinetics of DNA hairpins. A time-dependent voltage (or force) (*blue curve*) is applied to the molecule, and its response (ion current) (*green trace*) is measured by detecting the voltage (or force) at which unzipping occurs (V_U). Initially the pore is open (**A**), as indicated by the open level. DNA threading results in an abrupt blocking in the current (**B**). Threading of the single-stranded overhang only lasts a few tens of microseconds, after which a steady voltage ramp is applied (**C**). At ~11 ms the current abruptly jumps from the blocked state to the open state as indicated by an asterisk. This transition corresponds to the unzipping of the hairpin structure and the quick translocation of the unzipped part. This experiment is repeated hundreds of times to obtain the critical voltage at which hairpins are unzipped for a given voltage ramp slope. [Modified from Mathé et al. (2004).]

detection of unlabeled nucleotides. Moreover, since the DNA is hybridized in solution before being probed in the nanopore, biasing surface effects and crowding effects on DNA hybridization, which may lead to false positives in DNA microarray technology (Reiner et al., 2003), for example, are circumvented. In addition, NFS can be used to probe unmodified nucleic acids, circumventing the need to tag biopolymers with biotin or other biomolecular complexes.

10.3.4. Probing DNA–Protein Interactions: DNA Exonucleases and Polymerases

The ability of nanopores to probe native (unlabeled, unmodified) biomolecular complexes over an extended range of timescales (microseconds to seconds) lends itself to numerous studies of DNA–protein interactions and motor enzymes. Labeling a protein for

single-molecule studies is tedious and requires precise knowledge of its structure/function (e.g., labeling near a catalytic site can severely affect enzyme activity), whereas using a native protein is tremendously convenient and trustworthy. Recently, there have been several reports on the activity and interactions of a number of DNA–protein systems, including ssDNA interactions with exonuclease I, sequence-specific detection of DNA–Klenow fragment polymerase complexes, and incorporation of single nucleotides by a DNA Taq polymerase.

Exonuclease I–DNA Interactions

Hornblower and coworkers studied the interactions of exonuclease I (Exo I) with a ssDNA target using NFS (Hornblower et al., 2007). DNA molecules bound to a protein cannot translocate through the pore and would normally cause long ion-current blockades at the level of roughly 10%–20% (due to ssDNA occlusion in the pore). The digestion reaction was monitored by measuring the event rate (the rate of full-size ssDNA entries to the nanopore) as a function of time before and after the addition of 5 mM Mg^{2+}, indicated by an arrow in Figure 10.9A. As a control, a 3′-phosphorylated oligo was used, allowing binding but no digestion (circles). The presence of Exo I shifts the dwell times of a fraction of the events (proportional to the fraction of DNA–Exo I complexes) from ∼70 μs to 1 ms (Figure 10.9B). In this experiment, DNA digestion was prevented by not adding Mg^{2+} cofactor to the buffer. Close inspection of the dwell-time distributions of the 3′-phosphorylated DNA revealed the existence of a subpopulation of translocation events characterized by a longer dwell time (∼10 ms) as compared with the no-Mg^{+2} conditions, indicative of tighter DNA–Exo I binding mediated by the Mg^{2+} ions.

Nanopore force spectroscopy was used in a similar way in the studies of DNA hairpin unzipping described earlier (Mathé et al., 2004): The voltage-induced dissociation of the complex is signaled by an abrupt increase in the ion current through the nanopore, when the DNA strand exits the pore to the *trans* side. Each rupturing event yields the rupturing voltage

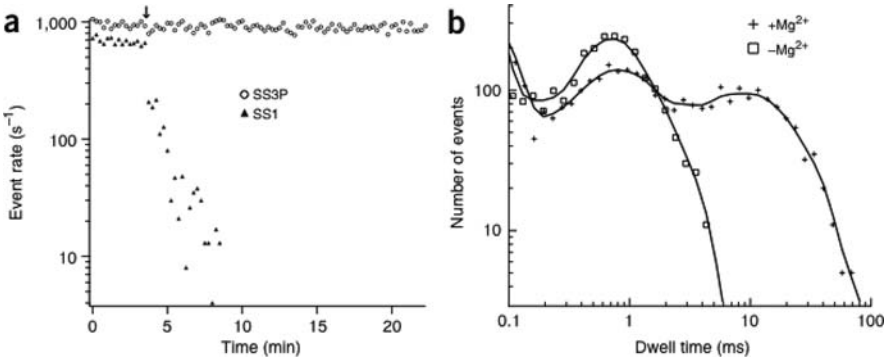


Figure 10.9. Monitoring exonuclease I (Exo I) digestion of DNA using a nanopore. **A.** The DNA capture rate in the nanopore as a function of time. After addition of 5 mM Mg^{2+} (*arrow*), the event rate sharply declines for the SS1 (3′-OH end) but not for the SS3P (3′-phosphorylated end). **B.** Translocation time distributions for complexes of SS3P–Exo I in the presence and absence of 5 mM Mg^{2+}. [Adapted from Hornblower et al. (2007).]

V_R, and a distribution of the rupturing voltages accumulated from ~1,000 SM events is used to determine the most probable rupturing voltage V_C. The dependence of V_C on v (plotted on a semilog scale) provided a measure of the association and dissociation rates of the complex, as interpreted using a simple semiempirical model.

Nanopore Probing of Deoxyribonucleotide Triphosphate Incorporation by a Klenow Fragment DNA Polymerase

Benner and coworkers used nanopores as a biophysical sensor to measure the binding dynamics of a Klenow fragment (KF) of *Escherichia coli* DNA polymerase I with its dsDNA substrate (Benner et al., 2007). The KF catalyzes DNA replication by sequential addition of nucleotides to the primer strand. Based on the use of nanopores as the detection device, the presence of dsDNA with and without bound KF protein, as well as the bound and unbound states of deoxyribonucleotide triphosphate (dNTP) to the DNA–KF binary complex, were distinguished.

Following capture of the DNA complex (14-bp hairpin with a 36-nucleotide 5′ overhang terminated by 2′,3′-dideoxycytidine) inside α-HL and the addition of KF (2 μM) to the solution, a distinct cluster of events was recorded with lower blockade and longer dwell times, the fraction of which was shown to be dependent on the KF concentration. The average pore current increased from 20 pA for dsDNA to 23 pA for DNA–KF binary complexes, whereas dwell times increased from 1 ms median duration for dsDNA to ~3 ms for the DNA–KF complex (Figure 10.10 A, B). The individual current traces revealed two distinct states: a shallow blocked state attributed to current blocked by the enzyme–DNA complex where the ssDNA template overhang blocks the nanopore, and a deeper blocking, short-lived state that corresponds to residency of dsDNA in the vestibule of the α-HL pore following enzyme dissociation. Following enzyme dissociation, a return to the open-pore current state marks translocation of the DNA to the *trans* side. The voltage dependence of enzyme–DNA complex dissociation was established at a reduced unbinding force by reducing the transmembrane voltage from 180 mV to 165 mV, in which dwell times of the final DNA dissociation step increased. This observation is consistent with the model in which following enzyme dissociation, the dsDNA in the pore dissociates and translocates through the pore rather than escapes back to the *cis* chamber.

Although these changes in blocked current and dwell times were subtle, the current and dwell-time signatures dramatically improved when the correct complementary dNTP to the KF catalytic site was added to the solution in the presence of Mg^{2+} ions (Figure 10.10, column II), resulting in the formation of a more stable ternary complex with the 2′,3′-dideoxy-terminated template. Because dideoxy termination prevents enzymatic extension by the polymerase, the ternary complex exhibits an order-of-magnitude larger dwell time in the pore. The authors have confirmed the consistency of this effect by using different templates that accommodate the four dNTPs.

Probing DNA Polymerase Activity Using a Nanopore

Another example of biophysical study using nanopores is the study of the incorporation of single nucleotides by DNA polymerase (Cockroft et al., 2008). In this work, Cockroft and coworkers developed a mechanically locked supramolecular assembly capable of monitoring the activity of DNA polymerase as it operates on a single strand of DNA threaded through

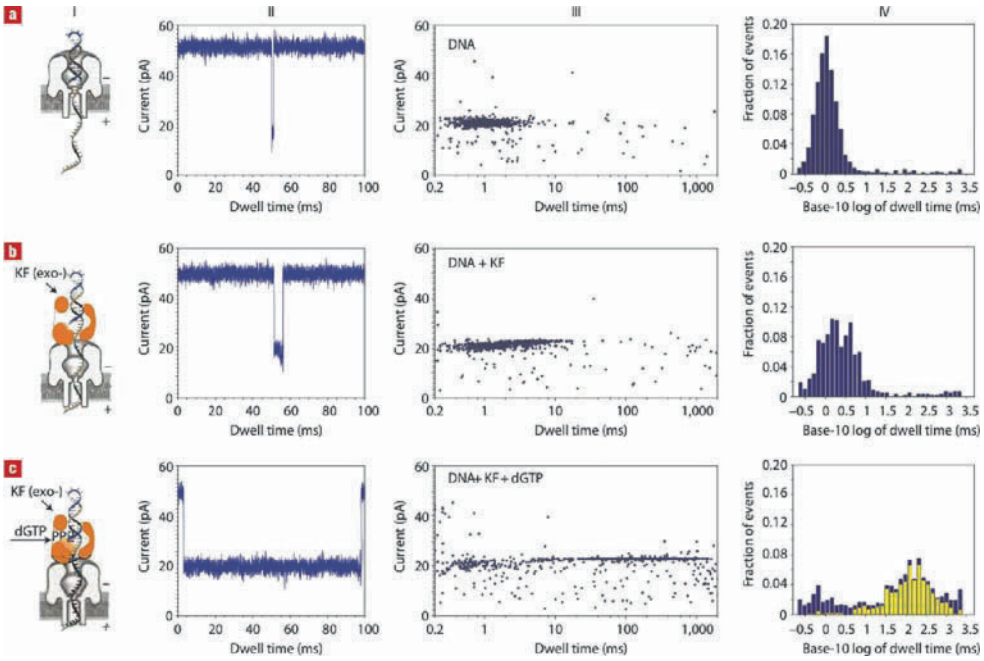


Figure 10.10. Distinguishing DNA, DNA–Klenow fragment (KF) complexes, and DNA-KF- deoxyribonucleotide triphosphate (dNTP) complexes in the nanopore device. **A.** Translocation through the nanopore of DNA alone (14-bp hairpin with a 36-nucleotide 5′ overhang and 2′ –3′ dideoxycytidine terminus; template base at $n = 0$ is C). **B, C.** Translocation of the 14-bp hairpin from complexes with KF **(B)** or from complexes with KF and deoxyguanosine triphosphate (dGTP) **(C)**. Diagram are shown of the nanopore with the associated complex (column I), a current trace (column II), and a dwell-time event plot (column III). In column IV, probability histograms of the base-10 logarithm of dwell-time data are shown in blue. Close examination of the event plot in column III of panel C reveals that most long-dwell-time events are within 22 to 24 pA. A yellow subset histogram for the events within 22–24 pA is overlaid on the probability histogram in panel C, revealing that the chosen range is dominated by long-dwell-time events. [From Benner et al. (2007).]

a protein pore. The experimental geometry consists of carefully chosen lengths of ssDNA–PEG copolymer threaded through a lipid-embedded α-HL nanopore. The anionic character of the ssDNA–PEG phosphate copolymer drives them through the nanopore under an applied transmembrane potential. The copolymers are then locked on both ends by streptavidin and dsDNA caps that are larger than the pore diameter. After locking on both ends, the relative position of the two polymers can be switched simply by reversing the polarity of the applied transmembrane potential. In this configuration, the ion current through the nanopore is a reporter of the type of polymer inside the pore (ssDNA or PEG) (Sanchez-Quesada et al., 2004). The authors show that the PEG chains block ∼50% of the current (with or without the streptavidin cap), whereas the ssDNA chain blocked ∼70% of the current (with or without the dsDNA cap).

The mechanically locked dsDNA-ssDNA-PEG-biotin-streptavidin molecule inside the α-HL nanopore was further used to detect base-by-base nucleotide extension of a DNA primer by DNA polymerase. The rotaxane molecule was designed such that each nucleotide addition to the primer by the DNA polymerase changes the relative proportion of ssDNA to PEG residing inside the pore, thus moving from PEG conductance state to ssDNA conductance

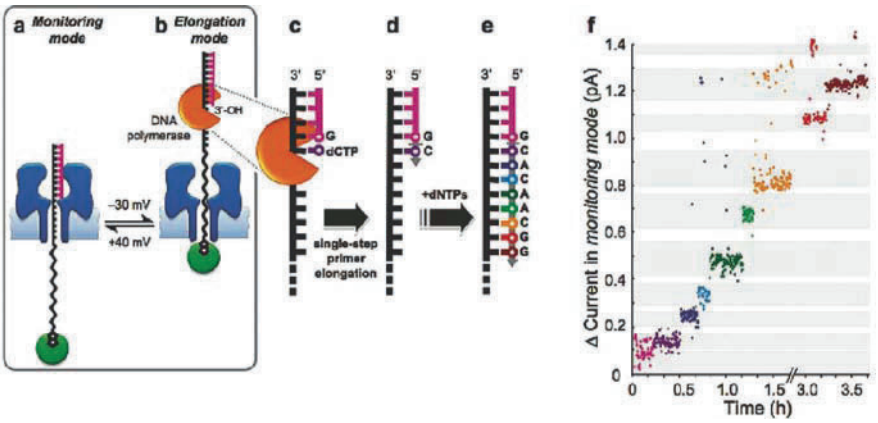


Figure 10.11. Monitoring DNA polymerase–catalyzed single-nucleotide primer extensions. **A.** The length of the DNA primer in a fully interlocked α-hemolysin DNA–polyethylene glycol complex is reported by ion current measurements in the monitoring mode (+40 mV). **B.** In the elongation mode (–30 mV), the 3′-OH of the primer is accessible to the DNA polymerase. **C–E.** DNA polymerase bound to the starting primer–template complex and templated extension of the primer sequence after correct bases are incorporation. **F.** A plot of cumulative change in the ion channel current recorded in the monitoring mode versus time (note the broken time axis). dCTP, deoxycytidine triphosphate; dNTP, deoxyribonucleotide triphosphate. [From Crockroft et al. (2008).]

state in steps of single nucleotide addition. The threading component of the nanopore device consists of a 31-nucleotide ssDNA segment bearing the primer-binding site and the template sequence, an 11-nucleotide poly(dA) sequence, an eight-unit PEG phosphate chain, and a terminal biotinyl group. Using a primer and a streptavidin molecule to lock the DNA–PEG chain on the *cis* and *trans* sides of α-HL, respectively, the authors switched between two modes: (1) In the monitoring mode, a transmembrane potential of +40 mV is applied, and the current provides a measure of the length of ssDNA region inside the pore. Elongation of the DNA primer in this mode is inhibited because the 3′-OH elongation site of the primer is concealed within the pore vestibule, preventing access by the DNA polymerase. (2) In the elongation mode, a transmembrane potential of –30 mV is applied, in which the ssDNA exits the pore with the 3′-OH of the primer strand exposed (Figure 10.11A, B). On adding TopoTaq DNA polymerase in the presence of deoxycytidine triphosphate (dCTP) and waiting for 30 min in the elongation mode, the authors switched the polarity of the voltage to the monitoring mode, after which they observed a step increase in the average current, attributed to single-base incorporation in the threaded template strand. By successive addition of other dNTPs, stepwise increases in the current were observed (Figure 10.11C–F). Nucleotide incorporation steps were then assigned based on the dNTP added from the cumulative ion current. The single-base sensitivity achieved in this work provides an important technological step toward DNA sequencing using protein nanopores, although significant advances are required for this method to be used in *de novo* DNA sequencing.

10.4. Biophysical Studies Using Solid-State Nanopores

As previously discussed, studies using lipid-embedded nanopores have provided a wealth of information on biopolymer structure and dynamics and have been applied for *in vitro* studies of numerous biomolecular systems. Protein pores are extremely versatile sensors

and are generally inexpensive, although they suffer from a number of shortcomings: (1) limited lipid bilayer stability and challenging single-channel insertion protocols complicate the integration of bilayer-based sensors into robust devices; (2) inevitable lateral diffusion of the protein channels in lipid bilayers interferes with the implementation of complementary detection modalities, such as optical detection (Hemmler et al., 2005); and (3) the fixed geometry of protein nanopores may not be suitable for different applications, the α-HL nanopore being a rare example of serendipity. In contrast, nanopores fabricated in thin synthetic membranes can offer superior chemical and mechanical robustness, as well as provide size tunability. Over the last decade, efforts have been made to produce synthetic nanopores using different materials and state-of-the-art techniques (Healy, 2007; Wanunu and Meller, 2008). The vast majority of studies involving biopolymer analysis have been carried out using solid-state nanopores fabricated in ultrathin Si-based membranes (Dekker, 2007), which are the focus of this section.

10.4.1. Nanopore Fabrication

The fabrication of nanopores in solid-state materials was introduced by Li and coworkers using ion-beam sculpting (Li et al., 2001). In this technique, an energetic beam of Ar^+ ions was used to sputter a silicon nitride (SiN) membrane, which contains a bowl-shaped cavity. The sputtering caused a gradual thinning of the film, eventually resulting in the formation of a nanopore, as detected by transmission of Ar^+ ions through the membrane. Nanopore sizes down to 1.5 nm have been fabricated using this method, as verified by transmission electron microscopy (TEM). Storm and coworkers employed an alternative procedure in which a focused electron beam (e-beam) of a commercial field-emission TEM was used (Storm et al., 2003). Using this technique, a large ($\sim$100 nm) SiO_2 pore was irradiated with an intense beam of 300-kV electrons, resulting in fluidization of the SiO_2 layer and shrinking of the pore. Direct use of the TEM for shaping the nanopores allows both fabrication and analysis of the nanopore shape in real time with TEM imaging precision (<1 nm). Direct drilling of intact SiO_2 and/or SiN membranes using the TEM e-beam is currently the most common method for fabricating solid-state nanopores. The planar ultrathin membrane used as substrate for solid-state nanopores allows nanopore arrays to be fabricated at specified locations using scanning TEM schemes (Kim et al., 2006, 2007), further enabling increased throughput by the use of other detection modalities (e.g., parallel optical detection).

10.4.2. Experimental Considerations

Although protein channels have well-defined geometry and chemistry, there are several challenges associated with manually fabricated synthetic nanopores. Pore-to-pore variability in both physical structure and chemical makeup has been impeding progress in the development of solid-state nanopore platforms. An example of the outcome of this variability is the different average velocities of double-stranded DNA translocation reported with 8- to 15-nm solid-state pores (Healy, 2007). Second, solid-state membranes exhibit larger capacitive noise than the lipid-embedded α-HL pores, limiting practical experimental bandwidths (i.e., the apparent time resolution). As a result, many studies with solid-state pores have been performed using lower bandwidths than protein-pore studies, limiting the range of biopolymer lengths that could be reliably detected.

Nanopore structure has a tremendous impact on its function, already observed with the asymmetric current–voltage (I-V) curves of α-HL (Figure 10.3) and its high capture rate for

ssDNA from the *cis* side as compared with the *trans* side (attributed to the vestibule funnel structure) (Henrickson et al., 2000). In making the transition to solid-state nanopore studies, precise characterization of the nanopore structure and surface properties is essential. The three-dimensional (3D) structure of solid-state nanopores fabricated using an e-beam has been recently revealed using TEM tomography (Kim et al., 2006). The pore structure resembles a truncated double-cone (hourglass) shape, with a half-angle of ~30° and an effective nanopore length of about one third of the SiN membrane thickness (Figure 10.12). To investigate the ion transport properties, a series of nanopores in the range 3–15 nm were fabricated and assembled into two-chamber devices. Ion-current measurements typically yield ohmic I-V curves, the slopes of which are used to define the pore conductivity, G, described by the following equation:

$$G = \frac{\pi d^2}{4}\sigma\left(\frac{\delta \tan\alpha + 1}{h + h_{eff}\delta \tan\alpha}\right) \tag{10.1}$$

where $\sigma = (\mu_K + \mu_{Cl})n_{KCl}e$ is the specific conductance at a number density n_{KCl}, equal to 15.04 and 3.01 $(\Omega\ \mathrm{m})^{-1}$ for 1.0 and 0.2 M KCl, respectively, with μ_K and μ_{Cl} the electrophoretic mobilities of K^+ and Cl^-, respectively, and e is the elementary charge unit, $\delta = (h - h_{eff})/d$, with h_{eff} is the width of the cylindrical region in the nanopore, and α is the cone half-angle. Figure 10.13 depicts the dependence of G on the pore diameter as determined by TEM imaging. The solid lines are fits to Eq. (10.1), taking into account the pore structure, as determined by TEM 3D tomography and parameterized by the geometrical factors α and h_{eff}.

Surface charge effects were recently studied by investigation of salt dependence on the pore conductivity and the current signature during DNA translocation (Chang et al., 2004;

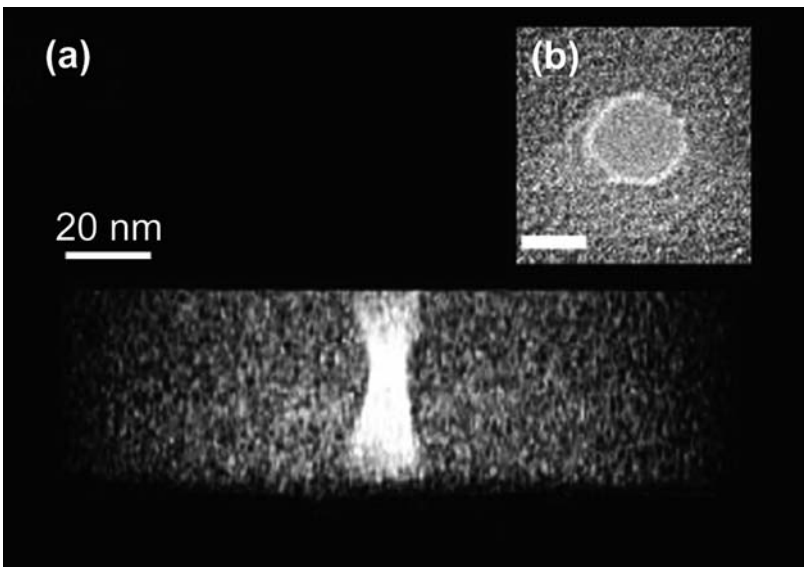


Figure 10.12. A. A reconstructed side view obtained by transmission electron microscopic tomography of a 7-nm pore in a 50-nm SiN membrane, revealing a truncated double-cone structure. **B.** A high-resolution top-view image of the same pore. [Modified from Kim et al. (2006).]

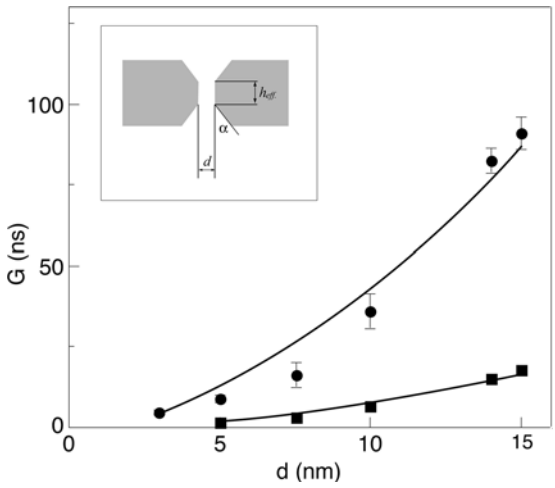


Figure 10.13. Ion conductance G of solid-state nanopores as a function of their diameter measured in 1.0 (*circles*) and 0.2 M (*squares*) KCl solution. The lines represent fits to Eq. (10.1) assuming a truncated double-cone structure, as shown in the inset. [From Kim et al. (2006).]

Ho et al., 2005; Smeets et al., 2006). Smeets and coworkers found that the surface charge of 10-nm-diameter pores fabricated in SiO_2 dominates the pore conductivity at low salt concentration (<0.1 M), leading to a nonlinear dependence of G on salt concentration. Of interest, at low salt concentrations, DNA translocation results in a transient current enhancement (rather than a current drop), associated with an increase in the number of charge carriers inside the pore volume on DNA entry.

Solid-state nanopores have been recently used to detect a variety of biopolymers, including DNA and proteins, antibody–virus interactions, and other immunocomplexes. In particular, several groups have focused on understanding the dynamics of dsDNA translocation through solid-state nanopores. We briefly discuss the results of these studies because they have shown utility in identifying DNA fragments by length and conformation and studying DNA–protein interactions.

10.4.3. DNA Translocation through Solid-State Nanopores

As with α-HL pores, the translocation dynamics of dsDNA is a function of several key parameters: (1) the force applied on the DNA molecule, which is a product of the effective charge of the DNA portion inside the nanopore and the applied voltage; (2) the properties of the DNA molecule (e.g., persistence length, cross section, conformation); (3) the properties of the solution (e.g., viscosity, ionic strength, temperature); and (4) interactions between the DNA and the nanopore.

Keyser et al. (2005) directly measured the force on a single DNA molecule pulled by an applied voltage by coupling the nanopore setup with an optical tweezers system. They brought a λ-DNA tethered bead to the vicinity of a nanopore (6–15 nm in diameter), and, following the application of voltage, they observed a reduction in conductance that corresponds to DNA threading. Based on the use of the force-calibrated optical tweezers setup, measurement of the displacement of the bead with applied voltage yielded a force of 0.24 ± 0.02 pN/mV, corresponding to an effective charge of 0.50 ± 0.05 electron per base pair. This

elegant experiment confirmed that the DNA effective charge is reduced when confined in a nanopore, in qualitative agreement with theories and other experimental results performed with α-HL and ssDNA (Jan Bonthuis et al., 2006).

Numerous studies involving ssDNA and dsDNA translocation through solid-state nanopores have been reported in recent years (Chang et al., 2004; Fologea et al., 2005; Heng et al., 2004; Kim et al., 2006; Li et al., 2003; Storm et al., 2005a). Quantitative polymerase chain reaction (PCR) has been used to confirm translocation, demonstrating that the number of current blockade signatures correlates with the amount of DNA in the *trans* chamber (Heng et al., 2004). Double-stranded DNA was found to translocate in both singly and multiply folded conformations through 7- to 15-nm pores, indicated by multiple blocked ion-current levels (Chen et al., 2004a; Gershow and Golovchenko, 2007; Li et al., 2003; Storm et al., 2005a). Storm and coworkers studied the effect of DNA length on the translocation dynamics through 10-nm pores (Storm et al., 2005a). When restricting the analysis to only linear DNA translocation (by analysis of single-level events only), they found a nonlinear dependence of the translocation dynamics on DNA length. This result was explained by considering viscous drag through the solution as the DNA coil shrinks near the nanopore during the translocation process. This argument leads to a scaling of the translocation time of $\tau \propto R_g^2$, where R_g is the radius of gyration of the DNA coil. Since $R_g \propto l^{\nu}$, where $\nu = 0.61$ is the Flory exponent and l is the DNA contour length, the calculated power law scaling of 1.22 is similar to the observed power law of 1.27.

Fologea and coworkers recently investigated the effect of DNA conformation on the translocation signal (Fologea et al., 2007). Using an 18-nm pore, they showed that linear, circular relaxed, and coiled circular DNA can be distinguished by the current blockade and translocation time. The circular DNA form exhibited current blocking that was double the magnitude of that for the linear DNA. The length dependence was also determined from the product of the dwell time and the current amplitude of each event. This product (called the event charge deficit [ecd]) was found to be independent of the DNA conformation inside the pore. These authors reported average translocation times for the DNA molecules on the order of 50–100 ns/base pair, almost two orders of magnitude faster than ssDNA translocation through the α-HL pore.

Slowing down the voltage-driven DNA dynamics in solid-state pores remains an important challenge. It is expected that rates on the order of 1–10 μs/base pair or slower are needed for the electronic implementation of sequence readout of DNA. Environmental factors such as increased solution viscosity or reduced voltage have been reported to increase the average translocation time by a factor that is commensurate with either viscosity increase or voltage reduction (Fologea et al., 2005). However, these parameters have negative effects on the capture rate of DNA to the pore: DNA collision with the pore is to a first approximation diffusion limited, strongly depending on the voltage gradient in the pore vicinity and the solution viscosity. Therefore, the use of viscosity and voltage to slow down DNA translocation results in reduced capture rate and can limit the practicality of the nanopore technique.

On the other hand, increasing the nanopore–DNA interactions can be used to better regulate the DNA sliding process with little or no negative effects on the rate of DNA capture into the pore. In addition, changing the nanopore chemical characteristics can affect the nanopore selectivity to analytes and prevent unwanted irreversible sticking to the nanopore surface. Chen and coworkers showed that nanopores modified with thin alumina coatings exhibit enhanced DNA capture rate and lower noise compared with uncoated pores, attributed to increased positive surface charge (Chen et al., 2004b). Iqbal and coworkers showed that

ssDNA transport through hairpin-functionalized nanopores is sequence dependent, suggesting that surface interactions can be tailored for selectivity (Iqbal et al., 2007). Wanunu and Meller (2007) developed a general method for applying monolayers of organic coating inside the nanopores, with various functional end groups, such as amine, aldehyde, PEG, or carboxylic acid. Chemical coating can open up various options, from tailoring the coating to specifically slowing down nucleic acids transport, to nanopore functionalization with enzymes, antibodies, and so on. If needed, the organic coating can be stripped off and the nanopore can be recycled many times, with little effect on the inorganic substrate.

More recently the effect of nanopore size on DNA translocation dynamics has been studied with high detail, focusing on the size range in which unfolded DNA is forbidden (3–5 nm) and DNA is expected to thread in a single-file configuration (Wanunu et al., 2008b). These studies have revealed that the DNA capture rate, as well as its sliding dynamics in the pore, is strongly affected by small variations in the nanopore diameter. Remarkably, it was found that a decrease of 1 nm in the pore size leads to nearly an order-of-magnitude increase in the translocation time (Figure 10.14). For the same nanopore size range, the blocked ion current I_B followed very closely the geometrical approximation $I_B = 1 - (a/d)^2$, with $a = 2.2$ nm being the DNA hydrodynamic cross section. These discoveries establish that in this size range, DNA–pore interactions *govern* the sliding process. Other factors that support this idea are the broad and well-characterized dwell-time distributions and the strong power-law dependence of the characteristic time scale on DNA length (for DNA longer than ~3 Kb). While the nanopore size was clearly shown to reduce average translocation speeds, a tradeoff between the translocation probability and average translocation speeds was revealed: When nanopore sizes decreased below 4 nm, a major fraction of events was collisions rather than translocations, while for nanopores larger than 4 nm, most events were translocations. The fraction of collisions to translocations was determined by the current-blockage levels and the corresponding dwell-time dependence.

Despite remarkable progress in the study of biopolymer transport through solid-state nanopores, it is not clear whether the sliding progress occurs at a steady velocity or a more complex dynamics is present. This question is associated with the nonlinear dependence of the translocation time on DNA length and its broad temporal distributions. In contrast, results from α-HL discussed in the previous section revealed a much slower average translocation rate

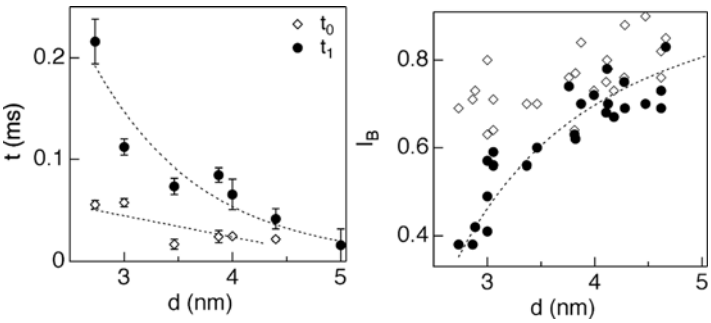


Figure 10.14. Left. The dependence of the translocation timescale t_1 (*solid circles*) and the collision timescale t_0 (*open diamonds*) on the nanopore diameter measured for 400-bp double-stranded DNA. **Right.** The dependence of the two current blocking states I_{B1} (*solid circles*) of the translocation events and I_{B0} (*empty diamonds*) of the collision events on nanopore diameter for 400-bp double-stranded DNA. [From Wanunu et al. (2008b).]

(approximately several microseconds per base as opposed to ∼50 ns/base pair reported for solid-state nanopores), and, more important, a strictly linear dependence of the translocation time on DNA length.

In addition to the differences in nanopore geometry, which undoubtedly lead to vastly different DNA–nanopore interactions, structural differences in ssDNA and dsDNA must play a significant role in translocation dynamics. The vast difference in the average velocities may be accounted for by differences in the ssDNA and dsDNA structures. Double-stranded DNA has a much larger coiled structure than ssDNA, and its self-relaxation time (Zimm time) is much longer than the typical translocation time (Storm et al., 2005b; Wanunu et al., 2008b). Thus, with solid-state pores, dsDNA structure is nearly frozen at the time of translocation, leading to large molecule-to-molecule variations. In contrast, ssDNA (and ssRNA) are characterized by a very short persistence length and short relaxation times, in particular much faster than the characteristic translocation times measured with the α-HL system. These differences may partly explain the vastly different dynamics. Studies with smaller solid-state nanopores, and in particular pores that are only slightly larger than the biopolymer cross section, reveal longer translocation time (Wanunu et al., 2008b), approaching those obtained using the protein pore. As our understanding of the translocation dynamics of dsDNA and ssDNA through solid-state nanopores matures, the role of biopolymer–nanopore interactions is becoming better appreciated, leading to new ways to control and tune these complex processes. These will undoubtedly provide novel tools for biomolecular manipulations in the future.

10.5. Summary and Future Prospects

The concept of threading and sequential analysis of biopolymers, coupled with the simplicity of single-molecule detection using nanopores, offers a wide array of potential applications in biophysics and biotechnology. Over the past decade, the toxin channel α-HL has been extensively used to probe secondary structure in nucleic acids and DNA–enzyme interactions, illustrating the great utility of the approach. More recently, synthetic nanopores have begun to emerge, offering new capabilities and features, albeit not free of challenges. Protein channels self-assemble by nature, implying a high level of pore-to-pore reproducibility. In contrast, synthetic nanopores still suffer from pore-to-pore variations due to imperfect fabrication processes, differences in material composition, and variations in postfabrication handling methods. As materials technology advances, more-accurate methods to control these parameters will quickly be developed. We envision that progress with nanopores will rapidly progress to deliver exciting possibilities and applications for biology and genomics. We conclude this chapter by discussing current efforts and future outlooks and challenges for novel nanopore applications, focusing on some of the single-molecule DNA sequencing approaches using nanopores.

10.5.1. DNA Sequencing by Ionic Blockade Measurement

Studies using the protein pore α-HL have demonstrated that the blocked ion current levels of homopolymers of purines can be discriminated from those of pyrimidines (Akeson et al., 1999; Meller et al., 2000). However, the inherent signal averaging of about ten nucleotides residing in the α-HL protein pore at any given time of the DNA sliding (Meller

et al., 2001) and the intrinsic noise of the system (ionic conduction fluctuation) mask the possibility of discriminating individual nucleotides.

Solid-state nanopores have opened new possibilities for this concept because finer nanopore depths (set by the local thickness of the membrane) approaching a single nucleotide (∼5 Å) can be envisioned. This concept, however, must be coupled to a way of controlling the DNA sliding speed: If the latter is governed by nanopore–DNA interactions (Meller, 2003), an ultrathin pore would result in extremely large translocation speeds, presenting a major challenge for this approach. In addition, the mechanical robustness of sub-nanometer-thick pores, which would play a role in reproducibility and durability, is yet to be determined.

Beyond the normal electronic measurement of DNA translocation, concepts of specially designed pores seem to be much more promising. Astier et al. (2006) beautifully demonstrated that the noncovalent incorporation of a positively charged cyclodextrin adapter into an α-HL nanopore results in unique, discernible signals on stochastic binding of each of the four mononucleotides with 93%–98% accuracy. Sequencing using this method requires the incorporation of an exonuclease that can sequentially cleave the DNA bases to be sequenced in the close vicinity of the nanopore. A major challenge to this approach is the synthesis and incorporation of an α-HL/exonuclease hybrid, which is yet to be demonstrated. Another challenge is to ensure that binding of each cleaved mononucleotide into the cyclodextrin adapter occurs exactly once, which may require a specially designed cyclodextrin adapter that facilitates the transport of each captured mononucleotide through the channel.

Sequencing by hybridization is an indirect method for DNA sequencing, which relies on the measurements of single-stranded gaps between hybridized oligonucleotide probes (with known sequences) to the unknown ssDNA target (Ling, 2008). The method is more tolerant than those discussed previously because the nanopore is used only to distinguish between single-stranded and double-stranded current states. However, this approach requires translation of the temporal nature of the current signal to spatial information (number of bases in each gap) with great precision (a few base pairs), a highly challenging feat. Another complication is the possibility of secondary-structure formation in the ssDNA, which is expected to affect the translocation dynamics and to increase the possibilities of folded DNA entry.

10.5.2. DNA Sequencing by Transverse Electronic Measurement

In this proposed method, as DNA slides through the nanopore under an applied voltage, a second set of "nano-electrodes" fabricated perpendicular to the DNA axis is used to identify each translocating nucleotide based on its electronic tunneling probability, which was theoretically calculated to be different for each base. Calculations and molecular dynamics simulations (Zwolak and Di Ventra, 2008) show that under specific experimental conditions (1.5-nm transverse electrode spacing, the bases aligned with the electrodes, and good base coupling to electrodes), transverse current for the nucleotides can be distinguished from each other by a factor of two (between G and C) to a few orders of magnitude (between A and T). In practical experimental conditions, however, coupling of the bases to the transverse nano-electrodes is anticipated to be poor due to hugely varying orientation of the bases between the electrodes, causing order-of-magnitude differences in transverse current and diminishing the contrast among the four nucleotides (Lagerqvist et al., 2006). This problem may be potentially circumvented by applying a strong enough alignment field by the nano-electrodes, as well as by slowing DNA translocation to provide sufficient residence time for alignment along the electric field. Under such external control, the statistical current distributions of the four bases are disjoint, which may allow sequencing.

An interesting alternative was proposed by He et al. (2007), in which hydrogen bond–mediated tunneling currents of individual bases are measured using scanning tunneling microscope (STM) tips functionalized with one of the four different nucleosides. Initial experiments show a difference in the tunneling current profiles for each of the four nucleoside monolayers measured with four different STM tips. Based on the STM measurements, the authors predict that the use of 13 independent reads will reach 99.99% accuracy. However, the fabrication of a nucleoside-functionalized nanopore that permits local tunneling current measurements with a single-DNA-base resolution remains a major challenge of this approach.

10.5.3. DNA Sequencing by Optical Readout of DNA Bases

Meller and coworkers described the feasibility of DNA sequencing based on ultrafast optical readout of labeled DNA polymers sequentially translocating through a nanopore (Lee and Meller, 2007; Soni and Meller, 2007). DNA is first enzymatically converted to a design DNA polymer in which each nucleotide is substituted with a specific combination of an oligo pair. Fluorescently tagged molecular beacons complementary to the two oligos are then hybridized to the concatenated oligo polymers, and the molecule is electrophoretically fed through a small nanopore (<2 nm). The nanopore is designed to sequentially peel off the molecular beacons one after the other, yielding flashes of light in two colors that correspond to the DNA base in the original DNA molecule. The DNA translocation speed is regulated by the DNA unzipping kinetics, and the contrast among bases is achieved through the use of optical probes, circumventing the problems with the electronic nanopore sequencing method. The optical readout lends itself to straightforward parallelism by using high-density arrays of tens or hundreds of nanopores. The major challenges of this approach are the creation of nearly uniform ~2-nm nanopore arrays suitable for DNA unzipping (McNally et al., 2008) and the conversion of long DNA strands to the design DNA format.

Acknowledgments

We acknowledge insightful discussions with M. Akeson, H. Bayley, D. Branton, D. Deamer, A. Kolomeisky, D. Nelson, and B. Shklovskii on various topics related to nanopore science and with many others who contributed ideas and thoughts. We thank members of the Meller group at Boston University for stimulating discussions. AM acknowledges support from National Institutes of Health award HG004128, National Science Foundation awards NIRT 0403891 and PHY-0646637, and Human Frontier Science Program award RGP0036/2005.

References

Akeson, M., D. Branton, J. Kasianowicz, E. Brandin, and D. Deamer. 1999. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* 77:3227–3233.

Aksimentiev, A., and K. Schulten. 2005. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophys. J.* 88(6):3745–3761.

Arnott, S., R. Chandrasekaran, and A. G. W. Leslie. 1976. Structure of the single-stranded polyribonucleotide polycytidylic acid. *J. Mol. Biol.* 106:735–748.

Astier, Y., O. Braha, and H. Bayley. 2006. Toward single molecule DNA sequencing: Direct identification of ribonucleoside and deoxyribonucleoside 5′-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *J. Am. Chem. Soc.* 128:1705–1710.

Bantel, H., B. Sinha, W. Domschke, G. Peters, K. Schulze-Osthoff, and R. U. Jänicke. 2001. α-Toxin is a mediator of *Staphylococcus aureus*–induced cell death and activates caspases via the intrinsic death pathway independently of death receptor signaling. *J. Cell Biol.* 155:637–647.

Bayley, H., and P. Cremer. 2001. Stochastic sensors inspired by biology. *Nature* 413:226–230.

Benner, S., R. J. A. Chen, N. A. Wilson, R. Abu-Shumays, N. Hurt, K. R. Lieberman, D. W. Deamer, W. B. Dunbar, and M. Akeson. 2007. Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. *Nat. Nanotechnol.* 2(11):718–724.

Bezrukov, S. M., I. Vodyanoy, and V. A. Parsegian. 1994. Counting polymers moving through a single ion channel. *Nature* 370:279–281.

Bhakdi, S., and J. Tranum-Jensen. 1991. *Microbiol. Rev.* 55:733–751.

Bustamante, J. O., H. Oberleithner, J. A. Hanover, and A. Liepins. 1995. Patch clamp detection of transcription factor translocation along the nuclear pore complex. *J. Memb. Biol.* 146:253–261.

Chang, H., F. Kosari, G. Andreadakis, M. A. Alam, G. Vasmatzis, and R. Bashir. 2004. DNA mediated fluctuations in ionic current through silicon oxide nano-channels. *Nano Lett.* 4(8):1551–1556.

Chen, P., J. J. Gu, E. Brandin, Y. R. Kim, Q. Wang, and D. Branton. 2004a. Probing single DNA molecule transport using fabricated nanopores. *Nano Lett.* 4(11):2293–2298.

Chen, P., T. Mitsui, D. B. Farmer, J. Golovchenko, R. G. Gordon, and D. Branton. 2004b. Atomic layer deposition to fine-tune the surface properties and diameters of fabricated nanopores. *Nano Lett.* 4(7):1333–1337.

Cockroft, S. L., J. Chu, M. Amorin, and M. R. Ghadiri. 2008. A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *J. Am. Chem. Soc.* 130(3):818–+.

Coulter, W. H. 1953. Means for counting particles suspended in a fluid US. 2,656,508.

Deamer, D. W., and M. Akeson. 2000. Nanopores and nucleic acids: Prospects for ultrarapid sequencing. *Tibtech* 18:147–150.

Deamer, D. W., and D. Branton. 2002. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* 35:817–825.

DeBlois, R. W., and C. P. Bean. 1970. Counting and sizing of submicron particles by the resistive pulse technique. *Rev. Sci. Instrum.* 41:909–915.

Dekker, C. 2007. Solid-state nanopores. *Nat. Nanotechnol.* 2(4):209–215.

Driselkelmann, B. 1994. Translocation of DNA across bacterial membranes. *Microbiol. Rev.* 58:293–316.

Dudko, O. K., J. Mathe, A. Szabo, A. Meller, and G. Hummer. 2007. Extracting kinetics from single-molecule force spectroscopy: Nanopore unzipping of DNA hairpins. *Biophys. J.* 92(12):4188–4195.

Dudko, O.K., G. Hummer, and A. Szabo. 2008. Theory, analysis and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci.* 105:15755–15760.

Evans, E. 2001. Probing the relation between force-lifetime and chemistry in single molecule bonds. *Annu. Rev. Biophys. Biomol. Struct.* 30:105–128.

Evans, E., K. Ritchie, and R. Merkel. 1995. Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces. *Biophys. J.* 68(6):2580–2587.

Fologea, D., E. Brandin, J. Uplinger, D. Branton, and J. Li. 2007. DNA conformation and base number simultaneously determined in a nanopore. *Electrophoresis* 28(18):3186–3192.

Fologea, D., J. Uplinger, B. Thomas, D. S. McNabb, and J. L. Li. 2005. Slowing DNA translocation in a solid-state nanopore. *Nano Lett.* 5(9):1734–1737.

Gershow, M., and J. A. Golovchenko. 2007. Recapturing and trapping single molecules with a solid-state nanopore. *Nat. Nanotechnol.* 2(12):775–779.

Gouax, J. E., O. Braha, M. R. Hobaugh, L. Song, S. Cheley, C. Shustak, and H. Bayley. 1994. Subunit stoichiometry of staphylococcal α-hemolysin in crystals and membranes: A heptameric transmembrane pore. *Proc. Natl. Acad. Sci. USA*. 91:12828–12831.

Grayson, P., L. Han, T. Winther, and R. Phillips. 2007. Real-time observations of single bacteriophage lambda DNA ejections *in vitro*. *Proc. Natl. Acad. Sci. USA*. 104(37):14652–14657.

Guan, X. Y., L. Q. Gu, S. Cheley, O. Braha, and H. Bayley. 2005. Stochastic sensing of TNT with a genetically engineered pore. *Chembiochem* 6(10):1875–1881.

Hanss, B., E. Leal-Pinto, L. A. Bruggeman, T. D. Copeland, and P. E. Klotman. 1998. Identification and characterization of a cell membrane nucleic acid channel. *Proc. Natl. Acad. Sci. USA*. 95:1921–1926.

He, J., L. Lin, P. Zhang, and S. Lindsay. 2007. Identification of DNA basepairing via tunnel-current decay. *Nano Lett.* 7(12):3854–3858.

Healy, K. 2007. Nanopore-based single-molecule DNA analysis. *Nanomedicine* 2(4):459–481.

Hemmler, R., G. Bose, R. Wagner, and R. Peters. 2005. Nanopore unitary permeability measured by electrochemical and optical single transporter recording. *Biophys. J.* 88(6):4000–4007.

Heng, J. B., C. Ho, T. Kim, R. Timp, A. Aksimentiev, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. 2004. Sizing DNA using a nanometer-diameter pore. *Biophys. J.* 87(4):2905–2911.

Henrickson, S. E., M. Misakian, B. Robertson, and J. J. Kasianowicz. 2000. Driven DNA transport into an asymmetric nanometer scale pore. *Phys. Rev. Lett.* 85:3057–3060.

Ho, C., R. Qiao, J. B. Heng, A. Chatterjee, R. J. Timp, N. R. Aluru, and G. Timp. 2005. Electrolytic transport through a synthetic nanometer-diameter pore. *Proc. Natl. Acad. Sci. USA*. 102(30):10445–10450.

Hornblower, B., A. Coombs, R. D. Whitaker, A. Kolomeisky, S. J. Picone, A. Meller, and M. Akeson. 2007. Single-molecule analysis of DNA-protein complexes using nanopores. *Nat. Methods* 4(4):315–317.

Hugel, T., and M. Seitz. 2001. The study of molecular interactions by AFM force spectroscopy. *Macromol. Rapid Commun.* 22(13):989–1016.

Ippen-Ihler, K. 1989. Bacterial conjugation. In *Gene Transfer in the Environment*, Levy SB, Miller RV, editors. McGraw-Hill, New York, pp. 33–72.

Iqbal, S. M., D. Akin, and R. Bashir. 2007. Solid-state nanopore channels with DNA selectivity. *Nat. Nanotechnol.* 2(4):243–248.

Jan Bonthuis, D., J. Zhang, B. Hornblower, J. Mathe, B. I. Shklovskii, and A. Meller. 2006. Self-energy-limited ion transport in subnanometer channels. *Phys. Rev. Lett.* 97:128104–205302.

Jonas, D., I. Walev, T. Berger, M. Liebetrau, M. Palmer, and S. Bhakdi 1994. Novel path to apoptosis: Small transmembrane pores created by staphylococcal alpha-toxin in T lymphocytes evoke internucleosomal DNA degradation. *Infect. Immun.* 62:1304–1312.

Kasamatsu, H., and A. Nakanishi. 1998. How do animal DNA viruses get to the nucleus? *Annu. Rev. Microbiol.* 52:627–686.

Kasianowicz, J., E. Brandin, D. Branton, and D. Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA*. 93:13770–13773.

Keyser, U. F., D. Krapf, B. N. Koeleman, R. M. M. Smeets, N. H. Dekker, and C. Dekker. 2005. Nanopore tomography of a laser focus. *Nano Lett.* 5(11):2253–2256.

Kim, M.J., B. McNally, K. Murata, and A. Meller. 2007. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology* 18(20).

Kim, M.-J., M. Wanunu, C. D. Bell, and A. Meller. 2006. Rapid fabrication of uniform size nanopores and nanopore arrays for parallel DNA analysis. *Adv. Mater.* 18:3149–3153.

Kokjohn, T. A. 1989. Transduction: Mechanism and potential for gene transfer in the environment. In *Gene Transfer in the Environment*, S. B. Levy and R. V. Miller, editors. McGraw-Hill, New York, pp. 73–97.

Lagerqvist, J., M. Zwolak, and M. Di Ventra. 2006. Fast DNA sequencing via transverse electronic transport. *Nano Lett.* 6(4):779–782.

Lee, J. W., and A. Meller. 2007. Rapid DNA sequencing by direct nanoscale reading of nucleotide bases on individual DNA chains. In *Prospective in Bioanalysis*, K. R. Mitchelson, editor. Elsevier, Burlington, MA, pp. 245–264.

Li, J., M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. 2003. DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* 2(9):611–615.

Li, J., D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko. 2001. Ion-beam sculpting at nanometre length scales. *Nature* 412:166–169.

Ling, X. S. 2008 (Private communication).

Mathé, J., A. Aksimentiev, D. R. Nelson, K. Schulten, and A. Meller. 2005. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc. Natl. Acad. Sci. USA*. 102(35):12377–12382.

Mathé, J., A. Arinstein, Y. Rabin, and A. Meller. 2006. Equilibrium and irreversible unzipping of DNA in a nanopore. *Europhys. Lett.* 73(1):128–134.

Mathé, J., H. Visram, V. Viasnoff, Y. Rabin, and A. Meller. 2004. Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.* 87(5):3205–3212.

McNally, B., M. Wanunu and A. Meller. 2008. Electro-mechanical unzipping of individual DNA molecules using synthetic sub.2 nm pores. *Nano Lett*. 8:3418–3422.

Meller, A. 2003. Dynamics of polynucleotide transport through nanometre-scale pores. *J. Phys. Condens. Matter* 15:R581–R607.

Meller, A., and D. Branton. 2002. Single molecule measurements of DNA transport through a nanopore. *Electrophoresis* 23:2583–2591.

Meller, A., L. Nivon, E. Brandin, J. Golovchenko, and D. Branton. 2000. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. USA*. 97:1079–1084.

Meller, A., L. Nivon, and D. Branton. 2001. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* 86:3435–3438.

Miller, R.V. 1998. Bacterial gene swapping in nature. *Sci. Am.* 278:67–71.

Nakane, J., M. Wiggin, and A. Marziali. 2004. A nanosensor for transmembrane capture and identification of single nucleic acid molecules. *Biophys. J.* 87(1):615–621.

Noskov, S.Y., W. Im, and B. Roux. 2004. Ion permeation through the alpha-hemolysin channel: Theoretical studies based on Brownian dynamics and Poisson-Nernst-Plank electrodiffusion theory. *Biophys. J.* 87(4):2299–2309.

Reiner, A., D. Yekutieli, and Y. Benjamini. 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 19(3):368–375.

Saenger, W. 1988. *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Salman, H., D. Zbaida, Y. Rabin, D. Chatenay, and M. Elbaum. 2001. Kinetics and mechanism of DNA uptake into the cell nucleous. *Proc. Natl. Acad. Sci. USA* 98(13):7247–7252.

Sanchez-Quesada, J., A. Saghatelian, S. Cheley, H. Bayley, and M. R. Ghadiri. 2004. Single DNA rotaxanes of a transmembrane pore protein. *Angew. Chem. Int. Ed.* 43(23):3063–3067.

Sauer-Budge, A. F., J. A. Nyamwanda, D. K. Lubensky, and D. Branton. 2003. Unzipping kinetics of double-stranded DNA in a nanopore. *Phys. Rev. Lett.* 90(23):238101.

Schatz, G., and B. Dobberstein. 1996. Common principles of protein translocation across membranes. *Science* 271:1519–1525.

Smeets, R. M. M., U. F. Keyser, D. Krapf, M. Y. Wu, N. H. Dekker, and C. Dekker. 2006. Salt dependence of ion transport and DNA translocation through solid-state nanopores. *Nano Lett.* 6(1):89–95.

Song, L., M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, and J. E. Gouax. 1996. Structure of staphylococcal α-hemolysin a heptameric transmembrane pore. *Science* 274:1859–1865.

Soni, G. V., and A. Meller. 2007. Progress toward ultrafast DNA sequencing using solid-state nanopores. *Clin. Chem.* 53(11):1996–2001.

Storm, A. J., J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. 2003. Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* 2(8):537–540.

Storm, A. J., J. H. Chen, H. W. Zandbergen, and C. Dekker. 2005a. Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E* 71(5)051903.

Storm, A. J., C. Storm, J. H. Chen, H. Zandbergen, J. F. Joanny, and C. Dekker. 2005b. Fast DNA translocation through a solid-state nanopore. *Nano Lett.* 5(7):1193–1197.

Szabò, I., G. Bàthori, F. Tombola, M. Brini, A. Coppola, and M. Zoratti. 1997. DNA Translocation across planar bilayers containing *Bacillus subtilis* ion channels. *J. Biol. Chem.* 272:25275–25282.

Vercoutere, W., S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, and M. Akeson. 2001. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotech.* 19:248–252.

Wanunu, M., B. Chakrabarti, J. Mathe, D.R. Nelson, and A. Meller. 2008a. Orientation-dependent interactions of DNA with an alpha-hemolysin channel. *Phys. Rev. E* 77(3)031904.

Wanunu, M., and A. Meller. 2007. Chemically modified solid-state nanopores. *Nano Lett.* 7(6):1580–1585.

Wanunu, M., and A. Meller. 2008. Single-molecule analysis of nucleic acids and DNA–protein interactions using nanopores. In *Single-Molecule Techniques: A Laboratory Manual*, P. Selvin and T. J. Ha, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 395–420.

Wanunu, M., J. Sutin, B. McNally, A. Chow, and A. Meller. 2008b. Unfolded DNA translocation governed by interactions with solid state nanopore. *Biophys. J.* 95, 4716–4725.

Whittaker, G. R., and A. Helenius. 1998. Nuclear import and export of viruses and virus genomes. *Virology* 246: 1–23.

Winters-Hilt, S., W. Vercoutere, V. S. DeGuzman, D. Deamer, M. Akeson, and D. Haussler. 2003. Highly accurate classification of Watson-Crick basepairs on termini of single DNA molecules. *Biophys. J.* 84(2):967–976.

Zwolak, M., and M. Di Ventra. 2008. Colloquium: Physical approaches to DNA sequencing and detection. *Rev. Mod. Phys.* 80(1):141–165.

11

Nanopores: Generation, Engineering, and Single-Molecule Applications

Stefan Howorka and Zuzanna Siwy

Abstract Nanopores enable the sensing of individual molecules based on the temporary blockades in ionic pore current. Initially conducted a decade ago with a biological protein pore, electrical recordings are now routinely performed with synthetic pores sculptured into polymeric and inorganic membranes. Assisted by channel engineering, the range of analytes has been expanded from nucleic acids to peptides, proteins, organic polymers, and small molecules. Apart from being an attractive analytical approach, nanopore recording has developed into a general platform technology with which it is possible to examine the biophysics, physicochemistry, and chemistry of individual molecules. Nanopores can also be exploited for separation technologies and nanofluidics due to their ability to control the flow of solvated ions. The combined use with atomic force and fluorescence microscopy is extending the versatility of nanopores for single-molecule research.

11.1. Introduction

The ability to measure ionic current flowing through individual membrane protein channels [1–4] has had a tremendous impact on the understanding of biological ion channels [5–7]. More recently, single-channel current recordings have been applied to nanopore analysis. In this approach, pores with an inner diameter of more than 1 nm are employed as instruments to detect individual molecules. The analytes pass through or bind to the pore, thereby modulating the ionic pore current in a characteristic fashion. First demonstrated with poly(ethylene glycol) [8], nucleic acid strands [9], and metal ions [10], the electrical and label-free sensing strategy has gained in popularity and is now used for a wide range of different analytes, including small molecules, organic polymers, proteins, and biomolecular complexes [11a]. Apart from being an attractive analytical approach, nanopore recording has developed into a general platform technology with which it is possible to examine the biophysics, physicochemistry, and chemistry of individual molecules. The scientific advances have been accompanied by an increase in the variety of nanopores. While initial experiments

S. Howorka • Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

Z. Siwy • Department of Physics and Astronomy, University of California at Irvine, 210G Rowland Hall, Mail Code: 4575, Irvine, CA 92697, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_11,

were solely performed with natural protein pores, synthetic nanopores in organic polymers or inorganic materials such as silicon nitride, silicon dioxide, silicon, or silicate are now used on a routine basis as described in several review articles [11b].

This chapter summarizes the achievements of nanopore analytics over the last 10 years. It covers, first, how biological, chemical, and solid-state nanopores are obtained and engineered to enhance their functionality. Second, an overview is provided on how nanopores are exploited to detect molecules and to study their behavior. Particular emphasis is given to molecules other than nucleic acids because these are covered in a dedicated book chapter by Meller within this book [12], and several review articles [13–16c]. Emerging applications of nanopores in nanofluidics and separation are also described. The chapter closes with an outlook on how the field of nanopore analytics may develop.

11.2. Principles of Nanopore Analytics

The principle of nanopore analytics is related to the Coulter counter method. In Coulter counting, microscale objects pass through a microscale aperture and cause detectable short modulation in the ionic current flowing through the aperture [17]. As implied by its name, a Coulter device counts but also sizes microparticles and cells [18,19]. Nanopore analytics uses the same sensing principle (Figure 11.1) and translates the movement of an analyte through the pore into short modulations in electrical conductance. Given the similarity of both approaches, the overarching term "resistive-pulse sensing" has been coined [18]. Nanopores feature, however, a much smaller inner diameter, between 1 and ~500 nm. The narrow width makes nanopores ideally suited to detect molecules, in that molecular-sized objects block a nanopore to a much bigger extent than a microscale pore.

The experimental setup for a nanopore recording is simple and comprises a biological, inorganic, or polymeric membrane with a single nanopore (Figure 11.1). The nanopore connects two electrolyte reservoirs on both sides of the membrane. Two immersed electrodes and a potentiostat are used to set up a transmembrane potential that drives an ionic current through the pore. A current amplifier measures the ionic pore current as a function of time.

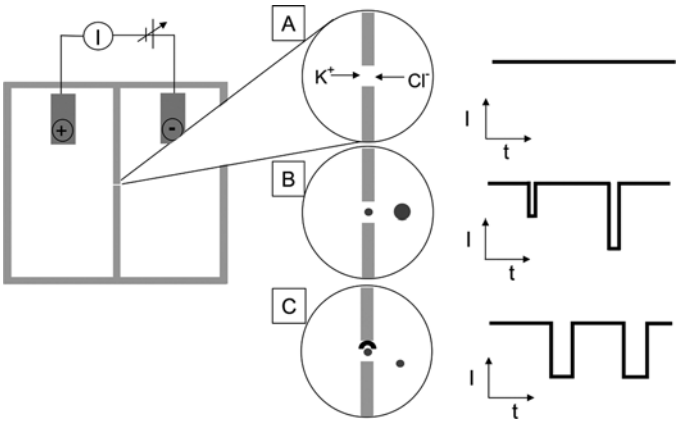


Figure 11.1. Scheme illustrating the principle of nanopore analytics. **A.** A transmembrane potential causes a constant flow of ions through a single pore. **B**. Analytes that pass the pore cause temporary current blockades. **C.** Analytes bind reversibly to an engineered binding site within the pore and produce current fluctuations.

For many widely used pores, the conductance is unitary, as indicated by a constant-current trace (Figure 11.1A). When analyte molecules are added to one reservoir, the ensuing passage through the pore leads to transient conductance blockades (Figure 11.1B). The modulations in current may also reflect molecules that are binding to molecular recognition sites at the pore wall (Figure 11.1C). Depending on the nature and timescale of the analyte pore interactions, current modulations other than discrete blockades are possible. For example, a strongly interacting analyte might block the pore current permanently, whereas rapidly passing molecules might be too fast to be resolved, thereby giving rise to an increased current noise. Any discrete current blockades can be measured with a temporal resolution of down to 50 μs and a maximal conductance resolution of 0.2% of the open-channel current. The actual resolution depends on many factors, such as the current amplifier, electrodes, membrane, pore, analyte, and the electrolyte concentration.

A crucial step in the analysis of nanopore traces is to ascribe the discrete current blockades to individual molecules. The assignment is unambiguous when the pore dimensions allow the passage of only one molecule at a time [9]. In the case of much smaller molecules, low concentrations of analytes can help to avoid the simultaneous passage of multiple molecules and the possible masking of individual signals. The genuine molecular origin of individual current blockades can also be supported by varying the analyte concentration because a lower concentration should proportionally decrease the frequency with which the blockades occur in the traces.

Nanopore recording is a simple technique for detecting individual molecules. Unlike other single-molecule approaches, including fluorescence-based techniques [20–23], optical tweezers [24–29], and atomic force microscopy [30], nanopore recording does not require that the analyte be fluorescence labeled or immobilized on a solid surface. This benefit is brought about because unlabeled and water-soluble molecules give rise to detectable current blockades. Due to the simple electrical sensing principle, nanopores probably offer the least complex and cheapest method for detecting single molecules.

As will be described further below, nanopore analytics represents a general platform technique with which to examine the physicochemistry, biophysics, chemistry, and electrostatics of a broad spectrum of molecules. Reflecting this versatility, many different aspects of molecular dynamics can be studied. For example, using the model system of poly(ethylene glycol) (PEG), nanopores can help investigate how hydrophilic linear polymers enter nanoscale pores and stay within an entropically disfavored confined space [31–33]. In biophysics, pores carrying analyte-specific molecular recognition sites [34] can be applied to examine the kinetics and thermodynamics on the formation of DNA duplexes [35] and biomolecular interaction between proteins and small-molecule ligands [36a]. Furthermore, sensing can also be extended to stoichiometric complexes between nucleic acids and proteins [37–40]. In this case, the transmembrane potential can be used as an experimental tool to control the force with which a binding partner is pulled through the pore. This opens up the opportunity to probe the binding strength within the biomolecular complexes via single-molecule force spectroscopy. In chemistry, the ability to detect molecules without any label can help examine the kinetics of processes that would otherwise remain undetected. Examples are the formation of short-lived intermediates in the breaking and making of bonds [41] and the fast switching between different isomeric forms [42]. Finally, the electrostatic interaction between charged peptide and the pore walls enables the study of electrostatics at the nanoscale [43,44a].

The methodology of nanopore recording can also be of biological relevance. In many biological processes, molecules pass through confined openings and interact with the walls

of the nanoscale channel. Relevant examples include the passage of DNA, polypeptides, and nutrients through bacterial and eukaryotic membrane pores and—in a wider sense—also the movement of nucleic acids strands through soluble ring-shaped polymerases and motor proteins. Nanopores can help shed light on the biophysical principles underlying these important biological systems.

Many widely used nanopores are "blank" and lack any intrinsic ability to bind analytes. In comparison, sensing can also be achieved with natural pores that were built by nature to recognize molecules. For example, bacterial maltoporin pores facilitate the passage of maltodextrin sugars [45], and the bacterial OmpF porin pores pass antibiotics [46]. In both cases, analytes permeating through the pore lead to current blockades. In addition, the mitochondrial VDAC channel exhibits conductance changes for the nucleotides adenosine triphosphate (ATP), uridine triphosphate (UTP), and nicotinamide adenine dinucleotide (NADH) [47] but not for the synthetic anions tetraglutamate and 1-hydroxypyrene-3,6,8-trisulfonate [48]. Another class of membrane proteins, ligand-gated channels, has the innate ability to translate the specific binding of ligands into the opening and closing of the channel. Indeed, acetylcholine receptors are able to detect low levels of acetylcholine [18, 49, 50]. Although these pores and channels are valuable biosensor elements, this chapter focuses on solid-state nanopores, as well as on blank and reengineered protein pores.

11.3. Pores

Nanopore analytics was pioneered with protein pores [9,51], but the field now encompasses solid-state nanopores formed in inorganic materials such as Si_3N_4, Si, and SiO_2 and polymeric membranes composed of poly(ethylene terephthalate) and the polyimide Kapton. Despite differences in their geometry and chemistry, most pores tend to share four characteristic features. First, the pore width is in the same size range as the dimension of the analyte. Molecules that are less than 1/20 in size are usually too small to cause an appreciable modulation in pore current. Second, the pores are able to sustain a constant current for a prolonged period of time. They do not open or close in the absence of analyte. In structural terms, this translates into stable pores devoid of any moving parts. A constant current is preferred over a fluctuating current because the unitary conductance makes it easier to detect analyte-induced blockades. Third, most nanopores do not feature any intrinsic ability to specifically bind analytes. In combination with a diameter of at least 1 nm, this enables nanopores to accommodate a wide range of analytes. Nanopores can, however, be engineered to specifically bind analytes. Fourth, the channel walls of the pores are amenable to modification with chemicals. Altering the properties of the wall provides a useful strategy to either introduce the specific binding to analytes or avoid nonspecific interactions with other nontarget components present in the analyte-containing sample.

11.3.1. Biological and Chemical Pores

Several different biological membrane pores have been applied for nanopore sensing. The most widely used are the bacterial α-hemolysin toxin, porins from the bacterial outer membrane, the gramicidin peptide antibiotics, and gramicidin-derived channels. Less frequently used pore structures of transient nature, such as those formed by α-helical alamethicin peptides, are not discussed in detail [8]. *De novo*-designed proteins pores are available [52,53] but their potential for single-molecule sensing has not been exploited. Similarly, among the

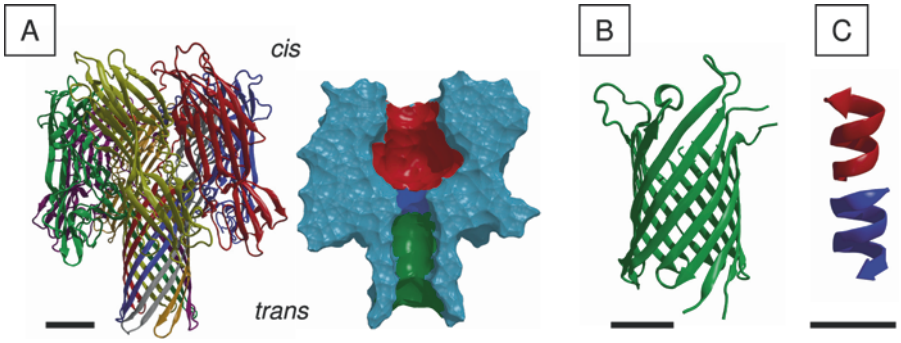


Figure 11.2. Biological nanopores. **A.** The heptameric α-hemolysin pore. The cross-sectional view on the right displays the inner cavity (*red*), inner constriction (*blue*), and β-barrel (*green*). **B.** The OmpG porin. **C.** the gramicidin dimer. Scale bars, 2 nm.

great variety of synthetic chemical pores with biologically inspired [54] and nonbiological scaffolds [55,56], few have been applied for nanopore analytics at the single-molecule level. Their discussion in this chapter is restricted to two exemplary pores.

The α-hemolysin (αHL) protein pore is formed by self-assembly of seven identical polypeptides (Figure 11.2A). The assembled heptameric αHL pore is of known structure and resembles a mushroom with a wide cap and a narrow stem that spans the lipid bilayer [57]. The external dimensions of the pore are 10 × 10 nm. The central channel is 2.9 nm in diameter at the *cis* entrance and widens to 4.1 nm in the internal cavity (Figure 11.2A). In the transmembrane region, the lumen narrows to 1.3 nm at the inner constriction and broadens to 2 nm at the *trans* entrance of the β-barrel. The β-barrel is composed of 14 antiparallel strands, two from each subunit. The defined yet robust structure of αHL has been key to achieving its successful engineering [34]. One peculiarity of αHL is its sevenfold symmetry. Due to its assembly into heptamers, a single mutation in the subunit is usually multiplied in the pore structure. Pores carrying a single modification can be obtained by isolating heptamers composed of one modified and six unmodified subunits [10,35,58].

Porins are pores from the outer membrane of gram-negative bacteria. Porin pores have a diameter of around 2 nm and are formed from a single polypeptide chain folded into a β-barrel of 16 or 18 antiparallel β-strands [59]. Most porins are trimers of β-barrels, but monomeric porins such as OmpG are also known (Figure 11.2B) [60]. Its monomeric nature is attractive compared to multimeric αHL because a single point mutation can be easily achieved within the pore. The lumen of porins can, however, contain flexible peptide loops that give rise to undesired gating.

Gramicidin channels are composed of two β-helical peptides that dimerize to form an elongated membrane-spanning tube (Figure 11.2C) [61]. The sequence of the peptides alternates between hydrophobic L- and D-amino acids, thereby positioning the hydrophobic residues at the outside of the channel that interacts with the lipid chains. The two helical peptides of gramicidin channels reside in the two membrane leaflets and can dissociate to migrate laterally within the membrane plane. The transient nature of the channels is obvious in electrical recordings and can be correlated to changes in single-molecule fluorescence energy transfer [62]. Covalently cross-linking of the two helical peptides decreases the current fluctuations [63].

Organic synthesis from scratch is another source of pores. The synthetic route is attractive because readily available, tailored chemical building blocks can be exploited to obtain

structures with enhanced functional properties. Ghadiri's peptide-based pores are one type of synthetic structure [64]. The pores are composed of stacks of peptide rings whose sequences alternate between L- and D-amino acids, similar to gramicidin. The pore diameter can be tuned by varying the size of the cyclic peptides [54]. The rings are held together by a network of hydrogen bonds, but the interaction can be stabilized by covalent linkages between the amino acid side chains [65]. Other examples of synthetic constructs are pores from Matile's group [56]. These pores are made from rigid octi(*p*-phenylene) rods that are laterally functionalized with peptides. The peptide arms interact and self-assemble the rods into a barrel-like structure. The pentapeptides can carry different chemical groups to enhance their functionality [55]. Although not made of synthetic materials, nanopores composed of lipids have also been generated [66]. The pores with a diameter of 50–150 nm have served as tracks for the electrophoretic migration of individual DNA molecules as observed via fluorescence microscopy.

11.3.2. Engineering of Biological and Chemical Nanopores

What are the benefits of engineering biological and chemical nanopores? Which functional properties can be given to the pores? In very general terms, pores are engineered to remove undesired properties that would interfere with sensing, such as the gating caused by a moving polypeptide loop. Engineering is also applied to alter how the pore interacts with electrolyte ions. The most widely used objective of engineering is to introduce new receptor or binding sites for analytes.

In the case of protein pores, these goals can be achieved by genetic engineering. Recombinant DNA technology is routinely used to add, delete, or replace amino acids [67]. Genetic engineering in combination with targeted chemical modification is a particularly powerful approach because, for example, site-specifically engineered cysteines can be covalently coupled to synthetic chemical compounds. Amino acids with synthetic moieties can alternatively be incorporated with a suppressor tRNA system [68] or expressed protein ligation [69,70], but these approaches have so far not been fully exploited for the engineering of nanopores. Structural information about the pore, such as an X-ray structure, is crucial to helping achieve rationally designed pores. Equally important is that the protein withstands the genetic and chemical modification, as well as the subsequent purification procedures. The protein pore αHL is exemplary in its ability to accommodate drastic genetic and chemical modification. In addition, the αHL pore withstands purification via sodium dodecyl sulfate gel electrophoresis, which usually denatures other proteins [10,58]. This has enabled the isolation of protien pores carrying a single modified subunit [10,35]. In the case of small peptide gramicidin channels, modification can be easily achieved by chemically coupling extensions to the N- or C-terminus of an existing peptide [71] or via *de novo* synthesis using solid-phase peptide synthesis and optionally native chemical ligation [72].

Which favorable properties have been embedded into protein pores with genetic engineering and/or chemical modification? Following the foregoing list of desirable properties, engineering has been essential to quiet the noisy porin OmpG pore for single-molecule sensing. This was achieved by stabilizing and trimming movable peptide loops in the OmpG pore through the introduction of a disulfide bond and by the deletion of amino acids [73]. Furthermore, engineering has been successfully applied to alter the interaction of pores with electrolytes. For example, changes in the voltage-dependent pore conductance were achieved by placing engineered cysteines into αHL and modifying them with charged chemical reagents [74]. In addition, genetic substitutions to charged amino acid have turned the nonrectifying

OmpF porin into a nanofluidic diode [75a]. Adding two glutamates into the constriction zone of the OmpF porin changed its selectivity towards calcium over sodium ions [75b]. This calcium selectivity is observed with biological L-type calcium channels.

Engineering also enhances pores by the introduction of artificial analyte-binding sites. Three different approaches toward the engineering of binding sites have been developed. Most have been pioneered on αHL by the Bayley group [34]. The first approach uses amino acid replacement to create a receptor site. For example, a histidine patch is able to bind metal ions [10] while a ring of small hydrophobic side chains interacts with aromatic explosives such as trinitrotoluene (TNT) [76]. Alternatively, the extension of one of the protein chains can create a specific attachment site for a kinase [77]. Second, nonproteinogenic receptors can be covalently attached to the pore via one or more engineered cysteines. This approach has been demonstrated using pores carrying a single DNA oligonucleotide [35,78] or a ligand-terminated PEG chain [79,80] to capture complementary DNA and antibodies, respectively. Finally, binding sites for small analytes such as drugs and organic solvents have been created by noncovalent adapters. Adapters are ring-shaped molecules that lodge inside a pore while simultaneously mediating the binding of smaller analytes [81]. Noncovalent adapters can be attached to the channel wall [82], but their main advantage lies in the ability to transplant a known binding site into a new structural framework. This has been shown for embedding an adapter originally used for αHL into the OmpG porin structure [73]. The concept of noncovalent adapters can be applied to different pores, such as synthetic channels [83].

11.3.3. Biological Nanopores in Lipid Membranes

Unlike solid-state nanopores, protein pores are usually first generated in aqueous solution and then inserted into a lipid bilayer that is traditionally formed across 30- to 100-μm-diameter orifices in hydrophobic polymer-based membranes [84]. The procedure involves adding a pore solution to the electrolyte-filled chamber and waiting until a single channel inserts into the lipid bilayer membrane. While this method is sufficient for many applications, it can also be desirable to exert more control over when and how many channels incorporate. Holden et al. developed an approach that enables the direct introduction of single protein channels and pores into lipid bilayers [85]. The approach is based on a hand-operated hydrogel probe that is coated with a layer of proteins and mechanically engaged with the lipid bilayer. The methodology has been extended to directly transfer membrane proteins from bacteria to planar bilayers for rapid screening by single-channel recording [86].

Lipid bilayer instability is another topic of relevance to recordings with protein pores. Ideally, the membrane should be stable to permit long recordings or to create durable sensing devices that withstand pressure fluctuations or mechanical vibrations. However, due to their noncovalent nature, lipid bilayers are fragile and often rupture. The tendency to break can be contained by decreasing the size of the lipid bilayer, as demonstrated by using a 2-μm orifice in a polymeric membrane [87] and by forming a lipid bilayer across a 100- to 400-nm-wide aperture in a glass nanopipette [88]. In the latter case, the glass surface had been made hydrophobic by chemical modification of 3-cyanopropyl-dimethylchlorosilane to facilitate the formation of the lipid bilayer [88]. The glass nanopore system combined several advantageous properties, such as seal resistances of approximately 70 GΩ, electrical breakdown voltages of approximately 0.8 V, small bilayer capacitance, mechanical stability, and enhanced bilayer lifetime. The functionality of this system was demonstrated by inserting α-hemolysin and detecting individual ring-shaped cyclodextrin molecules.

The use of chemical reagents can also help to stabilize membranes with reconstituted protein pores. Chemical strategies can encompass polymerizable phospholipids that are cross-linked after the protein has inserted [89] or the encapsulation of the native lipid bilayer membrane into a polymer hydrogel via *in situ* photopolymerization [90]. Alternatively, the bilayer can be sandwiched between two layers of agarose gel [91,92] or by supporting the membrane structure with a polymeric PEG layer [93]. A potential drawback of hydrogels is the need for the analyte to diffuse across the support layer, which can lead to limited temporal sampling responses of the pore systems. This potential drawback can be mitigated by using a hydrogel only on one side of the bilayer and adding the analysis sample to the non-supported side.

Another approach to stabilizing membranes involves covalently linking the bilayer to a solid support structure, such as gold. The process starts with tethering a lipid monolayer to a support via a spacer unit. The outer leaflet of a bilayer is then obtained by depositing vesicles, which fuse into a contiguous layer. Tethered support structures generated by this procedure were described to increase the stability of the bilayer while preserving the ability to incorporate proteins [94–96]. The usefulness of the supported bilayers for current recordings depends on the size of the electrolyte reservoir between the conducting substrate and the membrane.

Solid support structures without chemical linkage represent another strategy. For example, porous alumina has been used to stabilize pore-suspended lipid bilayers [97]. One of the challenges in this system was to form a tight interface between membrane and alumina to obtain a GΩ seal required for current recordings. This might be achieved by using a polymer cushion between solid support and lipid bilayer [97,98].

Independent of the support strategy, the long-term stability of bilayers is compromised by the diffusion of lipids molecules out of the membrane into the solvent. Recently, this limitation has been addressed by replenishing lost membrane components using a pool of lipids. The lipid reservoir consisted of an oil/lipid mixture with suspended lipid-coated water droplets [99]. When the droplets were joined together, the lipid monolayers surrounding them combined at the interface to form a robust lipid bilayer. Various protein channels and pores were incorporated into the droplet-interface bilayer (DIB). By applying a potential difference with electrodes embedded within the droplets, the ionic current flowing through the pores was measured. In addition, the lipid composition of the vesicles was tuned and α-helical, as well as β-barrel, membrane proteins inserted into the DIB [100]. Based on their inherent stability, the droplets may be used in future as devices for bilayer recording with greatly reduced electrolyte volume.

11.3.4. Solid-State and Polymer Nanopores

Experiments with protein pores have stimulated the scientific community to develop techniques for preparing single solid-state nanopores. Solid-state nanopores have potential advantages because they are mechanically and chemically more stable, as well as easier to handle, than biological channels. For example, a membrane with a solid-state nanopore can be used for weeks and exchanged between different laboratories. Apart from facilitating scientific research, synthetic pores are beneficial for the development of rugged commercial sensor devices. Solid-state nanopores can also offer flexibility in the pore diameter, length, and shape. Techniques that are currently used for single-nanopore preparation can be divided into two major categories (Figure 11.3): (1) the ion-track etching method and (2) the drilling and shaping of pores in silicon nitride and silicon oxide materials. There has also been

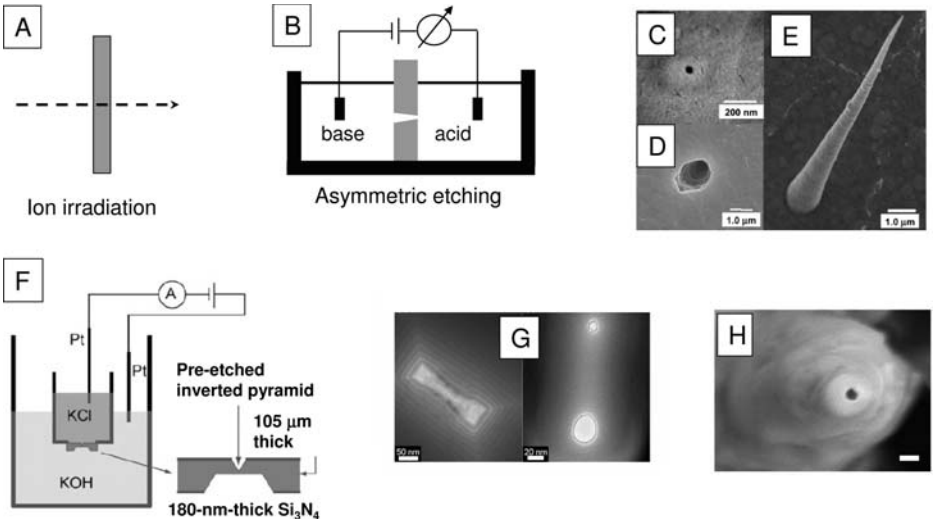


Figure 11.3. Preparation of solid-state nanopores by asymmetric chemical etching and glass nanopipettes. **A–E.** Preparation of single conically shaped nanopores by the track-etching technique, which consists in irradiating a polymer film with a single energetic heavy ion **(A)**, followed by asymmetric etching performed in a conductivity cell **(B)**. The narrow **(C)** and the big opening **(D)** of the resulting conical nanopore are shown. (Reprinted with permission from Harrell et al. [161]. Copyright 2006 American Chemical Society.) **E.** Gold replica of a single conical nanopore. (Reprinted with permission from Harrell et al. [161]. Copyright 2006 American Chemical Society.) **F.** Experimental setup used for the preparation of pores in silicon [115]. (Reprinted with permission from Park et al. [115]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.) **G.** A pre-etched silicon pyramid is placed between solutions of KOH and KCl. This process typically results in the formation of twin pores. Scale bars are 50 nm (left panel) and 20 nm (right panel). (Reprinted with permission from Park et al. [115]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.) **H.** Scanning electron micrograph of a glass nanopipette; the scale bar is 100 nm. (Reprinted with permission from Umehara et al. [126]. Copyright 2006 American Chemical Society.)

research on preparing glass pipettes with nanometer openings, which have been applied to study the translocation of single molecules and examine transport properties of nanometric channels.

Track-Etching Technique

Preparing pores by the track-etching technique is the oldest approach and was first applied to micrometer-sized pores [101]. The technique uses energetic heavy ions that pass through a thin polymer foil to create a localized damaged zone, the so-called latent track. Each heavy ion creates one latent track that can be subsequently developed into a pore using wet chemical etching. Typically, ions of Xe, Pb, Au, and U with kinetic energies in the giga eV range are applied. At the Gesellschaft für Schwerionenforschung (GSI), Darmstadt, Germany, a linear accelerator (UNILAC) produces a beam of ions of 11.4 MeV/u energy that is equivalent to $\sim$10% of the velocity of light [102]. This energy allows the heavy ions to penetrate a 100-µm-thick polymer material or simultaneously irradiate eight foils of 12 µm thickness. Heavy-ion irradiation can also be performed in several accelerator facilities (e.g., Brookhaven National Laboratory, Oak Ridge National Laboratory, and Lawrence Berkeley National Laboratory).

Given that the irradiation current densities determine the number of pores, there are currently two ways of preparing single-pore membranes. The first involves irradiating a polymer film with a defocused heavy-ion beam resulting in densities of $\sim 10^4$ ions/cm^2. The tracks are subsequently chemically etched, and membrane areas with just one pore are isolated. The process of identifying a single pore can be facilitated by using a light microscope for larger pores and by indirect visualization with fluorophores for smaller pores [103]. Another way of preparing single-pore membranes involves tuning the intensity of heavy-ion irradiation to limit the number of passing ions to one [105]. GSI has worldwide facilities designed for single-ion irradiation, and these irradiated samples are readily available.

The size and shape of the pores are controlled in the etching step by varying the type and concentration of the etchant, the reaction temperature, and the duration of the etching process. Conically shaped single nanopores have attracted special interest (Figure 11.3A–E). Compared to a cylindrical pore of the same limiting diameter [105,106], a conical pore has a much lower electrical resistance and is therefore easier to analyze. The method of obtaining conically shaped nanopores is based on asymmetric etching of irradiated films. In the case of poly(ethylene terephthalate) (PET) and polycarbonate membranes, conically shaped pores are obtained by placing an irradiated foil between a concentrated base on one side and an acidic solution on the other [107–109]. The method can be further improved by simultaneously recording the electrical current. In a feedback loop, the etch rate is altered by changing the effective concentration of hydroxide ions in the pore via a different voltage [105,110–112].

Conically shaped pores in polyimide foils (50HN Kapton, DuPont) are also prepared by asymmetric etching but performed in sodium hypochlorite when using potassium iodide as a stopping medium [113]. The track-etching technique allows preparation of pores with effective diameters as small as 2 nm [106,113,114]. Nanopores with diameters greater than 30 nm can be easily imaged using scanning electron microscopy. Most probably due to polymer surface roughness, imaging of smaller pores has not been yet successful, and their diameter is determined by conductivity methods [105].

Pores in Silicon Obtained by Asymmetric Etching

The concept of asymmetric chemical etching was also applied to silicon films [116]. Crystalline silicon etches along lattice planes, and pyramidal indentations can be easily obtained by using a square aperture in the etch mask. Using this procedure, Park et al. obtained a deep indentation in a 105-μm-thick wafer, which was subsequently transferred to a conductivity cell [115] (Figure 11.3F–G). The wafer was put in contact with KOH on one side and KCl on the other. The transmembrane current was measured during the etching process to detect the moment of breakthrough and stop the etching step. The experimental setup was more complicated than in the case of the ion-track–etched samples due to conductive properties of silicon, as well as to an electrochemical potential drop between the silicon surface and the solutions. This nanofabrication process typically resulted in the preparation of twin pores with a diameter of $\sim$20 nm (Figure 11.3G).

Ion-Beam Sculpting

Another technique for the preparation of single nanopores was developed by Li et al. [116] The approach is based on ion beams of much lower energies than those used in the track-etching technique. This ion-beam sculpting method was first applied to generate single

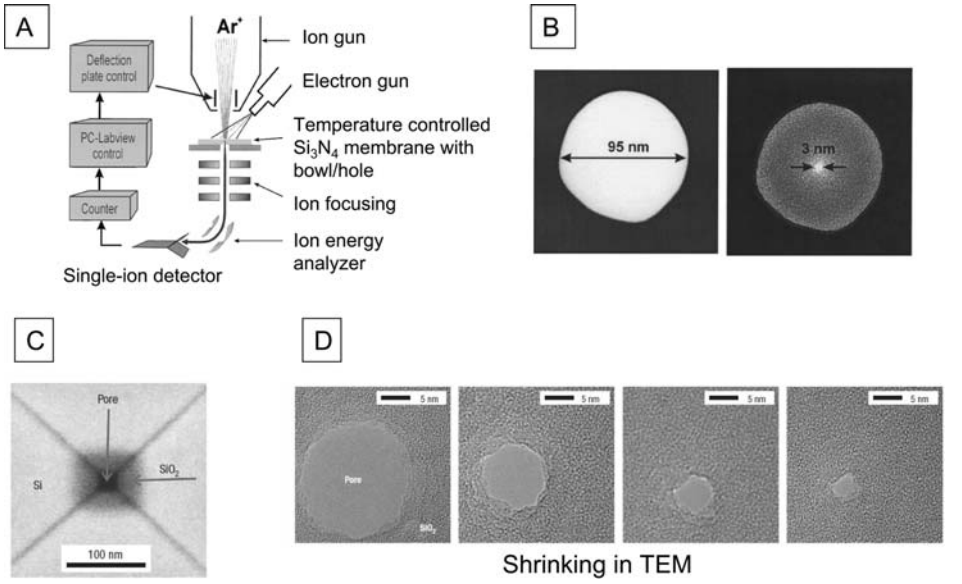


Figure 11.4. Drilling single nanopores in silicon nitride and silicon oxide. A single nanopore in silicon nitride was prepared by a focused ion beam **(A)** and was subsequently closed **(B)** by a diffuse beam. (Reprinted with permission from Stein et al. [119]. Copyright 2004 American Institute of Physics.) A pore was drilled by a focused electron beam in silicon oxide **(C)** that was closed by an electron beam in transmission electron microscopy (TEM) **(D)**. Scale bars are 5 nm. (Reprinted by permission from Storm et al. [120]. Copyright Macmillan Publishers Ltd.)

nanopores in silicon nitride. The use of an argon beam with 3-keV intensity had two opposing effects on silicon nitride. First, the ion sputter erosion led to the removal of the material and pore opening. Second, matter transport due to surface diffusion, viscous flow, and redeposition caused the closing of the pore. Based on this insight, the authors first drilled a 50- to 100-nm-wide pore using a focused ion beam and subsequently narrowed the pore with a diffuse beam (Figure 11.4A, B). The feedback mechanism allowed a real-time monitoring of the process so that the pore closing could be stopped when the desired diameter was reached. This technique has a reported precision of $\sim$1 nm. The physical mechanism behind the nanostructuring method was also studied in a great detail [116,117], along with the influence of various noble gases on the pore-shrinking process [118]. While this method is primarily used for generating pores in silicon nitride, it can also be used for other materials, such as silica, Al, Cr, poly(methyl methacrylate), and polyimide [119].

Electron Beam as a Nanofabrication Tool

The approach of drilling a large hole and its subsequent shrinking was also performed with electron beams. Similar to ion-beam sculpting, drilling is achieved with a focused beam, and pore narrowing is accomplished with a diffuse electron beam in a transmission electron microscope (TEM) [120] (Figure 11.4C–D). Using the TEM for the fabrication process gives the unique possibility of observing the pore *in situ* during the shrinking process. Small rates of pore closing of 0.3 nm/min can be achieved, which allows for fine tuning of the pore diameter.

Electron-beam sculpting was first applied for the fabrication of pores in silicon dioxide, but other materials also have been structured in this way [121,122].

Direct drilling of pores with diameters as small as 3 nm was also reported [121,122]. Due to the wide availability of TEM, the electron-beam–based technique for nanopore preparation is currently used by several groups. Lo et al. reported a method for fabricating sub-5 nm nanopores by using both ion beam and electron beam shaping [123].

Glass Nanopipettes

In the attempt to create solid-state pores with nanometer openings, glass nanopipettes have also been used with success (Figure 11.3H). Commercially available pipette pullers can prepare glass nanopipettes with diameters as small as several tens of nanometers [124–126]. In this technique, a glass capillary is heated in the middle with a laser and mechanically pulled to the point at which the narrow neck breaks, thereby creating two open-ended nanopipettes with diameters as small as ∼40 nm. These pipettes are usually used in the patch-clamp technique for studying single biological channels and in the scanning ion conducting microscope to image nanostructured surfaces [127,128].

White et al. reported an alternative way of preparing glass nanopipettes [129]. Briefly, the method entails sealing an electrochemically etched Pt wire into one end of a soda-lime glass. The glass capillary is subsequently polished to expose the wire's cross-sectional disk, whose area is determined by measuring the diffusion-limited voltammetric current for the oxidation of ferrocene. The Pt disk is then electrochemically etched and the Pt wire carefully removed. The pipette opening diameter is then again checked by measuring the pipette conductance in 1 M KCl. A pipette opening prepared in this way can have a radius down to sub-100 nm.

Submicrometer Pores with Diameters Larger Than 200 nm

Pores that have diameters greater than 100 nm are of equal importance, for example, in detecting viruses and large antibody–antigen complexes. The Mayer group popularized the preparation of single pores in borosilicate glass coverslips using femtosecond-pulsed layers. This technique does not require masks, etching, or high vacuum. The pores are prepared in glass, which has been found to have superior noise properties, as well as mechanical robustness. In addition, the silanol groups on the glass surface offer the possibility for chemical modification [130,131].

A approach for generating poly(dimethylsiloxane) (PDMS) pores with diameters of several hundreds of nanometers involves well-established lithographic techniques. First, a negative master of a pore and two reservoirs is fabricated. The master is prepared by patterning a 200-nm wide, 200-nm-thick polystyrene line on a silicon substrate, which becomes the negative of the pore, and subsequently patterning an SU-8 photoresist, which is the negative of the reservoirs. In a second step, a replica of the master is obtained by casting and curing PDMS. A 200-nm-wide PDMS channel was used to detect λ-phage DNA molecules [132].

11.3.5. Chemical Engineering of Solid-State Nanopores

There are three major reasons to chemically modify solid-state nanopores: (1) to alter the surface charges on the channel walls, (2) to render the walls hydrophobic, and (3) to

attach recognition sites for a specific molecule suitable for nanopore analytics. We describe methods for chemically modifying solid-state pores in polymer, silicon, and glass and provide examples of how surface modification affects the transport of ions and molecules.

Chemical Modification of Nanopores Obtained by the Track-Etching Technique

As a result of heavy-ion irradiation and chemical etching, polymer nanopores obtained by the track-etching technique typically contain carboxyl groups on their walls. These groups present a very convenient starting point for the chemical modification of the pore walls. One way of changing the surface chemistry involves the formation of peptide bonds between COOH and amino groups using 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) as coupling agent [133] (Figure 11.5A). For example, attachment of diamines to PET pores changed the surface charge from negative to positive [134] and affected the transport properties of the nanopores. Conical nanopores with deprotonated carboxyl groups are cation selective and rectify the cation flow with preferential direction from the narrow opening to the big opening of a conical nanopore [135,136]. After the modification with amines, at neutral and acidic solutions, the pores become positively charged, which resulted in a change to anion selectivity and the inversion of the rectified current–voltage curves (Figure 11.5A). A similar surface modification has been performed with polyimide pores [137]. The ion current rectification of conically shaped nanopores with surface charges was explained by breaking of the symmetry of the electrochemical potential inside a conical nanopore [135,138–140a]. As a consequence, the electrical force acting on translocating ions is a function of the voltage polarity resulting in the rectification effect. The ion current rectification was qualitatively and quantitatively modeled using the Poisson-Nernst-Planck equations [140b–141b].

The conical shape of nanopores also provides the unique possibility of obtaining surface charge patterns, such as a zone with positive surface charges followed by a zone with negative surface charges [134]. Vlassiouk and Siwy developed a modification technique that allows for a targeted modification of the tip of a conical nanopore (Figure 11.5B). The modification protocol assures a high concentration c of a modifying reagent only at the tip of the pore and a rapid decay of this concentration along the pore axis x. Such an asymmetric concentration profile is very easily obtained in a conical nanopore when a modifying agent is placed only on one side of the membrane, for example, at the narrow opening. When the diameter of the narrow opening d is much smaller than the diameter of the big opening D of a conical nanopore, the steady-state solution of the diffusion equation reads

$$c(x) = c_0 \frac{d}{D}\left(\frac{L}{x} - 1\right),\ d/2 \cdot \cot\alpha \le x \le L$$

where α is the angle of a nanopore's opening and $c(x)$ is the concentration of the reagent at position x. Boundary conditions for the diffusion equation were $c(d\,/\,2\ \cot\alpha) = c_0$ and $c(L) = 0$.

To implement the localized chemical modification of the pore's tip region, a mixture of EDC and diamine was therefore placed on one side of the membrane with the small opening. The rest of the pore walls remained chemically unchanged. The resulting pore had the surface charge distribution shown in Figure 11.5B, which was confirmed by very high rectification degrees [i.e., value of the ratio I(+5 V)/I(–5 V)] of greater than several hundreds. A similar surface charge pattern between a positively charged part of a pore and a neutral zone was obtained in a nanofluidic channel [142] using a very quickly occurring reaction between biotin and avidin.

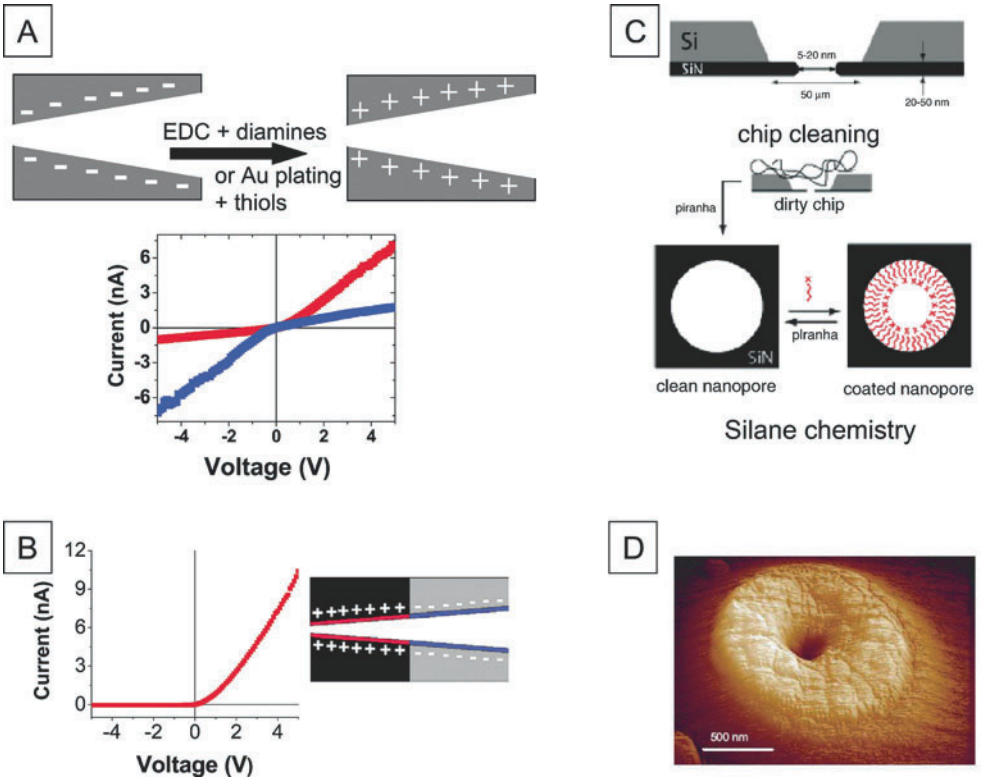


Figure 11.5. Modification of solid-state nanopores. **A.** Polymer pores have carboxyl groups that can be modified with amines using 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) as a coupling agent. This modification changes the surface charge on the pore walls from negative to positive. The change of the surface chemistry led to an inversion of ion current rectification. The displayed current–voltage curves were recorded for a single conical nanopore with negative surface charges (*blue*) and with positive surface charges (*red*) in 0.1 M KCl, pH 5.5. (Reprinted with permission from Vlassiouk and Siwy [134]. Copyright 2007 American Chemical Society.) Another method for changing the surface chemistry of polymer pores involves electroless plating of gold and subsequent chemisorption of thiols. **B.** A conical shape facilitates formation of surface charge patterns, for example, a zone with positive surface charges and a zone with negative surface charges. This surface pattern leads to a diode-like behavior observed in a current–voltage curve which was recorded for a single conical nanopore in 0.1 M KCl, pH 5.5. (Adapted with permission from Vlassiouk and Siwy [134]. Copyright 2007 American Chemical Society.) **C.** Chemical modification of silicon nitride pores with the application of silane chemistry. (Reprinted with permission from Wanunu and Meller [153]. Copyright 2007 American Chemical Society.) **D.** Local modification of silicon nitride pores by deposition of silica. (Reprinted with permission from Nilsson et al. [154]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.)

Another approach to the surface modification of polymer nanopores involves electroless plating of gold and subsequent chemisorption of thiols on the gold surface [143–146]. The electroless deposition process starts by immersing a membrane in a solution of $SnCl_2$, which results in Sn^{2+} being complexed by the surface COOH groups. During the next step, the membrane is immersed in a silver solution, which leads to reduction of silver on the pore walls. The third and final step involves immersing the membrane into a dilute gold solution in the presence of a reducing agent such as formaldehyde. The amount of gold that is reduced and added on to the pore walls is a linear function of time. The gold deposition has to be

performed at low temperatures (typically ∼4°C) to ensure a homogeneous deposition of gold and to avoid clogging of the pores [147]. Electroless deposition of gold offers another possibility of tuning the size of nanopores and achieving very small diameters down to 1 nm. This is important because the track-etching method does not allow for the preparation of pores smaller than 2 nm.

Gold nanotubes can also be very easily modified via thiol chemistry. Carboxyl- and amino-terminated thiols were chemisorbed on single gold nanotubes to render the surface negatively or positively charged, respectively [148]. The change of the surface charge was also monitored by observing changes in the current–voltage curve. Gold surfaces furthermore provide a very convenient way for attaching various biomolecules to the pore walls. This was first demonstrated using membranes containing many pores, and recently shown for single pores as well. Membranes with multiple gold tubes were modified with single-stranded DNA (ssDNA) [149], peptides [150], and proteins [151]. These systems achieved selective transport of molecules that specifically bind to the recognition agent on the pore walls. The selectivity was explained by the facilitated transport and the "relay" mechanism of the molecules hopping between neighboring binding sites.

Chemical Modification of Silicon-Based Nanopores

The first attempt to modify the surface of silicon nitride pores was performed by the deposition of an atomic layer composed of aluminum oxide [152]. The deposition of a ∼3-nm-thick Al_2O_3 film was found to improve the transport properties of the pores. In particular, the ion current rectification exhibited by the unmodified silicon nitride pores disappeared after the aluminum oxide had been deposited. The Al_2O_3 layer also reduced the $1/f$ noise level. In contrast to the cation-selective native pores, the coated pores were not ion selective, which was found to improve their ability to detect single DNA molecules.

Silicon-based nanopores can also be modified by chemical methods, as shown by Wanunu et al. [153] (Figure 11.5C). Individual silicon nitride nanopores were modified with organosilanes, rendering the surface positively or negatively charged. Two protocols for the chemical modification were applied. The first is based on simply immersing the single-pore membranes in a solution of a modifying agent. This approach was found to block smaller pores. To address this shortcoming, a protocol for the chemical modification *in situ* was developed. It involved mounting the silicon nitride membrane between two chambers of a conductivity cell. After introducing the modifying agent in an organic electrolyte dissolved in anhydrous solvent, a voltage was applied across the membrane. This voltage-assisted chemical modification was successfully used for pores as small as 5 nm in diameter. The clogging of nanopores was avoided because—most likely—the voltage drove unreacted silanes out of the pore. The surface modification was confirmed by electrochemically studying ion currents at various pH values, that is, at different protonation states of silanes. Successful chemistry was also confirmed by obtaining X-ray photoelectron spectroscopy (XPS) spectra of bare silicon nitride films subjected to the same modification procedures as the films with pores.

Silicon nitride pores can also be locally modified by a controlled deposition of silicon oxide, as shown by Nilsson et al. [154] (Figure 11.5D). These efforts are very important because they realize a more localized and controlled engineering of the pore wall related to what has been demonstrated at atomic precision with the α-hemolysin pore. Single nanopores in silicon nitride were prepared by the focused ion beam method using a setup with combined electron and ion-beam capabilities. The Ga^+ beam was the nanostructuring tool, whereas the

electron beam allowed for the *in situ* characterization and imaging of the pores. As the next step, a mixture of tetraethylorthosilicate (TEOS) gas and water vapor was introduced into a chamber, and the ion beam was focused on the area where the deposition of silicon oxide was desired. This deposition occurred through the decomposition of the precursor gas and subsequent silicon and oxide local condensation. This procedure resulted in the shrinking of the pores from several hundreds of nanometers to around 30 nm in diameter. The silicon oxide layer at the rim of the pore was subsequently modified via silane chemistry. Attachment of single-stranded DNA molecules to the silicon oxide layer was successfully performed, and the modified pores were characterized by recording current–voltage curves.

A similar approach toward the chemical modification of silicon-based pores was reported by Danelon et al. [155]. The authors started with smaller pores of ~50 nm in diameter prepared by the focused ion beam technique. The local decomposition of TEOS was induced by the electron beam rather than the ion beam used by Nilsson et al. [154]. Using the electron beam enabled the authors to avoid ion implantation, which is believed to affect the noise level in the conductivity measurements and influence the pores' ion selectivity. Deposition of silicon oxide in these pores resulted in shrinking the pore diameter to ~10 nm.

Chemical Modification of Glass Nanopipettes

Successful chemical modification was also demonstrated with glass nanopipettes. Umehara et al. changed the surface charge of ~50-nm pipettes with polylysine, which inverted the surface charge from positive to negative [126]. Glass nanopipettes, similar to conical nanopores, rectify the ionic current. Hence the change in the rectification direction was a probe for the successful chemical modification.

11.4. Applications of Nanopores

11.4.1. Sensing and Examining Individual Molecules

Current recordings on single ion channels were used to detect molecules prior to the advent of sensing with nanopores. For example, reversible current blockades caused by ion channel blockers were observed by Neher et al. using the acetylcholine receptor [156]. Single-channel ion blockades induced by metal cations were reported first with inward rectifier potassium channels [157a].

One of the first studies demonstrating that nanopores can detect nonnatural analytes was conducted by Bezrukov et al. on the transient alamethicin pores and PEG [8] (Figure 11.6A). The α-helical alamethicin peptide forms bundles of up to ten strands that give rise to pores with multiple conductance states in single-channel current traces [157b]. When PEG was added, the conductance states were of lower amplitude and featured increased current noise. These changes indicated that the flexible polymer molecules interacted with the pore [8]. The single-channel traces did not show discrete current steps characteristic of individual molecules, most likely because the PEG concentrations used in the experiment was very high.

Current blockades caused by individual polymer molecules were first published by Kasianowicz et al. [9] (Figure 11.6B). The recordings were performed with αHL, which has a unitary conductance rather than multiple current states. The addition of single-stranded DNA or RNA homopolymers to the pore led to short current deflections from the open-channel

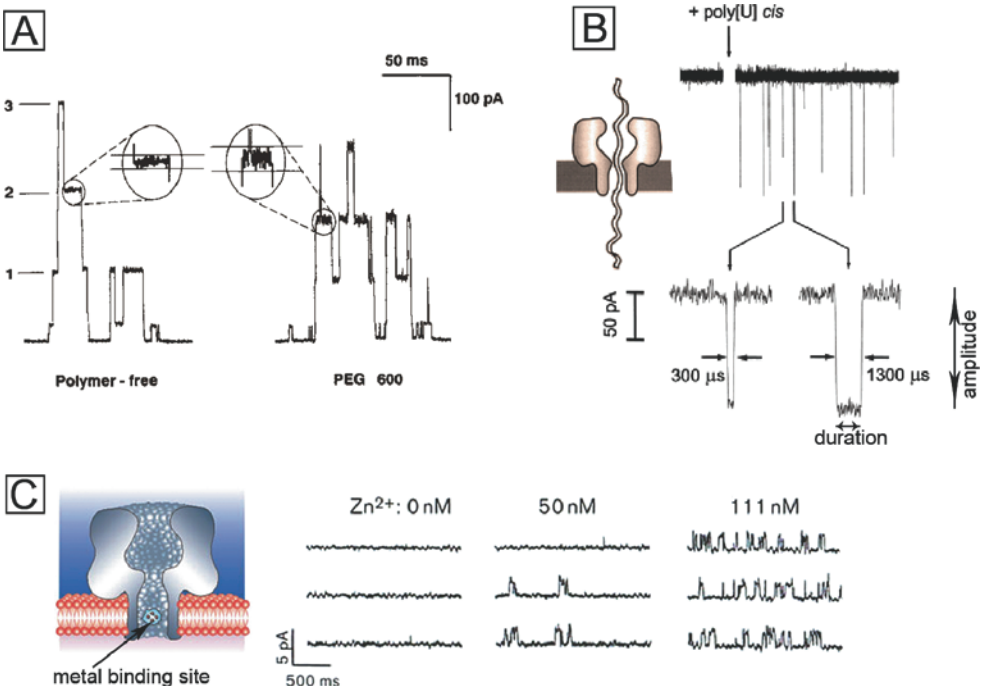


Figure 11.6. Sensing with nanopores. **A.** Ion current of a single alamethicin channel in the absence (*left*) and presence of poly(ethylene glycol) 600 (*right*). The polymer decreases the amplitude of difference conductance levels (1, 2, and 3) and induces additional current noise. (Reprinted with permission from Bezrukov et al. [8]. Copyright Macmillan Publishers Ltd.). **B.** Oligomers of poly[U] cause transient blockades in the ionic current of a single α-hemolysin pore. The trace after addition of the RNA strands with a mean length of 210 bases shows high-amplitude events, which are attributed to the translocation of individual nucleic acid strands as illustrated schematically. The inset shows two typical high-amplitude blockades with lifetimes of 300 and 1,300 ms, respectively. (Reprinted with permission from Kasianowicz et al. [9]. Copyright National Academy of Sciences, U.S.A. Reprinted with permission from Akeson et al. [158]. Copyright Biophysical Society.) **C.** Detection of metal ions with an α-hemolysin pore carrying a genetically engineered binding site composed of four histidines on one of the seven pore subunits. The single-channel current traces show the dependence of the blockade on the concentration of Zn(II). (Reprinted with permission from Bayley and Cremer [51]. Copyright Macmillan Publishers Ltd.)

current. The blockades represent the voltage-driven translocation of individual nucleic acids strands through the pore because the 1.2-nm-wide inner constriction of αHL can only accommodate one individual single-stranded DNA molecule at a time. The blocking events—or short "events"—were characterized by their frequency of occurrence, their duration, and their conductance amplitude (Figure 11.6B). In line with the proposed molecular origin, the frequency of occurrence of events increased with the concentration of the nucleic acids. Furthermore, the blockade duration—also called dwell time or event lifetime—was proportional to the length of the strands [9]. A more detailed examination found that the dependence can be more complex due to secondary structure elements. Finally, the current amplitude was specific for the homopolymer type [158] and the orientation of the nucleic acid strand during the fast translocation through the nonengineered pore [159]. A detailed review of nanopore analytics of nucleic acids can be found in the chapter by Meller within this book.

The binding of individual molecules to a pore with an engineered molecular recognition site was first demonstrated by Braha et al. [10]. Single bivalent metal cations were detected

using an αHL pore carrying several engineered histidines in the β-barrel (Figure 11.6C). The cationic metals bound reversibly to the electron-rich imidazole groups, thereby causing temporary current blockades. The frequency of the events was proportional to the concentration of metal ions in the solution (Figure 11.6C), and different cations gave rise to different current blockade signatures [160].

As in any other single-molecule experiment, the frequency of occurrence of the blockade events and the duration of the events described by Braha et al. [10,160] were Poisson distributed, reflecting that it is not possible to predict when a given single-molecule event occurs and how long it will last. To emphasize this degree of randomness, nanopore analytics has been described as "stochastic sensing" [51]. In stochastic sensing with nanopores, the frequency of occurrence of individual blockade events reveals the concentration of an analyte, whereas the current signature (the mean duration and amplitude of the events) reveals its identity [51]. The term was introduced in the context of analyte detection with engineered protein pores carrying molecular recognition sites [34,51] but has also been applied to describe the sensing with nonengineered nanopores [36a,76,96,161].

Whereas the initial experiments on nanopore analytics were conducted with PEG, DNA, and metal cations, the range of accessible analytes has been significantly widened to include peptides and proteins, biomolecular complexes, organic polymers, and small molecules. Table 11.1 provides a comprehensive list of which analyte has been detected with which pore system. As the details of the current modulation and indeed the pore–analyte interaction can vary considerably among different pores, the following survey distinguishes three sensing modes to enhance clarity. The categories differ with respect to whether the molecules pass through, bind to, or are covalently attached to the pore.

Coulter Counter Method

In the Coulter counter mode, analytes pass the pore without specifically binding to a predetermined site within the channel. Usually, unmodified pores are used for Coulter counter sensing. The approach was initially demonstrated with nucleic acids strands [9] and has continued to be used for DNA and RNA [12,15,16], as well as for proteins and protein assemblies, peptides, linear organic polymers, and complexes between nucleic acids and proteins.

Following the essential function of a Coulter device to count and size, nanopores have been used to determine the concentration of proteins and distinguish differently sized proteins. The experiments were conducted using a silicon nitride pore of 55 nm diameter in combination with bovine serum albumin (BSA) protein ($5 \times 4 \times 4$ nm) and fibrinogen ($34 \times 4 \times 3$ nm) [162]. The larger fibrinogen caused bigger current blockades than the smaller BSA protein. In another study, size-dependent current blockades were exploited to distinguish between BSA and the larger immunocomplex with anti-BSA antibody [163]. Detecting analytes via the formation of larger immune complexes is advantageous because the large aggregates produce an analyte-specific signal that can be easily detected. The latter benefit was demonstrated using glass pores of 650 nm and the enterotoxin B analytes [131]. Only specific large immunocomplexes gave rise to blockades, whereas smaller ubiquitous proteins in the sample did not cause current modulations. This helped eliminate the nonspecific background blockades that usually affect the signal-to-noise ratio. Independent of the specific experimental conditions, the pores used in these systems were of bigger dimensions than the analyte. Therefore, molecules preferentially passed the pore while showing little interaction with the channel walls.

Table 11.1. Single nanopore systems used in detecting single molecules

Analytes	Pore	Mode[a]	Comments[b]	Ref.
Proteins, enzymes, and peptides				
Antibody- staphylococcal enterotoxin B complexes	Silicate	I	Enterotoxin complexes but not antibodies or toxins give rise to current blockades	[131]
Antibody–microparticle complexes	PDMS	I	Antibody binding leads to detectable increase in microparticle diameter	[183]
Antibody–BSA complexes	Gold-plated Kapton	I	Discrimination of larger complexes from individual BSA and antibody proteins	[163]
BSA	SiN	I	Detection of single BSA molecules as they pass through a nanopore	[162]
Fibrinogen	SiN	I	Discrimination between larger fibrinogen and smaller BSA protein	[184]
Ricin	Gold-plated PET	II	Binding of ricin analyte leads to permanent and complete pore blockade	[180a]
Streptavidin, antibody-antigen	Glass nanopipette	II	Binding of analyte to walls of functionalised pipettes caused changes of the ion current signal	[180b]
Streptavidin, bovine serum albumin, lysozyme	PET and Kapton	II	Interactions of analytes with the pore walls were detected as a change in current-voltage curves	[180c,180d]
Antibody	αHL	II	Binding to biotin-terminated PEG leads to fewer PEG-induced current blockades	[80]
Antibody	Protective antigen, anthrax	II	Antibody prevents the binding of lethal factor to the PA63 pore, which otherwise would modify pore current	[185]
Streptavidin	αHL	II	Binding to biotinylated DNA	[172]
Protein kinase	αHL	II	Binding to inhibitory peptide fused to pore	[77]
Lectin	αHL	II	Monovalent and multivalent binding events to disaccharide-derivatized pore	[36a]
Enzyme-ligand interaction	Alamethicin channels	II	Sulfonamide ligand to carbonic II tethered to alamethicin peptide	[36b]
DNA exonuclease complex	αHL	I	Nanopore force spectroscopy reveals energy barriers affecting dissociation of complex between exonuclease I and single-stranded DNA	[37]
RNA translocase complex	αHL	I	Adenosine triphosphate–dependent binding of translocase to RNA	[40]

Table 11.1. (continued)

Analytes	Pore	Mode[a]	Comments[b]	Ref.
DNA polymerase activity	αHL	I	Discrimination between unbound DNA, binary DNA/polymerase complexes, and ternary DNA/polymerase/deoxynucleotide triphosphate complexes	[38,39]
Peptide	αHL	I	Peptide translocation influenced by electrostatic interactions and steric clash between charged peptides and pore	[43–44b]
Peptide, temperature-dependent conformational changes	αHL	III	Peptide sequences inserted into protein pore lumen collapse above the transition temperature, leading to fewer current blockades	[186]
Unfolding of peptides and proteins	αHL	I	Unfolded but not native maltose-binding protein can thread into the pore; β-hairpin peptide unfold during translocation	[170,171]
Organic polymers				
PEG, repeat units	αHL	I	Detection of the size distribution in a polydisperse PEG sample	[166]
Hydrodynamic volume and conformational flexibility	αHL	I	Detection of the size distribution in a polydisperse PEG sample	[31–33]
Small molecules and single-molecule chemistry				
Antibiotics	OmpF porin	I	Interaction of antibiotics with constriction zone within channel	[46]
Trinitrotoluene	αHL	II	Interaction with engineered aromatic amino acid side chains in pore lumen	[76]
Adamantane derivatives, tricyclic antidepressant drugs	αHL	II	Recognition by ring-shaped molecular adapter β-cyclodextrin	[81]
Tetrahydrofuran	αHL	II	Binding of analyte to ring-shaped cucurbituril adapter inside the pore	[178a]
Nucleotides	αHL	II	Sensing of ribonucleoside and deoxyribonucleoside 5′-monophosphates via binding to adapter	[178b,178c]
Inositol 1,4,5-trisphosphate (IP3)	αHL	II	IP3 and cAMP bind to ring of 14 arginines	[177]
Hoechst 33258 DNA-binding drug	Kapton	II	Binding of positively charged drug to negatively charged channel walls, detection via rectification but not via pore blockade	[181]
Porphyrin	Kapton	I	Detection of single molecules as they pass through a single nanopore as a function of applied voltage and porphyrin concentration	[106]
Photoisomerization of azo-benzene	αHL	III	*Cis–trans* isomerization induced by light	[42]
Polymerization	αHL	III	Stepwise growth of a single polymer chain using thiol-based linkages	[187]

Table 11.1. (continued)

Analytes	Pore	Mode[a]	Comments[b]	Ref.
Photo-induced fragmentation	αHL	III	Kinetics of a three-step reaction	[41]
Organoarsenic compounds	αHL	III	Reversible formation of covalent bonds between organoarsenic compounds and pore cysteines	[182]
Chemical modification on amino acids	Gramicidin	III	Detecting the presence and absence of Boc protecting group on the peptide terminus	[71,188a, 189b]
Divalent metal ions	αHL	II	Binding to histidine patches within the pore lumen	[10,160]

αHL, α-hemolysin; BSA, bovine serum albumin; cAMP, cyclic adenosine monophosphate; PDMS, poly(dimethylsiloxane); PEG poly(ethylene glycol); PET, poly(ethylene terephthalate).

[a]I, Coulter counter; II, molecular recognition site; III, covalently attached analytes.

[b]Unless stated otherwise, transient interactions between the pore and individual analytes gave rise to separate reversible current deflections from the open-channel current level whose frequency of occurrence increased with the concentration of analyte.

Analyte–pore contacts can become more predominant when analyte molecules are of similar size to the pore diameter. This was observed for the translocation of cationic α-helical peptides through the αHL protein pore. The peptides with a cross-sectional diameter of 1.4 nm permeated from one to the other side of the pore but were slowed down due to the electrostatic and steric interactions at the narrow β-barrel [43]. The resulting energetic barrier for the translocation of the positively charged peptide could be lowered, however, by placing negatively charged amino acids residues at the entry and exit of the β-barrel [164]. These findings are of relevance to the biologically important process of peptide translocation and contribute to a better mechanistic understanding of how electrostatic interactions alter the underlying kinetics of translocation.

The organic polymer PEG represents another class of analytes whose size matches the dimension of the αHL lumen. Several studies systematically examined the permeation of PEG chains into the protein pore by detecting the polymer-induced current blockade levels [31–33,165] or increased current noise [165]. The permeation was dependent on the size of PEG chain [31,165]. In addition, the residence time of the polymer showed a crossover behavior, first increasing and then decreasing with molecular weight [33]. The data were interpreted by assuming that for polymers that are too large to be accommodated within the pore, the out-of-the-pore part of the molecule pulls on the trapped part, thus acting as an entropic spring [33]. The matching dimensions between analyte and αHL pore lumen were also exploited to detect minor variations in the size distribution of PEG [166] (Figure 11.7A). The polydispersity arises from differences in the number of ethylene glycol units in the linear polymer chains. Similar to previous studies, PEG gave rise to discrete current blockades that stem from individual molecules temporarily residing within the pore. Slightly different current levels not only discriminated between the minor mass variations between the PEG chains, but the frequency of occurrence of the blockades also accurately represented the relative percentages in the mass distribution found in conventional mass spectra. The recordings resolved these minor differences because both analyte and αHL protein pore lumen have matching dimensions, thereby maximizing the effect of slight mass variations on current blockade.

Whereas the traditional Coulter counter concept envisages the pore diameter to be wider than or of similar size to the analyte, it is also possible to sense molecules that are too big to pass the pore. In its simplest embodiment, this approach can be exploited to follow the unfolding of charged bulky polymers. Unfolding occurs because the bulky folded polymer is too big for pore entrance, whereas the free dangling polymer end is electrophoretically pulled into the narrow pore. This method was first demonstrated using DNA hairpins composed of a duplex and a single-stranded extension [167]. The latter threaded into the β-barrel of αHL, whereas the 2-nm-wide duplex could not pass the 1.2-nm-wide inner constriction of the protein pore. Passage occurred after unfolding of the duplex, and the time required for dissociation of the complementary segments correlated with the free energy of the duplex [167] and kinetic rate constants [168,169]. The concept of unfolding induced via electrophoretic threading was also applied to β-hairpins [170] and proteins [171]. In the case of proteins, completely unfolded polypeptides led to short pore blockades, whereas partially folded conformations were characterized by long blockades. This is consistent with the idea that translocation requires the unfolding [171]. Indeed, the proportion of short events increased when a denaturing agent was added. Unlike the larger globular protein, partially, as well as fully, folded β-hairpins could traverse the αHL pore, albeit at greatly retarded speed [170].

In the second embodiment of sensing pore-impenetrable analytes, complexes between a single-stranded nucleic acid molecule and a larger protein can be detected. This approach

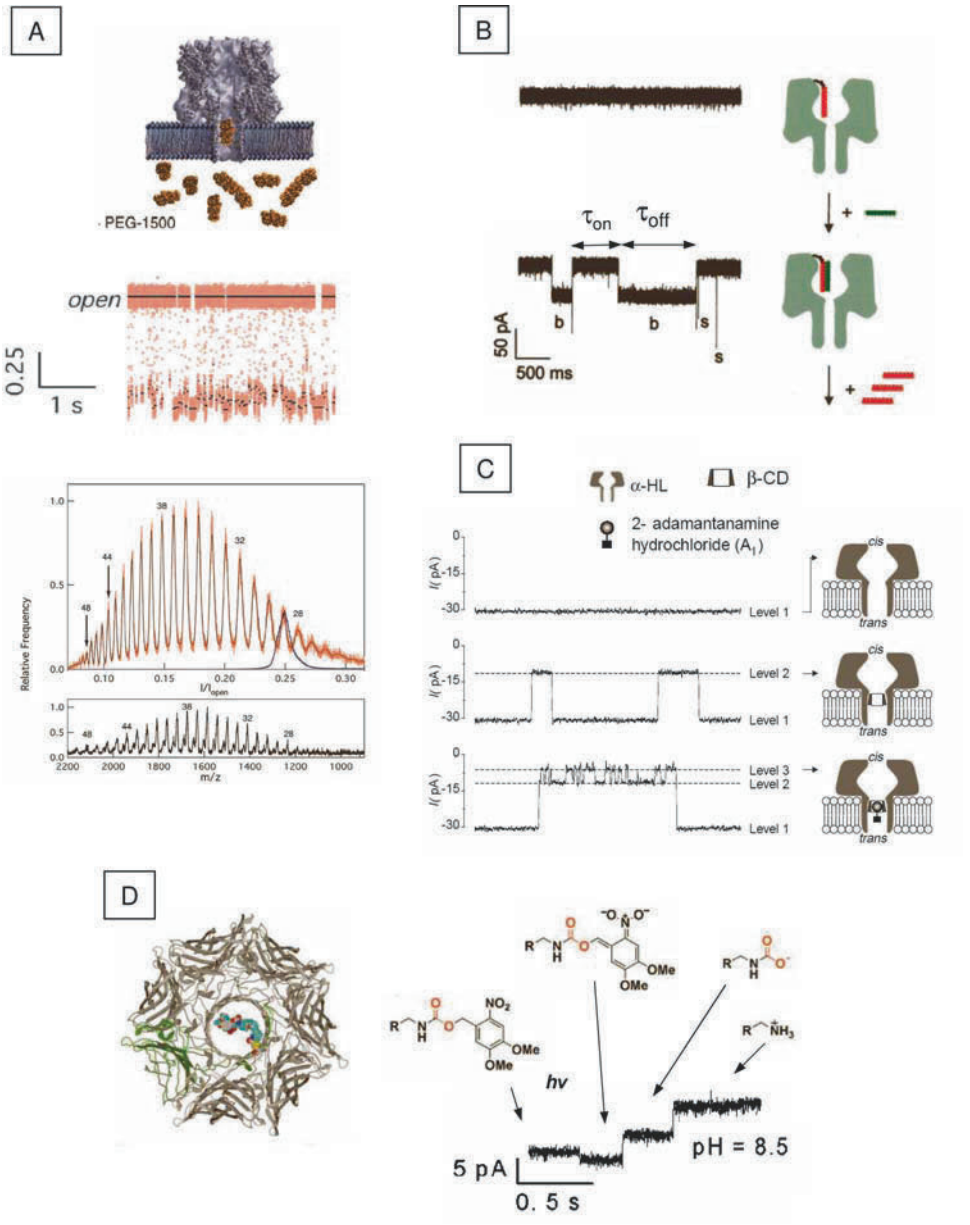


Figure 11.7. Sensing with nanopores. **A.** Single neutral poly(ethylene glycol) (PEG chains) partition into the α-hemolysin (αHL) nanopore (*top*) and cause well-defined reductions in the ionic current that represent the polymer's different molecular masses (*middle*). The relative blockade distribution obtained with a single nanopore (*lower*) is very similar to a distribution in a conventional matrix-assisted laser desorption/ionization time-of-flight mass spectrum for polydisperse PEG of molecular weight 1.5 kD (*bottom*). (Reprinted with permission from Robertson et al. [166]. Copyright National Academy of Sciences, U.S.A.) **B.** An αHL pore modified with a single DNA oligonucleotide responds to individual binding events with oligonucleotides of complementary sequence. Representative current traces before (*top*) and after (*bottom*) the addition of a free oligonucleotide 8 nucleotides in length (*green*) in the *cis* chamber. Negative current deflections (b) represent individual binding events of the free oligonucleotide (*green*) to the tethered oligonucleotide (*red*). The short downward spikes (s) in the trace are translocation events of

was initially presented for DNA–antibody dimers [172]. More recently, it has been exploited to sense nucleic acid complexes with a DNA exonuclease [37], a DNA polymerase [38,39], and an RNA-binding motor protein [40]. The free nucleic acid strand dangling off the complex was shown to enter the narrow pore. Fast strand translocation was blocked, however, due to the bulky protein that lodged on top of the pore. As the transmembrane potential pulled on the negatively charged nucleic acid strand, the protein nucleic acid complex was destabilized until dissociation terminated the current blockade. Detailed measurements of blockade duration and amplitude not only helped to determine the binding strength between DNA and exonuclease [37], but it also revealed the stepwise enzymatic addition of bases by a polymerase [38,39]. These studies are relevant and highlight the suitability of nanopores for conducting single-molecule force spectroscopy [168,173,174], which has previously been performed using optical tweezers [24–29] and atomic force microscopy [20–23]. Compared to these approaches, nanopore force spectroscopy does not require DNA or protein molecules to be immobilized on a bead, solid support, or cantilever.

Sensing Aided by Pore-Tethered Molecular Recognition Sites

In this second mode of nanopore analysis, molecules specifically bind to molecular recognition sites (MRS) engineered to the pore. The molecules bind reversibly to the MRS and block the pore before dissociating and restoring an open-channel current. This sensing mode has been inspired by biological channels that carry ligand-specific receptors to control the channel opening and closing [175]. Unlike natural channels, nanopores can be equipped with a great variety of nonproteinogenic MRS to broaden the pores' sensing spectrum. For example, protein [78] as well as silicon nitride pores [176] have been derivatized with ssDNA receptors capable of hybridizing to complementary DNA molecules. A further difference is that in nanopore analytics the molecules directly cause the current blockade while in biological channels the mode is indirect as the binding of ligands causes conformational changes in the protein's structure that lead to the blockades.

Many analytes can be detected with engineered pores, as summarized in Table 11.1. The pioneering studies were conducted with αHL because its known X-ray structure and rigid scaffold facilitates atom-precision engineering of MRS. As mentioned before, in the first report on a rationally designed biosensor element, amino acid patches were shown to interact with bivalent metal ions [10]. The signatures of the current blockades were dependent on

Figure 11.7. free oligonucleotides that did not bind to the tethered oligonucleotide. (Reprinted with permission from Howorka et al. [35]. Copyright National Academy of Sciences, U.S.A.) **C.** Bilayer recordings showing the interaction of a single αHL pore with the molecular adapter β-cyclodextrin (βCD) and the model analytes 2-adamantanamine. (a) The αHL pore is continuously open (level 1); (b) βCD produces transient partial blockades of the channel (level 2); (c) 2-adamantanamine does not affect the fully open channel (level 1) but produces an additional block of αHL × βCD (level 3). (Reprinted with permission from Gu et al. [81]. Copyright Macmillan Publishers Ltd.) **D.** Single-molecule chemistry photoinitiated within an αHL pore. The pore with an engineered cysteine carries a single 3,4-dimethoxy-6-nitrobenzylcarbamate protecting group covalently attached through an oligo(ethylene glycol) linker. Four levels in the single-channel current trace were observed, which were ascribed to the carbamate protecting group, the two intermediates, and the deprotected amine, as indicated. The transition from the first to the second level was triggered by exposure to near-ultraviolet light. (Reprinted with permission from Luchian et al. [41]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA.)

the type of metal cation, and this helped to distinguish among various metal analytes [160]. The genetic engineering of binding sites has also been used to make pores that respond to other analytes, such as the explosive TNT [76] or the cellular second messenger inositol-1,4, 5-trisphosphate [177].

Pores carrying covalently rather than genetically engineered attached molecular recognition sites have also been exploited for the sensing of analytes. For the sensing of single-stranded DNA, the αHL pore with a tethered DNA oligonucleotide of complementary sequence was used (Figure 11.7B) [35]. The free DNA strand reversibly bound to the tethered recognition site, leading to current blockades with durations of hundreds of milliseconds (Figure 11.7B). DNA sequences that form a perfectly matched duplex blocked the pore longer than duplexes with a single point mismatch [78], and this observation was exploited to detect a resistance-conferring mutation in an HIV-derived DNA sequence [78].

The work on DNA pores also proved useful for examining the kinetics and thermodynamics of the biomolecular recognition reaction at the single-molecule level. The kinetic rate constant for duplex dissociation, k_{off}, and the association, k_{on}, can be derived from the two characteristic temporal parameters of the binding events, τ_{off}, the event dwell times, and τ_{on}, the interevent interval (Figure 11.7B). k_{off} and k_{on} are related to τ_{off} and τ_{on} via

$$k_{off} = 1/\tau_{off}$$
$$k_{on} = 1/c \times \tau_{on}$$

where c is the concentration of the analyte. The derived kinetic constants for DNA duplex formation at the single-molecule level were found to be consistent with established literature values for macroscopic duplex formation [35]. In addition, the nanopore recordings revealed details of the kinetics difficult to obtain by conventional methods that measure ensemble properties. By investigating the temperature dependence of DNA duplex formation at the single-molecule level, the standard enthalpy and entropy of the interaction could be obtained. Other studies with engineered nanopores have also provided detailed kinetic and/or thermodynamic data for the biomolecular recognition between small ligands and lectins, antibodies, or regulatory enzyme subunits [36a,77,80]. These experiments highlight that engineered pores with recognition sites are useful tools for studying biophysical aspects of biomolecular interactions.

Binding sites for other molecules can be created by noncovalent adapters. Adapters are ring-shaped molecules that lodge inside a pore while simultaneously mediating the binding of smaller analytes [81]. Figure 11.7C illustrates that the adapter β-cyclodextrin (βCD) resides temporarily inside the transmembrane β-barrel of αHL. This reversible interaction gives rise to transient current blockades (Figure 11.7C, level 2). The organic ligand 2-adamantanamine can bind to the adapter, thereby decorating the longer βCD closures with shorter blockades (Figure 11.7C, level 3). Additional ligands such as 1-adamantanecarboxylic acid and the drugs promethazine and imipramine were also shown to interact with the βCD adapter and yield analyte-characteristic blockade signatures [81]. Reflecting the large number of other host–guest pairs, the sensing repertoire can be broadened to include organic solvents [178a]. The adapter approach has been very successfully applied to sense nucleoside and nucleotides. In work by Astier et al. [178b] and Clarke et al. [178C]. In the study, the adapter heptakis-(6-deoxy-6-amino)-β-cyclodextrin was non-covalently or covalently lodged into the αHL pore, respectively. Reversible binding of 2'-deoxyribonucleoside 5'-monophosphates to the adaptor gave rise to distinct current levels for all each base. This pioneering work is relevant for DNA sequencing. In the proposed exonuclease approach, a DNA strand is sequentially digested

by a pore-tethered processive exonuclease. The released nucleotides are then captured and sensed with the nanopore [16C]. Additional development is required in order to tether an enzymatically active enzyme to the pore, and to capture each released nucleotide into the pore for sensing with the adapter [16C].

The detection of analytes with nanopores does not necessarily lead to or require discrete and reversible current blockades. Depending on the nature and the timescale of the analyte pore interaction, other types of current modulation are also possible. For example, binding and dissociation events whose kinetics are faster than the temporal resolution of the current amplifier can give rise to an increased current noise. This has been mentioned for the interaction of PEG with alamethicin [8] but was observed for αHL pores too [165]. A pore with genetically engineered metal-ion-binding sites was also observed to give rise to increased noise on binding of bivalent metal ions [179].

Rather than causing reversible current blockades, the binding of the analytes to the MRS-engineered pore can also lead to a complete blockade. This has been demonstrated for a gold pore with a diameter of 4 nm. The pore walls were derivatized with tens of ant-iricin antibodies. Addition of ricin completely abolished the pore current [180a]. Given that the ricin toxin has a molecular weight of 60 kDa, it is possible that multiple molecules bound to the antibody-coated pore to completely obstruct the pore lumen. This approach provides a quick and robust way for the detection that is not based on the sensing of individual molecules. The study on ricin was the first to report a single solid-state nanopore featuring an engineered and fully functional MRS.

Another strategy for detecting analyte relies on the subtler change of pore rectification that describes the time-independent changes of pore current as a function of voltage. This sensing approach has been illustrated using positively charged drugs and a Kapton pore with negatively charged channel walls [181]. The drug was too small to cause detectable transient current modulations. The electrostatic binding of the drug to the channel wall altered, however, the current–voltage curve of the pore [181]. Table 1 provides a summary of the different sensing strategies.

Sensing of Covalently Attached Analytes

The third category of nanopore sensing is different from the previous modes, in that single molecules are covalently attached to the channel wall to follow the molecules' chemical changes. The covalent approach is ideally suited for small organic molecules that fit into the pore lumen. For example, the αHL protein pore was covalently modified with a single photolabile carbamate group (Figure 11.7D) [41]. Upon illumination with ultraviolet light the protecting group fragments, and the pore recordings revealed the intermediates of the decomposition, as shown in Figure 11.7D. The recordings were able to detect the various intermediates because the current depends on the size, shape, and polarity of the moieties tethered within the pore. Studying single-molecule chemistry with label-free nanopore recordings is advantageous because short-lived intermediates and their kinetics of formation and disappearance can be detected. This would be difficult to attain with ensemble methods or other single-molecule techniques that require labeling.

In another study, the light-induced isomerization of azobenzene was investigated at the single-molecule level [42]. Azobenzene carries a phenyl-substituted nitrogen–nitrogen double bond that can undergo a light-induced transition between the *cis* and *trans* isomers. When an azobenzene derivative was attached to the pore wall, the two different geometric isomers

could be distinguished based on their different current levels. Remarkably, continuous interconversion between the two isomeric states was observed at the wavelength of 330 nm but not in the dark state. The recordings constitute the first observation of the reversible photoisomerization of individual azobenzene molecules in an aqueous environment. The approach to follow chemical transitions by nanopore recordings was also applied to other types of chemical transformation, such as a covalent bond formation [182], and studying chemical modification at the single-molecule level can be extended to other pore systems, such as gramicidin [71].

Most cited studies on nanopore analytics fall within one of the three different sensing modes. The assignment is not always clear-cut, however, and there are ambiguous cases that could be assigned to two different modes. For instance, linear charged peptides fall between the Coulter counter and the MRS modes because they translocate through the pore but also exhibit electrostatic interactions to charged sites within the engineered channel. Similarly, the sensing of viruses via immunocomplexes is a Coulter counter approach but also incorporates non–pore-tethered MRSs. The three sensing modes—and in particular the first two categories—therefore constitute explicit formulations in a more continuous spectrum of possible nanopore detection approaches.

Controlling the Movement and Position of Molecules within Nanopores

Molecules permeate into nanopores due to electrophoresis [9], electro-osmosis [189], a chemical gradient, or a combination of these factors. Controlling the rate at which molecules enter the pore is of great practical value. By tuning the frequency of pore entrance, the number of events can be increased to improve the statistics on their characteristics. Modifying the speed of translocation is also of interest because molecules that would otherwise pass the pore very quickly can be slowed down to improve the signal-to-noise ratio for the current blockades. Furthermore, varying the transmembrane potential is an important tool for probing the binding strength within biomolecular complexes in single-molecule force spectroscopy experiments, as described in section Coulter Counter Method [37–39].

It can also be beneficial to trap a single molecule within a pore. Keeping a molecule stationary can improve the signal-to-noise ratio of the current blockade [190] or facilitate the examination of analyte association and dissociation kinetics [191]. Several approaches have been developed to capture a single molecule. One strategy employs physical barriers within a pore. The barriers were composed of noncovalent adapters that served as lids for a nanocavity within a protein pore [191]. Single organic, charged molecules were electrophoretically driven into the cavity, and complex kinetics for the binding to two different adapters could be examined [191]. A physical barrier can also assume the form of a duplex in DNA hairpins with a single-stranded extension. Because the duplexes cannot easily pass through the narrow inner constriction of αHL, the extension was kept threaded inside the β-barrel [167,190]. This approach helped to detect different bases in a single position of the stationary ssDNA [190].

Use of electrostatic barriers within a pore is another strategy to maintain charged molecules inside the pore. This was demonstrated with a negatively charged ring in the αHL pore that interacted with positively charged polypeptide analytes [192]. Theoretical and experimental work indicates further scope for the trapping of analytes using electro-osmotic flow in oppositely charged nanopores [193] and the electrokinetic capturing of larger objects [194].

Catching a single molecule and moving it back and forth through the pore repeated times is another strategy. By performing multiple measurements on a single molecule, the accuracy and precision of the analytical results may be improved. This concept was exemplified with a single DNA strand covalently attached to a short PEG chain [195]. After the polymeric chain was threaded into a pore, bulky terminal groups were latched onto the strand termini outside the pore [195]. The strand could be moved up and down the pore by using sequential switching of the transmembrane potential. Ultimately, this methodology may lead to a more reliable analysis of DNA sequences [196].

A completely different approach to controlling the position of molecules relies on the use of optical tweezers (Figure 11.8). In a study by Keyser et al., a single DNA strand was threaded with one end into a solid-state pore [197a]. The other DNA terminus was attached to a bead. By controlling the position of the bead with an optical tweezer, the direction and speed of DNA threading through the nanopore was varied. This strategy led to detailed measurements on the frictional forces of strand movement [197a]. The technique of optical tweezers also allowed for repeated characterisation of the same piece of DNA by "flossing" it through the pore [197b]. Capturing of the same DNA strand after its translocation through the pore was also achieved by a quick reversal of the transmembrane potential, which forced the molecule back into the pore [197c]. Note that the ultimate method of holding a single molecule stationary within a pore is covalent attachment (see section Sensing of Covalently Attached Analytes).

Theoretical Modeling

Theoretical modeling is a prerequisite to fully understand the behavior of molecules in a nanopore and also to build improved single-molecule sensors. The modeling of interactions within a subfemtoliter volume of nanopores has been performed using different approaches such as Monte Carlo, Brownian dynamics, and molecular dynamics simulations. For example, theoretical studies on translocating polymers have determined free energy barriers and

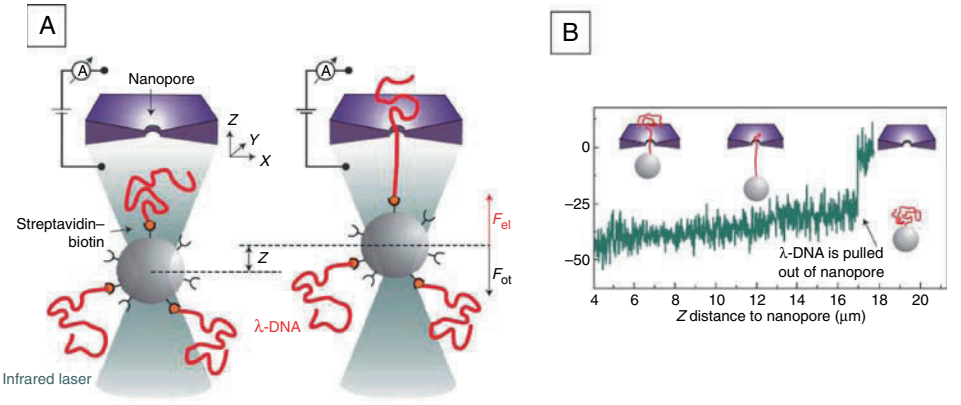


Figure 11.8. Combination of optical tweezers and electrochemical measurements for direct force measurements on DNA in a nanopore. **A.** A focused laser beam is used to trap a DNA-coated bead near a solid-state nanopare. An electric field applied across the nanopore drives the DNA molecule into the nanopore. **B.** The DNA is pulled out from the nanopore at a speed of 30 nm/sec. (Reprinted with permission from Keyser et al. [197a]. Copyright Macmillan Publishers Ltd.)

mean translocation time as a function of polymer length and channel dimensions, as well as the chemistry of the translocating molecules and the pore [198–200]. Muthukumar predicted that when the pore length increases, the translocation process becomes primarily dependent on the polymer–pore interactions [201]. Storm et al. found that for long DNA chains passing through a pore, the hydrodynamic force acting on the DNA part outside the pore counteracts the electric driving force and causes a power-law dependence of the translocation time with the polymer length L, following $L^{1.22}$ [202]. Atomistic molecular dynamics simulation of the α-hemolysin pore and DNA provided a deeper understanding of differences of DNA translocation with regard to the dependence of the 3′ and the 5′ orientation of the strand [159]. Computer simulation was also applied to study the translocation of a polypeptide through a nanopore [203]. These studies represent a selection of a much bigger body of work on the theoretical modeling, which is not completely covered here.

11.4.2. Separation and Molecular Filtration

Based on their nanoscale dimension, nanopores can be used for separation and filtration purposes. In the simplest understanding of the molecular sieve systems, molecules larger than the pore diameter are restricted, whereas smaller molecules pass (Figure 11.9A). On a single-pore level, this concept has been realized to detect DNA molecules of different size. Single nanopores in polycarbonate membrane were used to sense linearized single-stranded DNA molecules from a phage DNA in the presence of double-stranded DNA (dsDNA) plasmids [161]. The linear DNA passed through the pores, giving rise to voltage-dependent current blockades, whereas the plasmid did not translocate. The molecular sieve concept was also implemented on a much smaller size scale using ssDNA and dsDNA oligonucleotide molecules. Heng et al. prepared single pores in silicon nitride that were wide enough to allow ssDNA to pass through but were too narrow for dsDNA [204]. A similar size discrimination had previously only been reported for the α-hemolysin pore [167,173,174].

Another study showed that chemical modification of the α-hemolysin pore can fine-tune its molecular sieve properties. Martin et al. [205] covalently attached a sulfhydryl-reactive polyamido amine (PAMAM) dendrimer to cysteine residues of the protein pore. The PAMAM dendrimer blocked the passage of single-stranded RNA strands while permitting the

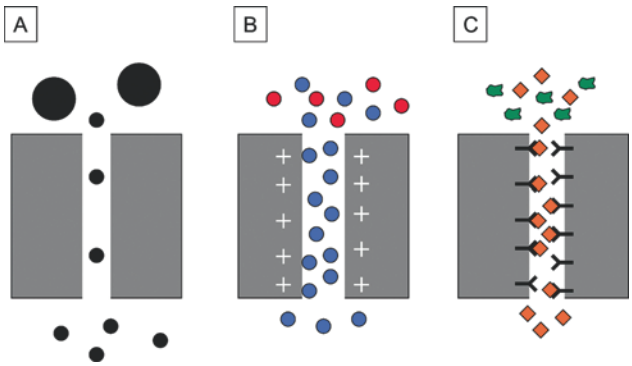


Figure 11.9. Three majors mechanisms used for separations with nanopores: **(A)** molecular sieving, **(B)** separation based on electrostatics, and **(C)** recognition sites and facilitated transport.

permeation of electrolyte ions. This molecular filter can be potentially applied in nanopore-based biosensors to selectively allow the passage of small analyte molecules while blocking out large components within the sample matrix that would lead to the jamming of the pore.

Exploiting the full potential of molecular sieves for separation processes requires big fluxes rather than transport through single channels. Table 11.2 summarizes a selection systems with multiple pores that have been used for the separation of various types of molecules. One of the systems used for separating molecules by size was based on multipore gold-plated polycarbonate membranes with a pore diameter of less than 1 nm. The porous membrane was tested with pyridine (molecular weight 79) and quinine (molecular weight 324) [147]. Only the smaller pyridine passed the pores, whereas the bigger quinine was rejected. Many other reports have achieved separation by exploiting steric factors in nano- and microstructures. Rather than relying on size exclusion, the designed structures exploited processes such as entropic trapping, Brownian ratchet action, hydrodynamic sorting, and continuous-flow molecular sieving [206–213]. A detailed discussion of these approaches can be found in a dedicated review article [214].

Molecules can also be separated based on electrostatic interactions (Figure 11.9B). Due to the attraction of opposite charges, negatively charged nanopores are expected to be filled predominantly with positively charged species, whereas anions are rejected. The concept of selectivity via electrostatic interactions has been applied for the separation of proteins based on their different isolectric points [215]. The study was performed with gold-plated polycarbonate membranes that carried a self-assembled monolayer of alkanethiols with terminal carboxyl groups. The pores with an hydraulic radius of $\sim$8.7 nm were applied to separate bovine serum albumin and bovine hemoglobin proteins that have nearly identical molecular weights but different isoelectric points of 4.7 and 7.0, respectively. The diffusion of the proteins through the porous membrane was dependent on the pH value of the solution. The fluxes reached their maximum at pH values close to the specific isoelectric points of the proteins, most likely because there are fewer electrostatic interactions to the channel wall at electroneutrality than with proteins of net charge. The electrostatic interactions were found to be more pronounced at low ionic strengths where the surface charges are less screened. For example, at 0.01 M KCl, the protein fluxes were almost an order of magnitude smaller than at pH values above and below the pI, whereas the difference was only threefold at 0.1 M KCl [216]. This suggests that protein diffusion through pores can be switched on or off by varying the pH.

The importance of electrostatic interactions for separating proteins in chromatographic and ultrafiltration processes has been discussed by Zydney and coworkers [217–219]. A very interesting concept for transporting amino acids against their concentration gradient has been recently put forward and experimentally proven [220,221]. A pH gradient was imposed across a gold nanotube membrane carrying carboxyl-terminated alkanethiols. One side of the membrane with an acidic pH of 2.4 contained tyrosine in its cationic form, whereas the other side of the membrane with a pH value of 11.5 contained negatively charged tyrosine molecules. Because the carboxyl modified pore walls are protonated at pH 2.4 and deprotonated at pH 11.5, the porous system had an electrochemical potential with broken symmetry. The uphill transport of tyrosine occurred from the acidic reservoir containing a lower concentration of tyrosine to the reservoir with basic and higher tyrosine concentration. The transport occurred because it was energetically more favorable for tyrosine in its cationic form to pass through a negatively charged pore and against the concentration gradient than for the anionic tyrosine to flow down its concentration gradient through a negatively charged nanopore. This coupling of

Table 11.2. Multiple-pore membranes used in separations (selected publications)

Analytes	Membrane	Mechanism for Separation	Ref.
DNA	Microlithographic arrays	Electrophoresis	[224]
DNA	Gold-plated polycarbonate membrane	Facilitated transport	[149]
Proteins	Ultrathin silicon membranes	Steric, electrostatic	[225]
Proteins	Polycarbonate modified with polymers by filtration	Steric, electrostatic	[222]
Proteins	Gold-plated polycarbonate membrane	Electrostatic	[151,215,216,226,227]
Proteins	Pyrex nanoslits	Electrostatic	[228]
Amino acids	Gold-plated polycarbonate	Electrostatic	[221]
Amino acids, chiral separations	Gold-plated polycarbonate	Enzymatic molecular recognition	[229]
Amino acids, dyes	Polycarbonate membranes integrated with microfluidics	Electrostatic	[230,231]
Urea, glucose, sucrose, raffinose, 2-dextrin, ribonuclease	Mica track-etched membranes	Steric	[232]
Drugs	Polycarbonate modified with polymers by filtration, gold-plated polycarbonate, chemically modified alumina membranes	Steric, hydrophobic, and electrostatic	[222,233]
Enantiomeric drugs	Silica nanotubes synthesized in alumina membranes	Molecular recognition	[223, 234]
Small molecules, e.g., pyridine, methyl viologen chloride, quinine, organic ions	Gold-plated polycarbonate	Steric, electrostatic	[146,147,234–238]
Small molecules	Gold-plated polycarbonate	Hydrophobic	[145,223]
Fluxes of water/ethanol	Silicon membranes	Steric	[239]
Glucose, insulin	Silicon membranes	Steric	[240]
Ions, electrostatic separation for inorganic ions	Single and multipore systems: polymer membranes, gold-plated polycarbonate membranes, silicon nanochannels, Pyrex nanoslits, silica channels	Electrostatic	[135,180,205,220,228,241–245b]

the tyrosine transport to a pH gradient is reminiscent of coupling between ionic and molecular fluxes that occurs in biological systems.

Steric and electrostatic factors can interplay in a synergistic way to enhance selectivity within the transport process. This was demonstrated by Savariar et al. by using a polycarbonate membrane with 50-nm pores [222]. The pore diameter and the surface charges were controlled by coating the pore walls using the layer-by-layer procedure, starting from adsorption of Sn^{2+} ions. Alternating treatment with polymers carrying negatively charged or electron-rich functional groups resulted in a layered structure held together by metal chelate bridges and electrostatic forces. The coated pores were applied to electrostatically separate positively charged rhodamine-6G and negatively charged calcein. Separation between the fluorescence dyes was more efficient when performed in a solution of low ionic strength and with pores of small diameter. When the layer-by-layer coating procedure was adapted to include polymers containing a hydrophobic part, separation based on hydrophobic interactions was achieved. Separation by hydrophobic interactions was also demonstrated by Jirage et al. [145]. This study used modified gold tubes with alkanethiols with increasing chain length and hydrophobic properties. Membranes that were modified with $HS\text{-}C_{16}H_{33}$ led to the preferential transport of very hydrophobic molecules and their separation from hydrophilic species.

Mitchell et al. and Lee et al. introduced the idea of separation by molecular recognition [150,223] (Figure 11.9C). Two enantiomers of the drug 4-[3-(4-fluorophenyl)-2-hydroxy-1-[1,2,4]triazol-1-ylpropyl]–benzonitrile were separated using the *R*- or *S*-enantiomer–specific Fab antibody fragments immobilized on the pore walls. In another example of separation by molecular recognition, Kohli et al. prepared a multipore membrane with hairpin DNA molecules immobilized on the pore walls. This system selectively transported DNA molecules that were complementary to the surface hairpin molecules and was sensitive to single-base mismatches. The selective transport of complementary DNA was explained by the process of facilitated diffusion [149].

Finally, nanoporous membranes and single pores have been engineered to attain ion selectivity. The ability to distinguish between simple inorganic ions such as Cl^- and K^+ is of great scientific interest and has fostered attempts to emulate the ion-selective properties of biological channels in cell membranes. Artificial systems that allow one to switch ionic selectivity in response to external stimuli are especially attractive. For example, Liu et al. [246] and Nishizawa et al. [241] prepared polymer membrane systems with ionic selectivity regulated by pH and external voltage, respectively. At the single-pore level, Gu et al. showed that the α-hemolysin pore could be made anion or cation selective by using charged noncovalent molecular adapters. The ring-shaped β-cyclodextrins docked into the narrowest part of the pore lumen, thereby maximizing the effect on ion transport [244]. In another study, Merzlyak et al. enhanced and tuned the ionic selectivity of the α-hemolysin pore by covalently attaching positively or negatively charged reagents to engineered cysteine residues [74].

An interesting system that is responsive to the presence of polyvalent cations such as calcium, magnesium, or cobalt(III) was reported by van der Heyden et al. [247a]. Due to intrinsic negative charge of the silica surface, silica nanochannels are selective for polyvalent cations. Cation selectivity was observed at low cation concentrations, but at higher concentrations the effect inverted and anionic-selective silica channels were obtained. Charge inversion was also observed in OmpF pores [247b] and PET conical nanopores [247c]. The inversion of charge selectivity was explained by Shklovskii [248] by assuming a two-dimensional strongly correlated liquid that changes the effective surface charge from negative to positive.

11.4.3. Nanofluidics

The exquisitely designed natural voltage-gated ion-selective channels have served as inspiration to recreate synthetic ionic devices that regulate the passage of charged molecules in response to an external voltage stimulus. Man-made ionic devices would also represent an adventurous attempt to advance the concept of diodes and transistors from semiconductors into the realm of liquids. Being able to control ions and charged molecules in aqueous systems, these devices would help to create ionic circuits for basic science and applied areas, including lab-on-the-chip, nanofluidic, and biosensing, where switching, redirecting, and amplifying ionic fluxes is necessary or of operational benefit. Nanopores are a perfect structural template for generating these ionic systems. Their large surface-to-volume ratio is of advantage. In addition, the ionic transport can be influenced by steric, chemical, or electrostatic interactions with the pore walls.

Rectifying pores in glass nanopipettes and polymer films were the first systems that exhibited diode-like behavior for ions in solution [111,124,126,135]. Rectification indicates that ionic transport through a pore occurs preferentially in one direction. Thorough theoretical studies on the solid-state pores showed that ion current rectification requires the electrochemical potential within the pore to be asymmetric. An asymmetric potential can be realized either by using, for example, a conical pore shape or by a charge pattern within the pore lumen. Indeed, an ionic switch that is reminiscent of a semiconductor PN junction has been theoretically suggested [249] and experimentally realized in polymeric pores and nanofluidic channels [134,142]. The ionic diode consisted of a circular pore region with positive surface charges followed by a zone of negative charges (Figure 11.10A–C). It was found that the application of a positive voltage drove the ions out of the pore, leading to the formation of the zone depleted of charges. This state represents the "off-state" of a nonconducting diode. The voltage of other polarity caused an increase of the ion concentrations inside the pore lumen and a higher current flow. This constitutes the "on-state" of the pore. A very similar surface pattern was obtained in the biological protein pore OmpF. Single-point mutations were used to introduce four positively charged arginines at one pore entrance and four negatively charged glutamic acids on the other entrance [75a]. The mutations converted the nonrectifying wild-type porin into an ionic rectifier. In a separate approach toward ion current rectification, a difference in the surface charges at the pore openings of wild-type OmpF was achieved by using a transmembrane pH gradient [250]. It was mentioned that when only one part of the pore was charged and the other one neutral, a unipolar diode was formed [142,274].

A transistor is another electrical circuit element that, unlike the diode, makes it possible to control current in two directions. In a typical three-electrode setup, a gate element influences electron current between the two other electrodes of the transistor. Usually, a small change of the gate potential leads to substantial changes in the transport properties of the device. This attractive feature allows for very easy tuning of the electron flow in the transistor. Applied to nanopores, a gate can provide control over ionic flux to either enhance or switch it off (Figure 11.10A,D). The concept of ionic-flow transistor was first realized on a micrometer-scale pore by Schasfoort et al. [251]. A single microchannel of 25 μm height and 25 μm width was prepared in silicon and was equipped with two gate elements. While the gate and the channel surface were insulated from the solution by silicon nitride, a change in the gate voltage affected the surface potential of the silicon nitride layer and the structure of the electrical double layer. Applying a voltage parallel to the channel axis caused the ions in the electrical double layer to move, dragging along the fluid through viscous coupling. Depending on the

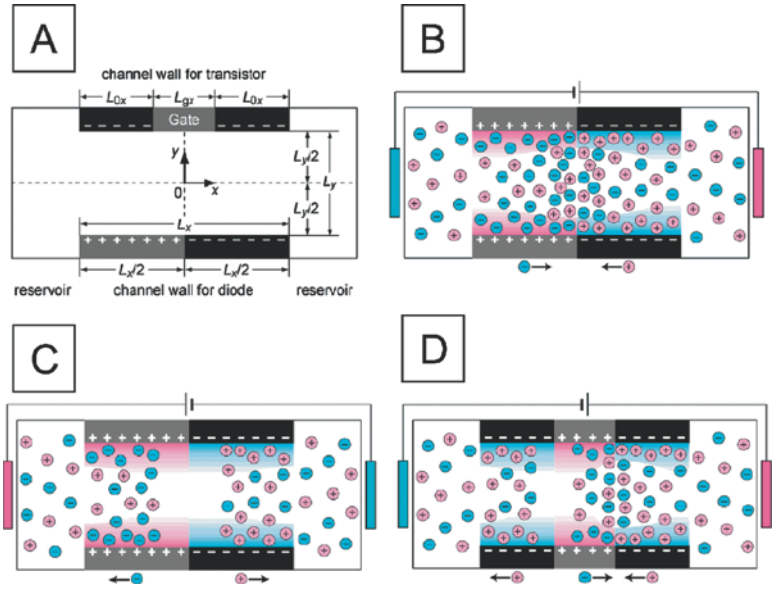


Figure 11.10. A. Scheme of an ionic transistor with a gate. **B.** An open, conducting state of an ionic diode. **C.** A closed (off) state of an ionic diode with the depletion zone created at the junction between positively and negatively charged zones. **D.** Ionic transistor consisting of three charged regions of a pore. (Reprinted with permission from Daiguji et al. [249]. Copyright 2005 American Chemical Society.)

polarity of the voltage applied to the gate, the direction of this electro-osmotic flow could be reversed.

An ionic field-effect transistor based on nanostructures was also prepared (Figure 11.11A) [252]. The transistor was generated using silicon dioxide nanochannels with 30–40 nm in height and silicon dioxide nanotubes with internal diameters of 10–100 nm [253]. The device was applied to control the flux of DNA [254] and proteins [255a]. A potential difference across the channel caused an electrophoretic flux of avidin, while the gate voltage (Figure 11.11B) enhanced or stopped the protein transport. To control the transport of avidin, gate voltages between –1 and +1 V were sufficient to induce a severalfold change in the protein flux.

While successful with proteins, the transistor devices based on silicon nanochannels did not show superior properties with regard to the control of ion fluxes of simple inorganic ions. The most probable reason for this observation is that the pore dimensions are too big compared to the thickness of the electrical double layer. For example, large gate voltages of 20 V had to be applied to change the current by ~40%. Recently a system of conical nanopores with a gold layer at the entrance of the pore was prepared. Gate voltages of ~1V were shown to regulate the ion current in electrolyte concentrations of 0.1 M [255b]. So far, only one bipolar transistor for ions has been developed (Figure 11.10D). The transistor is based on a single double-conical nanopore with a limiting diameter as small as 5 nm [256]. The pores contained a zone with positive surface charges at the two pore openings and a zone with negative charges in the middle. It is a chemical transistor because the charge on the gate can be changed only by chemical methods, namely by altering the ionic concentration and pH.

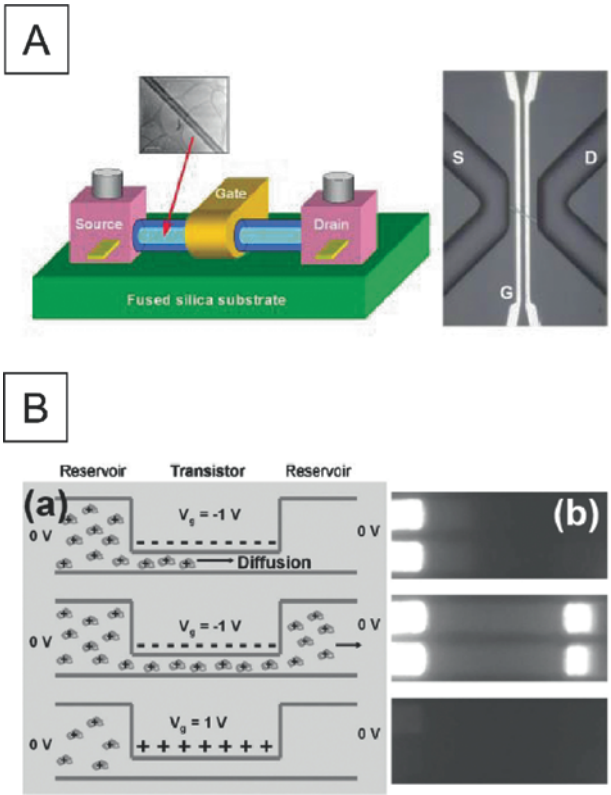


Figure 11.11. Nanofluidic field-effect transistor. **A.** A nanofluidic field-effect transistor with a gate. (Reprinted with permission from Fan et al. [252]. Copyright 2005 American Physical Society.) **B.** Application of the nanofluidic transistor to control the transport of avidin. (Reprinted with permission from Karnik et al. [255a]. Copyright 2006 American Institute of Physics.)

11.5. Outlook

Nanopore analytics will likely benefit from research advances in three areas: new nanopore materials and chemical modifications, integration of nanopores with non–current-based sensing techniques, and the improvement in the current noise of solid-state pores.

Although nanopore analysis has so far been conducted with a variety of biological or synthetic polymeric, inorganic, and metallic pores, other nanopore materials such as individual single-walled carbon nanotubes (SWCNTs) have so far not been exploited. Experimental studies were performed, however, on single multiwalled carbon nanotubes (MWCNTs) with a diameter of ~100 nm [257]. Isolated MWCNTs were used for detecting charged single particles by the Coulter counter method. SWCNTs are expected to be especially useful in nanopore analytics. This expectation is based on the defined diameter of 1 nm and a wide body of theoretical modeling. Indeed, experimental studies on membranes with 10^{10}–10^{11} carbon nanotubes per cm^2 [258,259] along with molecular simulations [260,261] have demonstrated that water and solvated ions can pass through the lumen of the pores. Nanotube

membranes with carboxyl groups at the entrance, however, were found to reject electrolytes, indicating that the permeation properties can be tuned [263]. DNA translocation through SWCNT has also been modeled [263] and implicated with nanopore analysis [264]. There are, however, no experimental accounts on the detection of single molecule with individual SWCNTs.

Single-molecule detection with nanopores has predominantly been carried out using ionic current recordings. There are no limitations on the possible use of other types of readout technology. For example, fluorescence sensing can follow the passage of labeled compounds through nanopores either by detecting the accumulated compounds on one membrane side [265] or by tracking individual molecules as they pass through the pore [266]. In the case of fluorescence-based readout, the pore could also act as a guide that unfolds polymeric analyte strands such as DNA as they enter the pore. This could facilitate the stepwise interrogation of the sequence as the strand passes through the pore. For example, it has been suggested that a narrow nanopore can unzip short fluorescence-tagged oligonucleotides from a longer analyte DNA strands to help reveal sequence information [264].

Nanopores can also be equipped with an *in situ* sensor to read analytes independent of ionic current measurements. In one type of nanopore sensor, built-in electrodes are set up on opposite sides of the channel wall to monitor the transversal tunnel current. Given that the tunneling current would be affected by molecules passing through the pore, it might be possible to detect and analyze individual molecules such as DNA strands [264]. Theoretical studies have shown that each of the four nucleotide carries a unique signature of tunneling current and that this signature is independent of the nearest-neighbor nucleotides [267] and occurs under various pore diameters [268] and field strengths [268,269]. Very recently, a nanopore detector with a gap of 9 nm has been built and tested on double-stranded DNA using the related transverse electrical current [270a]. Nanofabrication of electrodes placed as close as 2 nm has recently been reported [270b,270c]. This integrated concept may be expanded to accommodate other detection schemes, such as fluorescence or electron spin resonance [271], or applied to analytes different than DNA.

Improvements in the quality of recordings can also advance nanopore analytics. Single-molecule detection usually relies on current signals of small conductance amplitude and short duration. The issue of noise in the ion current recordings and the ability to distinguish the molecule-specific signal from the intrinsic noise of the nanopore system are therefore very important. Understanding the electronic reason for noise is the first step to improve the pores and achieve accurate recordings with a high signal-to-noise ratio. Given that solid-state nanopores tend to be "noisy" even in contact with simple electrolytes such as KCl [272], Uram et al. presented a thorough overview of the bandwidth requirements and the main sources of noise in current recordings with glass and polymer nanopores [273]. Further theoretical studies will shed light on the electronic and molecular reasons for noise and help to improve the properties of solid-state nanopores.

In conclusion, nanopore research is a highly interdisciplinary and fast-growing area in nanotechnology. It bridges biological pores with robust solid-state nanopores that can be applied in sensing and lab-on-the chip devices. Nanopores form a unique platform suitable for single-molecule sensors and for studying chemical reactions, biomolecular recognition, and electrostatic interactions at the nanoscale. Further advances in nanopore research will meet remaining challenges, which include the atomistic-localized chemical modification of solid-state nanopores, a fundamental understanding of their noise properties, and the characterization of single nanopores with physical methods other than electrochemical techniques.

Acknowledgments

Z.S. acknowledges financial support from the National Science Foundation (CHE 0747237).

References

1. Hladky, S. B., and D. A. Haydon. 1970. Discreteness of conductance change in bimolecular lipid membranes in the presence of certain antibiotics. *Nature* 225:451–453.
2. Mueller, P., D. O. Rudin, H. T. Tien, and W. C. Wescott. 1962. Reconstitution of cell membrane structure *in vitro* and its transformation into an excitable system. *Nature* 194:979–980.
3. Neher, E., and B. Sakmann. 1976. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* 260:799–802.
4. Hamill, O. P., A. Marty, E. Neher, B. Sakmann, and F. J. Sigworth. 1981. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Arch.* 391:85–100.
5. Hille, B. 2001. *Ion channels of excitable membranes*. Sinauer Associates, Sunderland, MA.
6. Sakmann, B., and B. Neher. 1995. *Single-channel recording*. Plenum Press, New York.
7. Miller, C. 1986. *Ion channel reconstitution*. Springer, New York.
8. Bezrukov, S. M., I. Vodyanoy, and V. A. Parsegian. 1994. Counting polymers moving through a single ion channel. *Nature* 370:279–281.
9. Kasianowicz, J. J., E. Brandin, D. Branton, and D. W. Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* 93:13770–13773.
10. Braha, O., B. Walker, S. Cheley, J. J. Kasianowicz, L. Song, J. E. Gouaux, and H. Bayley. 1997. Designed protein pores as components for biosensors. *Chem. Biol.* 4:497–505.

11a. Griffiths, J. 2008. The realm of the nanopore. Interest in nanoscale research has skyrocketed, and the humble pore has become a king. *Anal. Chem.* 80:23–27.

11b. Rhee, M., and M. A. Burns. 2007. Nanopore sequencing technology: nanopore preparations. *Trends Biotechnol*. 25:174–181; Dekker, C. 2007. Solid-state nanopores. *Nature Nanotechnol*. 2:209–215.

12. Meller, A. 2008. Nucleic-acid analysis at the single-molecule level. In *Handbook of single-molecule biophysics*. P. Hinterdorfer and A. M. Van Oijen, editors. Springer, New York.
13. Vercoutere, W., and M. Akeson. 2002. Biosensors for DNA sequence detection. *Curr. Opin. Chem. Biol.* 6:816–822.
14. Deamer, D. W., and D. Branton. 2002. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* 35:817–825.
15. Healy, K. 2007. Nanopore-based single-molecule DNA analysis. *Nanomedicine* 2:459–481.

16a. Healy, K., B. Schiedt, and A. P. Morrison. 2007. Solid-state nanopore technologies for nanopore-based DNA analysis. *Nanomedicine* 2:875–897.

16b. Zwolak, M., and M. Di Ventra. 2008. *Colloquium:* Physical approaches to DNA sequencing and detection. *Rev. Mod. Phys.* 80:141–165.

16c. Branton, D., D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. H. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, and J. A. Schloss. 2008. The potential and challenges of nanopore sequencing. *Nature Biotechnol.* 26:1146–1153.

17. Coulter, W. H. 1953. Means for counting particles suspended in a fluid. U.S. Patent No. 2656508.
18. Bayley, H., and C. R. Martin. 2000. Resistive-pulse sensing—From microbes to molecules. *Chem. Rev.* 100:2575–2594.
19. Stanley-Wood, N. G., and R. W. Lines. 1992. Particle size analysis. The Royal Society of Chemistry, Cambridge.
20. Bianco, P. R., L. R. Brewer, M. Corzett, R. Balhorn, Y. Yeh, S. C. Kowalczykowski, and R. J. Baskin. 2001. Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* 409:374–378.
21. Rasnik, I., S. A. McKinney, and T. Ha. 2006. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* 3:891–893.

22. Myong, S., I. Rasnik, C. Joo, T. M. Lohman, and T. Ha. 2005. Repetitive shuttling of a motor protein on DNA. *Nature* 437:1321–1325.
23. Greulich, K. O. 2005. Single-molecule studies on DNA and RNA. *ChemPhysChem* 6:2458–2471.
24. Smith, D. E., S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante. 2001. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* 413:748–752.
25. Chemla, Y. R., K. Aathavan, J. Michaelis, S. Grimes, P. J. Jardine, D. L. Anderson, and C. Bustamante. 2005. Mechanism of force generation of a viral DNA packaging motor. *Cell* 122:683–692.
26. Abbondanzieri, E. A., W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block. 2005. Direct observation of base-pair stepping by RNA polymerase. *Nature* 438:460–465.
27. Neuman, K. C., E. A. Abbondanzieri, R. Landick, J. Gelles, and S. M. Block. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* 115:437–447.
28. Herbert, K. M., A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick, and S. M. Block. 2006. Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell* 125:1083–1094.
29. Perkins, T. T., H. W. Li, R. V. Dalal, J. Gelles, and S. M. Block. 2004. Forward and reverse motion of single RecBCD molecules on DNA. *Biophys. J.* 86:1640–1648.
30. Hansma, H. G., R. Golan, W. Hsieh, S. L. Daubendiek, and E. T. Kool. 1999. Polymerase activities and RNA structures in the atomic force microscope. *J. Struct. Biol.* 127:240–247.
31. Movileanu, L., and H. Bayley. 2001. Partitioning of a polymer into a nanoscopic protein pore obeys a simple scaling law. *Proc. Natl. Acad. Sci. USA* 98:10137–10141.
32. Movileanu, L., S. Cheley, S. Howorka, O. Braha, and H. Bayley. 2001. Location of a constriction in the lumen of a transmembrane pore by targeted covalent attachment of polymer molecules. *J. Gen. Physiol.* 117:239–252.
33. Krasilnikov, O. V., C. G. Rodrigues, and S. M. Bezrukov. 2006. Single polymer molecules in a protein nanopore in the limit of a strong polymer–pore attraction. *Phys. Rev. Lett.* 97:018301(1)–018301(4).
34. Bayley, H., O. Braha, and L. Gu. 2000. Stochastic sensing with protein pores. *Adv. Mater.* 12:139–142.
35. Howorka, S., L. Movileanu, O. Braha, and H. Bayley. 2001. Kinetics of duplex formation for individual DNA strands within a single protein nanopore. *Proc. Natl. Acad. Sci. USA* 98:12996–13001.

36a. Howorka, S., J. Nam, H. Bayley, and D. Kahne. 2004. Stochastic detection of monovalent and multivalent protein–ligand interactions. *Angew. Chem. Int. Ed. Engl.* 43:842–846.

36b. Mayer, M., V. Semetey, I. Gitlin, J. Yang, and G. M. Whitesides. 2008. Using ion channel-forming peptides to quantify protein-ligand interactions. *J. Am. Chem. Soc.* 130:1453–1465.

37. Hornblower, B., A. Coombs, R. D. Whitaker, A. Kolomeisky, S. J. Picone, A. Meller, and M. Akeson. 2007. Single-molecule analysis of DNA–protein complexes using nanopores. *Nat. Methods* 4:315–317.
38. Cockroft, S. L., J. Chu, M. Amorin, and M. R. Ghadiri. 2008. A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *J. Am. Chem. Soc.* 130:818–820.
39. Benner, S., R. J. A. Chen, N. A. Wilson, R. Abu-Shumays, N. Hurt, K. R. Lieberman, D. W. Deamer, W. B. Dunbar, and M. Akeson. 2007. Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. *Nat. Nanotechnol.* 2:718–724.
40. Astier, Y., D. E. Kainov, H. Bayley, R. Tuma, and S. Howorka. 2007. Stochastic detection of motor protein–RNA complexes by single-channel current recording. *ChemPhysChem* 8:2189–2194.
41. Luchian, T., S. H. Shin, and H. Bayley. 2003. Kinetics of a three-step reaction observed at the single-molecule level. *Angew. Chem. Int. Ed. Engl.* 42:1926–1929.
42. Loudwig, S., and H. Bayley. 2006. Photoisomerization of an individual azobenzene molecule in water: an on-off switch triggered by light at a fixed wavelength. *J. Am. Chem. Soc.* 128:12404–12405.
43. Movileanu, L., J. P. Schmittschmitt, J. M. Scholtz, and H. Bayley. 2005. Interactions of peptides with a protein pore. *Biophys. J.* 89:1030–1045.

44a. Stefureac, R., Y. T. Long, H. B. Kraatz, P. Howard, and J. S. Lee. 2006. Transport of alpha-helical peptides through alpha-hemolysin and aerolysin pores. *Biochemistry* 45:9172–9179.

44b. Zhao, Q., D. A. Jayawardhana, D. Wang, and X. Guan. 2009. Study of peptide transport through engineered protein channels. *J. Phys. Chem. B.* 113:3572–3578.

45. Kullman, L., M. Winterhalter, and S. M. Bezrukov. 2002. Transport of maltodextrins through maltoporin: a single-channel study. *Biophys. J.* 82:803–812.
46. Nestorovich, E. M., C. Danelon, M. Winterhalter, and S. M. Bezrukov. 2002. Designed to penetrate: time-resolved interaction of single antibiotic molecules with bacterial pores. *Proc. Natl. Acad. Sci. USA* 99:9789–9794.
47. Rostovtseva, T. K., and S. M. Bezrukov. 1998. ATP transport through a single mitochondrial channel, VDAC, studied by current fluctuation analysis. *Biophys. J.* 74:2365–2373.

48. Rostovtseva, T. K., A. Komarov, S. M. Bezrukov, and M. Colombini. 2002. VDAC channels differentiate between natural metabolites and synthetic molecules. *J. Membr. Biol.* 187:147–156.
49. Hume, R. I., L. W. Role, and G. D. Fischbach. 1983. Acetylcholine release from growth cones detected with patches of acetylcholine receptor–rich membranes. *Nature* 305:632–634.
50. Allen, T. G. 1997. The 'sniffer-patch' technique for detection of neurotransmitter release. *Trends Neurosci.* 20:192–197.
51. Bayley, H., and P. S. Cremer. 2001. Stochastic sensors inspired by biology. *Nature* 413:226–230.
52. Bayley, H., and L. Jayasinghe. 2004. Functional engineered channels and pores. *Mol. Membr. Biol.* 21:209–220.
53. Bayley, H., O. Braha, S. Cheley, and L. Q. Gu. 2005. Engineered nanopores. In *Nanobiotechnology: Concepts, applications and perspectives*. C. M. Niemeyer and C. A. Mirkin, editors. Wiley-VCH, Weinheim, Germany, pp. 93–110.
54. Bong, D. T., T. D. Clark, J. R. Granja, and M. R. Ghadiri. 2001. Self-assembling organic nanotubes. *Angew. Chem. Int. Ed. Engl.* 40:988–1011.
55. Matile, S., A. Som, and N. Sorde. 2004. Recent synthetic ion channels and pores. *Tetrahedron* 60:6405–6435.
56. Sakaki, Y., J. Mareda, and S. Matile. 2007. Ion channels and pores, made from scratch. *Molecular Biosystems* 3:658–666.
57. Song, L., M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, and J. E. Gouaux. 1996. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* 274:1859–1866.
58. Gouaux, J. E., O. Braha, M. R. Hobaugh, L. Song, S. Cheley, C. Shustak, and H. Bayley. 1994. Subunit stoichiometry of staphylococcal alpha-hemolysin in crystals and on membranes: a heptameric transmembrane pore. *Proc. Natl. Acad. Sci. USA* 91:12828–12831.
59. Benz, R. 2004. *Bacterial and eukaryotic porins: Structure, function, mechanism*. Wiley-VCH, Weinheim, Germany.
60. Conlan, S., Y. Zhang, S. Cheley, and H. Bayley. 2000. Biochemical and biophysical characterization of OmpG: a monomeric porin. *Biochemistry* 39:11845–11854.
61. Menestrina, G. 2003. *Pore-forming peptides and protein toxins*. Taylor & Francis, London.
62. Borisenko, V., T. Lougheed, J. Hesse, E. Füreder-Kitzmüller, N. Fertig, J. C. Behrends, G. A. Woolley, and G. J. Schütz. 2003. Simultaneous optical and electrical recording of single gramicidin channels. *Biophys. J.* 84:612–622.
63. Armstrong, K. M., E. P. Quigley, P. Quigley, D. S. Crumrine, and S. Cukierman. 2001. Covalently linked gramicidin channels: effects of linker hydrophobicity and alkaline metals on different stereoisomers. *Biophys. J.* 80:1810–1818.
64. Ghadiri, M. R., J. R. Granja, and L. K. Buehler. 1994. Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* 369:301–304.
65. Clark, T. D., K. Kobayashi, and M. R. Ghadiri. 1999. Covalent capture and stabilization of cylindrical beta-sheet peptide assemblies. *Chem-Eur. J.* 5:782–792.
66. Tokarz, M., B. Akerman, J. Olofsson, J. F. Joanny, P. Dommersnes, and O. Orwar. 2005. Single-file electrophoretic transport and counting of individual DNA molecules in surfactant nanotubes. *Proc. Natl. Acad. Sci. USA* 102:9127–9132.
67. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, editors. 2008. *Current protocols in molecular biology*. Wiley, New York.
68. Xie, J., and P. G. Schultz. 2006. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell. Biol.* 7:775–782.
69. Muralidharan, V., and T. W. Muir. 2006. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* 3:429–438.
70. Schwarzer, D., and P. A. Cole. 2005. Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr. Opin. Chem. Biol.* 9:561–569.
71. Blake, S., T. Mayer, M. Mayer, and J. Yang. 2006. Monitoring chemical reactions by using ion-channel-forming peptides. *ChemBioChem* 7:433–435.
72. Macmillan, D. 2006. Evolving strategies for protein synthesis converge on native chemical ligation. *Angew. Chem. Int. Ed. Engl.* 45:7668–7672.
73. Chen, M., S. Khalid, M. S. Sansom, and H. Bayley. 2008. Outer membrane protein G: engineering a quiet pore for biosensing. *Proc. Natl. Acad. Sci. USA* 105:6272–6277.
74. Merzlyak, P. G., M. F. P. Capistrano, A. Valeva, J. J. Kasianowicz, and O. V. Krasilnikov. 2005. Conductance and ion selectivity of a mesoscopic protein nanopore probed with cysteine scanning mutagenesis. *Biophys. J.* 89:3059–3070.

75a. Miedema, H., M. Vrouenraets, J. Wierenga, W. Meijberg, G. Robillard, and B. Eisenberg. 2007. A biological porin engineered into a molecular, nanofluidic diode. *Nano Lett.* 7:2886–2891.

75b. Miedema, H., A. Meter-Arkema, J. Wierenga, J. Tang, B. Eisenberg, W. Nonner, H. Hektor, D. Gillespie, and W. Meijberg. 2004. Permeation properties of an engineered bacterial OmpF porin containing the EEEE-locus of Ca^{2+} channels. *Biophys. J.* 87:3137–3147.

76. Guan, X., L. Q. Gu, S. Cheley, O. Braha, and H. Bayley. 2005. Stochastic sensing of TNT with a genetically engineered pore. *ChemBioChem* 6:1875–1881.

77. Xie, H., O. Braha, L. Q. Gu, S. Cheley, and H. Bayley. 2005. Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem. Biol.* 12:109–120.

78. Howorka, S., S. Cheley, and H. Bayley. 2001. Sequence-specific detection of individual DNA-strands using engineered nanopores. *Nat. Biotechnol.* 19:636–639.

79. Howorka, S., L. Movileanu, X. Lu, M. Magnon, S. Cheley, O. Braha, and H. Bayley. 2000. A protein pore with a single polymer chain tethered within the lumen. *J. Am. Chem. Soc.* 122:2411–2416.

80. Movileanu, L., S. Howorka, O. Braha, and H. Bayley. 2000. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat. Biotechnol.* 18:1091–1095.

81. Gu, L. Q., O. Braha, S. Conlan, S. Cheley, and H. Bayley. 1999. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* 398:686–690.

82. Wu, H. C., Y. Astier, G. Maglia, E. Mikhailova, and H. Bayley. 2007. Protein nanopores with covalently attached molecular adapters. *J. Am. Chem. Soc.* 129:16142–16148.

83. Litvinchuk, S., H. Tanaka, T. Miyatake, D. Pasini, T. Tanaka, G. Bollot, J. Mareda, and S. Matile. 2007. Synthetic pores with reactive signal amplifiers as artificial tongues. *Nat. Mater.* 6:576–580.

84. Montal, M., and P. Mueller. 1972. Formation of bimolecular membranes from lipid monolayers and study of their electric properties. *Proc. Natl. Acad. Sci. USA* 69:3561–3566.

85. Holden, M. A., and H. Bayley. 2005. Direct introduction of single protein channels and pores into lipid bilayers. *J. Am. Chem. Soc.* 127:6502–6503.

86. Holden, M. A., L. Jayasinghe, O. Daltrop, A. Mason, and H. Bayley. 2006. Direct transfer of membrane proteins from bacteria to planar bilayers for rapid screening by single-channel recording. *Nat. Chem. Biol.* 2:314–318.

87. Mayer, M., J. K. Kriebel, M. T. Tosteson, and G. M. Whitesides. 2003. Microfabricated Teflon membranes for low-noise recordings of ion channels in planar lipid bilayers. *Biophys. J.* 85:2684–2695.

88. White, R. J., E. N. Ervin, T. Yang, X. Chen, S. Daniel, P. S. Cremer, and H. S. White. 2007. Single ion-channel recordings using glass nanopore membranes. *J. Am. Chem. Soc.* 129:11766–11775.

89. Shenoy, D. K., W. R. Barger, A. Singh, R. G. Panchal, M. Misakian, V. M. Stanford, and J. J. Kasianowicz. 2005. Functional reconstitution of protein ion channels into planar polymerizable phospholipid membranes. *Nano Lett.* 5:1181–1185.

90. Jeon, T. J., N. Malmstadt, and J. J. Schmidt. 2006. Hydrogel-encapsulated lipid membranes. *J. Am. Chem. Soc.* 128:42–43.

91. Kang, X. F., S. Cheley, A. C. Rice-Ficht, and H. Bayley. 2007. A storable encapsulated bilayer chip containing a single protein nanopore. *J. Am. Chem. Soc.* 129:4701–4705.

92. Shim, J. W., and L. Q. Gu. 2007. Stochastic sensing on a modular chip containing a single-ion channel. *Anal. Chem.* 79:2207–2213.

93. Hromada, L. P., B. J. Nablo, J. J. Kasianowicz, M. A. Gaitan, and D. L. DeVoe. 2008. Single molecule measurements within individual membrane-bound ion channels using a polymer-based bilayer lipid membrane chip. *Lab Chip* 8:602–608.

94. Atanasov, V., N. Knorr, R. S. Duran, S. Ingebrandt, A. Offenhausser, W. Knoll, and I. Koper. 2005. Membrane on a chip: a functional tethered lipid bilayer membrane on silicon oxide surfaces. *Biophys. J.* 89:1780–1788.

95. Atanasov, V., P. P. Atanasova, I. K. Vockenroth, N. Knorr, and I. Koper. 2006. A molecular toolkit for highly insulating tethered bilayer lipid membranes on various substrates. *Bioconjug. Chem.* 17: 631–637.

96. Andersson, M., H. M. Keizer, C. Zhu, D. Fine, A. Dodabalapur, and R. S. Duran. 2007. Detection of single ion channel activity on a chip using tethered bilayer membranes. *Langmuir* 23:2924–2927.

97. Drexler, J., and C. Steinem. 2003. Pore-suspending lipid bilayers on porous alumina investigated by electrical impedance spectroscopy. *J. Phys. Chem. B* 107:11245–11254.

98. Deme, B., and D. Marchal. 2005. Polymer-cushioned lipid bilayers in porous alumina. *Eur. Biophys. J. Biophys.* 34:170–179.

99. Holden, M. A., D. Needham, and H. Bayley. 2007. Functional bionetworks from nanoliter water droplets. *J. Am. Chem. Soc.* 129:8650–8655.

100. Hwang, W. L., M. Chen, B. Cronin, M. A. Holden, and H. Bayley. 2008. Asymmetric droplet interface bilayers. *J. Am. Chem. Soc.* 130:5878–5879.
101. Fleischer, R. L., P. B. Price, and R. M. Walker. 1975. *Nuclear tracks in solids. Principles and applications*. University of California Press, Berkeley, CA.
102. http://www.gsi.de/forschung/mf/index_e.html, Gesellchaft fuer Schwerionenforschung, Materials Research, Darmstadt, Germany.
103. Harrell, C. C., S. B. Lee, and C. R. Martin. 2003. Synthetic single-nanopore and nanotube membranes. *Anal. Chem.* 75:6861–6867.
104. Spohr, R. 1983. Methods and device to generate a predetermined number of ion tracks. German Patent No. DE 2951376 C2; U. S. Patent No. 4369370.
105. Apel, P. Y., Y. E. Korchev, Z. Siwy, R. Spohr, and M. Yoshida. 2001. Diode-like single-ion track membrane prepared by electro-stopping. *Nucl. Instrum. Meth. B* 184:337–346.
106. Heins, E. A., Z. S. Siwy, L. A. Baker, and C. R. Martin. 2005. Detecting single porphyrin molecules in a conically shaped synthetic nanopore. *Nano Lett.* 5:1824–1829.
107. Bean, C., and W. DeSorbo. 1968. U.S. Patent No. 3770532.
108. Dobrev, D., J. Vetter, R. Neumann, and N. Angert. 2001. Conical etching and electrochemical metal replication of heavy-ion tracks in polymer foils. *J. Vac. Sci. Technol. B* 19:1385–1387.
109. Harrell, C. C., Z. S. Siwy, and C. R. Martin. 2006. Conical nanopore membranes: controlling the nanopore shape. *Small* 2:194–198.
110. Siwy, Z., P. Apel, D. Baur, D. D. Dobrev, Y. E. Korchev, R. Neumann, R. Spohr, C. Trautmann, and K. O. Voss. 2003. Preparation of synthetic nanopores with transport properties analogous to biological channels. *Surf. Sci.* 532:1061–1066.
111. Siwy, Z. S. 2006. Ion-current rectification in nanopores and nanotubes with broken symmetry. *Adv. Funct. Mater.* 16:735–746.
112. Siwy, Z., P. Apel, D. Dobrev, R. Neumann, R. Spohr, C. Trautmann, and K. Voss. 2003. Ion transport through asymmetric nanopores prepared by ion track etching. *Nucl. Instrum. Meth. B* 208:143–148.
113. Siwy, Z., D. Dobrev, R. Neumann, C. Trautmann, and K. Voss. 2003. Electro-responsive asymmetric nanopores in polyimide with stable ion-current signal. *Appl. Phys. A* 76:781–785.
114. Wharton, J. E., P. Jin, L. T. Sexton, L. P. Horne, S. A. Sherrill, W. K. Mino, and C. R. Martin. 2007. A method for reproducibly preparing synthetic nanopores for resistive-pulse biosensors. *Small* 3:1424–1430.
115. Park, S. R., H. B. Peng, and X. S. S. Ling. 2007. Fabrication of nanopores in silicon chips using feedback chemical etching. *Small* 3:116–119.
116. Li, J., D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko. 2001. Ion-beam sculpting at nanometre length scales. *Nature* 412:166–169.
117. Wu, M. Y., D. Krapf, M. Zandbergen, H. Zandbergen, and P. E. Batson. 2005. Formation of nanopores in a SiN/SiO2 membrane with an electron beam. *Appl. Phys. Lett.* 87:113106(1)–113106(3).
118. Cai, Q., B. Ledden, E. Krueger, J. A. Golovchenko, and J. L. Li. 2006. Nanopore sculpting with noble gas ions. *J. Appl. Phys.* 100:024914(1)–024914(6).
119. Stein, D. M., C. J. McMullan, J. L. Li, and J. A. Golovchenko. 2004. Feedback-controlled ion beam sculpting apparatus. *Rev. Sci. Instrum.* 75:900–905.
120. Storm, A. J., J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. 2003. Fabrication of solid-state nanopores with single-nanometre precision. *Nat. Mater.* 2:537–540.
121. Kim, M. J., M. Wanunu, D. C. Bell, and A. Meller. 2006. Rapid fabrication of uniformly sized nanopores and nanopore arrays for parallel DNA analysis. *Adv. Mater.* 18:3149–3153.
122. Kim, M. J., B. McNally, K. Murata, and A. Meller. 2007. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology* 18:205302(1)–205302(5).
123. Lo, C. J., T. Aref, and A. Bezryadin. 2006. Fabrication of symmetric sub-5 nm nanopores using focused ion and electron beams. *Nanotechnology* 17:3264–3267.
124. Wei, C., A. J. Bard, and S. W. Feldberg. 1997. Current rectification at quartz nanopipet electrodes. *Anal. Chem.* 69:4627–4633.
125. Karhanek, M., J. T. Kemp, N. Pourmand, R. W. Davis, and C. D. Webb. 2005. Single DNA molecule detection using nanopipettes and nanoparticles. *Nano Lett.* 5:403–407.
126. Umehara, S., N. Pourmand, C. D. Webb, R. W. Davis, K. Yasuda, and M. Karhanek. 2006. Current rectification with poly-L-lysine-coated quartz nanopipettes. *Nano Lett.* 6:2486–2492.
127. Korchev, Y. E., C. L. Bashford, M. Milovanovic, I. Vodyanoy, and M. J. Lab. 1997. Scanning ion conductance microscopy of living cells. *Biophys. J.* 73:653–658.
128. Hansma, P. K., B. Drake, O. Marti, S. A. C. Gould, and C. B. Prater. 1989. The scanning ion-conductance microscope. *Science* 243:641–643.

129. White, R. J., B. Zhang, S. Daniel, J. M. Tang, E. N. Ervin, P. S. Cremer, and H. S. White. 2006. Ionic conductivity of the aqueous layer separating a lipid bilayer membrane and a glass support. *Langmuir* 22: 10777–10783.
130. Schaffer, C. B., A. Brodeur, J. F. Garcia, and E. Mazur. 2001. Micromachining bulk glass by use of femtosecond laser pulses with nanojoule energy. *Opt. Lett.* 26:93–95.
131. Uram, J. D., K. Ke, A. J. Hunt, and M. Mayer. 2006. Label-free affinity assays by rapid detection of immune complexes in submicrometer pores. *Angew. Chem. Int. Ed. Engl.* 45:2281–2285.
132. Saleh, O. A., and L. L. Sohn. 2003. An artificial nanopore for molecular sensing. *Nano Lett.* 3:37–38.
133. Grabarek, Z., and J. Gergely. 1990. Zero-length crosslinking procedure with the use of active esters. *Anal. Biochem.* 185:131–135.
134. Vlassiouk, I., and Z. S. Siwy. 2007. Nanofluidic diode. *Nano Lett.* 7:552–556.
135. Siwy, Z., and A. Fulinski. 2002. Fabrication of a synthetic nanopore ion pump. *Phys. Rev. Lett.* 89:198103(1)–198103(4).
136. Siwy, Z., and A. Fulinski. 2004. A nanodevice for rectification and pumping ions. *Am. J. Phys.* 72: 567–574.
137. Ali, M., B. Schiedt, K. Healy, R. Neumann, and A. Ensinger. 2008. Modifying the surface charge of single track-etched conical nanopores in polyimide. *Nanotechnology* 19:085713(1)–085713(6).
138. Hanggi, P., and R. Bartussek. 1996. Brownian rectifiers: how to convert Brownian motion into directed transport. *Lect. Notes Phys.* 476:294–308.
139. Astumian, R. D. 1997. Thermodynamics and kinetics of a Brownian motor. *Science* 276:917–922.
140a. Hanggi, P., and F. Marchesoni. 2009. Artificial Brownian motors: Controlling transport on the nanoscale. *Rev. Mod. Phys.* 81:387–442.
140b. Cervera, J., B. Schiedt, and P. Ramirez. 2005. A Poisson/Nernst-Planck model for ionic transport through synthetic conical nanopores. *Europhys. Lett.* 71:35–41.
141a. Cervera, J., B. Schiedt, R. Neumann, S. Mafe, and P. Ramirez. 2006. Ionic conduction, rectification, and selectivity in single conical nanopores. *J. Chem. Phys.* 124:104706(1)–104706(9).
141b. Kosinska, I. D., I. Goychuk, M. Kostur, G. Schmid, and P. Hanggi. 2008. Rectification in synthetic conical nanopores: A one-dimensional Poisson-Nernst-Planck model. *Phys. Rev. E* 77:031131(1)–031131(10).
142. Karnik, R., C. H. Duan, K. Castelino, H. Daiguji, and A. Majumdar. 2007. Rectification of ionic current in a nanofluidic diode. *Nano Lett.* 7:547–551.
143. Ulman, A. 1996. Formation and structure of self-assembled monolayers. *Chem. Rev.* 96:1533–1554.
144. Kobayashi, Y., and C. R. Martin. 1999. Highly sensitive methods for electroanalytical chemistry based on nanotubule membranes. *Anal. Chem.* 71:3665–3672.
145. Jirage, K. B., J. C. Hulteen, and C. R. Martin. 1999. Effect of thiol chemisorption on the transport properties of gold nanotubule membranes. *Anal. Chem.* 71:4913–4918.
146. Martin, C. R., M. Nishizawa, K. Jirage, M. S. Kang, and S. B. Lee. 2001. Controlling ion-transport selectivity in gold nanotubule membranes. *Adv. Mater.* 13:1351–1362.
147. Jirage, K. B., J. C. Hulteen, and C. R. Martin. 1997. Nanotubule-based molecular-filtration membranes. *Science* 278:655–658.
148. Siwy, Z., E. Heins, C. C. Harrell, P. Kohli, and C. R. Martin. 2004. Conical-nanotube ion-current rectifiers: the role of surface charge. *J. Am. Chem. Soc.* 126:10850–10851.
149. Kohli, P., C. C. Harrell, Z. Cao, R. Gasparac, W. Tan, and C. R. Martin. 2004. DNA-functionalized nanotube membranes with single-base mismatch selectivity. *Science* 305:984–986.
150. Lee, S. B., D. T. Mitchell, L. Trofin, T. K. Nevanen, H. Soderlund, and C. R. Martin. 2002. Antibody-based bio-nanotube membranes for enantiomeric drug separations. *Science* 296:2198–2200.
151. Yu, S. F., S. B. Lee, and C. R. Martin. 2003. Electrophoretic protein transport in gold nanotube membranes. *Anal. Chem.* 75:1239–1244.
152. Chen, P., T. Mitsui, D. B. Farmer, J. Golovchenko, R. G. Gordon, and D. Branton. 2004. Atomic layer deposition to fine-tune the surface properties and diameters of fabricated nanopores. *Nano Lett.* 4:1333–1337.
153. Wanunu, M., and A. Meller. 2007. Chemically modified solid-state nanopores. *Nano Lett.* 7:1580–1585.
154. Nilsson, J., J. R. I. Lee, T. V. Ratto, and S. E. Letant. 2006. Localized functionalization of single nanopores. *Adv. Mater.* 18:427–431.
155. Danelon, C., C. Santschi, J. Brugger, and H. Vogel. 2006. Fabrication and functionalization of nanochannels by electron-beam-induced silicon oxide deposition. *Langmuir* 22:10711–10715.
156. Neher, E., and J. H. Steinbach. 1978. Local anaesthetics transiently block currents through single acetylcholine-receptor channels. *J. Physiol.* 277:153–176.

157a. Fukushima, Y. 1982. Blocking kinetics of the anomalous potassium rectifier of tunicate egg studied by single channel recording. *J. Physiol.* 331:311–331.
157b. Blake, S., R. Capone, M. Mayer, and J. Yang. 2008. Chemically reactive derivatives of gramicidin A for developing ion channel-based nanoprobes. *Bioconjugate Chem.* 19:1614–1624.
158. Akeson, M., D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer. 1999. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* 77:3227–3233.
159. Mathe, J., A. Aksimentiev, D. R. Nelson, K. Schulten, and A. Meller. 2005. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc. Natl. Acad. Sci. USA* 102:12377–12382.
160. Braha, O., L. Q. Gu, X. Lu, S. Cheley, and H. Bayley. 2000. Simultaneous stochastic sensing of divalent metal ions. *Nat. Biotechnol.* 18:1005–1007.
161. Harrell, C. C., Y. Choi, L. P. Horne, L. A. Baker, Z. S. Siwy, and C. R. Martin. 2006. Resistive-pulse DNA detection with a conical nanopore sensor. *Langmuir* 22:10837–10843.
162. Han, A., G. Schurmann, G. Mondin, R. A. Bitterli, N. G. Hegelbach, N. F. de Rooij, and R. Staufer. 2006. Sensing protein molecules using nanofabricated pores. *Appl. Phys. Lett.* 88:093901(1)–093901(3).
163. Sexton, L. T., L. P. Horne, S. A. Sherrill, G. W. Bishop, L. A. Baker, and C. R. Martin. 2007. Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor. *J. Am. Chem. Soc.* 129:13144–13153.
164. Wolfe, A. J., M. M. Mohammad, S. Cheley, H. Bayley, and L. Movileanu. 2007. Catalyzing the translocation of polypeptides through attractive interactions. *J. Am. Chem. Soc.* 129:14034–14041.
165. Bezrukov, S. M., and J. J. Kasianowicz. 1997. The charge state of an ion channel controls neutral polymer entry into its pore. *Eur. Biophys. J. Biophy.* 26:471–476.
166. Robertson, J. W. F., C. G. Rodrigues, V. M. Stanford, K. A. Rubinson, O. V. Krasilnikov, and J. J. Kasianowicz. 2007. Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci. USA* 104:8207–8211.
167. Vercoutere, W., S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, and M. Akeson. 2001. Rapid discrimination among individual DNA hairpin molecules at single- nucleotide resolution using an ion channel. *Nat. Biotechnol.* 19:248–252.
168. Dudko, O. K., J. Mathé, A. Szabo, A. Meller, and G. Hummer. 2007. Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. *Biophys. J.* 92:4188–4195.
169. Howorka, S., and H. Bayley. 2002. Probing distance and electrical potential within a protein pore with tethered DNA. *Biophys. J.* 83:3202–3210.
170. Goodrich, C. P., S. Kirmizialtin, B. M. Huyghues-Despointes, A. Zhu, J. M. Scholtz, D. E. Makarov, and L. Movileanu. 2007. Single-molecule electrophoresis of beta-hairpin peptides by electrical recordings and Langevin dynamics simulations. *J. Phys. Chem. B* 111:3332–3335.
171. Oukhaled, G., J. Mathé, A. L. Biance, L. Bacri, J. M. Betton, D. Lairez, J. Pelta, and L. Auvray. 2007. Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys. Rev. Lett.* 98:158101(1)–158101(4).
172. Kasianowicz, J. J., S. E. Henrickson, H. H. Weetall, and B. Robertson. 2001. Simultaneous multianalyte detection with a nanometer-scale pore. *Anal. Chem.* 73:2268–2272.
173. Mathe, J., H. Visram, V. Viasnoff, Y. Rabin, and A. Meller. 2004. Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.* 87:3205–3212.
174. Nakane, J., M. Wiggin, and A. Marziali. 2004. A nanosensor for transmembrane capture and identification of single nucleic Acid molecules. *Biophys. J.* 87:615–621.
175. Martin, C. R., and Z. S. Siwy. 2007. Learning nature's way: Biosensing with synthetic nanopores. *Science* 317:331–332.
176. Iqbal, S. M., D. Akin, and R. Bashir. 2007. Solid-state nanopore channels with DNA selectivity. *Nat. Nanotechnol.* 2:243–248.
177. Cheley, S., L. Q. Gu, and H. Bayley. 2002. Stochastic sensing of nanomolar inositol 1,4,5-trisphosphate with an engineered pore. *Chem. Biol.* 9:829–838.
178a. Braha, O., J. Webb, L. Q. Gu, K. Kim, and H. Bayley. 2005. Carriers versus adapters in stochastic sensing. *ChemPhysChem* 5:889–892.
178b. Astier, Y., O.Braha, and H. Bayley. 2006. Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *J. Am. Chem. Soc.* 128:1705–1710.

178c. Clarke, J., H. C. Wu, L. Jayasinghe, A. Patel, S. Reid, and H. Bayley. 2009. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol.* DOI: 10.1038/nnano.2009.12 </http://dx.doi.org/10.1038/nnano.2009.12.>

179. Kasianowicz, J. J., D. L. Burden, L. C. Han, S. Cheley, and H. Bayley. 1999. Genetically engineered metal ion binding sites on the outside of a channel's transmembrane beta-barrel. *Biophys. J.* 76:837–845.

180a. Siwy, Z., L. Trofin, P. Kohli, L. A. Baker, C. Trautmann, and C. R. Martin. 2005. Protein biosensors based on biofunctionalized conical gold nanotubes. *J. Am. Chem. Soc.* 127:5000–5001.

180b. Umehara, S., M. Karhanek, R. W. Davis, and N. Pourmand. 2009. Label-free biosensing with functionalized nanopipette probes. *Proc. Natl. Acad. Sci. U S A.* 106:4611–4616.

180c. Ali, M., B. Yameen, R. Neumann, W. Ensinger, W. Knoll, and W. Azzaroni. 2008. Biosensing and supramolecular bioconjugation in single conical polymer nanochannels. facile incorporation of biorecognition elements into nanoconfined Geometries. *J. Am. Chem. Soc.* 130:16351–16357.

180d. Ali, M., V. Bayer, B. Schiedt, R. Neumann, and A. Ensinger. 2008. Fabrication and functionalization of single asymmetric nanochannels for electrostatic/hydrophobic association of protein molecules. *Nanotechnology.* 19:485711–485719.

181. Wang, J., and C. R. Martin. 2008. A new drug-sensing paradigm based on ion-current rectification in a conically shaped nanopore. *Nanomedicine* 3:13–20.

182. Shin, S. H., T. Luchian, S. Cheley, O. Braha, and H. Bayley. 2002. Kinetics of a reversible covalent-bond-forming reaction observed at the single-molecule level. *Angew. Chem. Int. Ed. Engl.* 41:3707–3709.

183. Saleh, O. A., and L. L. Sohn. 2003. Direct detection of antibody-antigen binding using an on-chip artificial pore. *Proc. Natl. Acad. Sci. USA* 100:820–824.

184. Fologea, D., B. Ledden, D. S. McNabb, and J. Li. 2007. Electrical characterization of protein molecules by a solid-state nanopore. *Appl. Phys. Lett.* 91:053901(1)–053901(3).

185. Halverson, K. M., R. G. Panchal, T. L. Nguyen, R. Gussio, S. F. Little, M. Misakian, S. Bavari, and J. J. Kasianowicz. 2005. Anthrax biosensor, protective antigen ion channel asymmetric blockade. *J. Biol. Chem.* 280:34056–34062.

186. Jung, Y., H. Bayley, and L. Movileanu. 2006. Temperature-responsive protein pores. *J. Am. Chem. Soc.* 128:15332–15340.

187. Shin, S. H., and H. Bayley. 2005. Stepwise growth of a single polymer chain. *J. Am. Chem. Soc.* 127: 10462–10463.

188a. Capone, R., S. Blake, M. R. Restrepo, J. Yang, and M. Mayer. 2007. Designing nanosensors based on charged derivatives of gramicidin A. *J. Am. Chem. Soc.* 129:9737–9745.

188b. Blake, S., R. Capone, M. Mayer, and J. Yang. 2008. Chemically reactive derivatives of gramicidin A for developing ion channel-based nanoprobes. *Bioconjugate Chem.* 19:1614–1624.

189. Gu, L. Q., S. Cheley, and H. Bayley. 2003. Electroosmotic enhancement of the binding of a neutral molecule to a transmembrane pore. *Proc. Natl. Acad. Sci. USA* 100:15498–15503.

190. Ashkenasy, N., J. Sanchez-Quesada, H. Bayley, and M. R. Ghadiri. 2005. Recognizing a single base in an individual DNA strand: a step toward DNA sequencing in nanopores. *Angew. Chem. Int. Ed. Engl.* 44: 1401–1404.

191. Gu, L. Q., S. Cheley, and H. Bayley. 2001. Capture of a single molecule in a nanocavity. *Science* 291:636–640.

192. Mohammad, M. M., S. Prakash, A. Matouschek, and L. Movileanu. 2008. Controlling a single protein in a nanopore through electrostatic traps. *J. Am. Chem. Soc.* 130:4081–4088.

193. Wong, C. T., and M. Muthukumar. 2007. Polymer capture by electro-osmotic flow of oppositely charged nanopores. *J. Chem. Phys.* 126:164903(1)–164903(6).

194. Kovarik, M. L., and S. Jacobson, C. 2008. Integrated nanopore/microchannel devices for ac electrokinetic trapping of particles. *Anal. Chem.* 80:657–664.

195. Sanchez-Quesada, J., A. Saghatelian, S. Cheley, H. Bayley, and M. R. Ghadiri. 2004. Single DNA rotaxanes of a transmembrane pore protein. *Angew. Chem. Int. Ed. Engl.* 43:3063–3067.

196. Kasianowicz, J. J. 2004. Nanopores—flossing with DNA. *Nat. Mater.* 3:355–356.

197a. Keyser, U. F., B. N. Koeleman, S. van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker. 2006. Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys.* 2:473–477.

197b. Trepagnier, E. H., A. Radenovic, D. Sivak, P. Geissler, and J. Liphardt. 2007. Controlling DNA capture and propagation through artificial nanopores. *Nano Lett.* 7:2824–2830.

197c. Gershow, M., and J. A. Golovchenko. 2007. Recapturing and trapping single molecules with a solid-state nanopore. *Nat. Nanotech.* 2:775–779.

198. Sung, W., and P. J. Park. 1996. Polymer translocation through a pore in a membrane. *Phys. Rev. Lett.* 77:783–786.

199. Chuang, J., Y. Kantor, and M. Kardar. 2001. Anomalous dynamics of translocation. *Phys Rev E* 65:011802(1)–011802(8).
200. Lubensky, D. K., and D. R. Nelson. 1999. Driven polymer translocation through a narrow pore. *Biophys. J.* 77:1824–1838.
201. Muthukumar, M. 2203. Polymer escape through a nanopore. *J. Chem. Phys.* 118:5174.
202. Storm, A. J., C. Storm, J. Chen, H. Zandbergen, J. F. Joanny, and C. Dekker. 2005. Fast DNA translocation through a solid-state nanopore. *Nano Lett.* 5:1193–1197.
203. Sikorski, A., and P. Romiszowski. 2005. Computer simulation of polypeptide translocation through a nanopore. *J. Mol. Model.* 11:379–384.
204. Heng, J. B., C. Ho, T. Kim, R. Timp, A. Aksimentiev, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp. 2004. Sizing DNA using a nanometer-diameter pore. *Biophys. J.* 87:2905–2911.
205. Martin, H., H. Kinns, N. Mitchell, Y. Astier, R. Madathil, and S. Howorka. 2007. Nanoscale protein pores modified with PAMAM dendrimers. *J. Am. Chem. Soc.* 129:9640–9649.
206. Rousselet, J., L. Salome, A. Ajdari, and J. Prost. 1994. Directional motion of Brownian particles induced by a periodic asymmetric potential. *Nature* 370:446–448.
207. Fu, J. P., P. Mao, and J. Y. Han. 2005. Nanofilter array chip for fast gel-free biomolecule separation. *Appl. Phys. Lett.* 87:263902(1)–263902(3).
208. Fu, J. P., R. B. Schoch, A. L. Stevens, S. R. Tannenbaum, and J. Y. Han. 2007. A patterned anisotropic nanofluidic sieving structure for continuous-flow separation of DNA and proteins. *Nat. Nanotechnol.* 2:121–128.
209. Han, J., S. W. Turner, and H. G. Craighead. 1999. Entropic trapping and escape of long DNA molecules at submicron size constriction. *Phys. Rev. Lett.* 83:1688–1691.
210. Huang, L. R., P. Silberzan, J. O. Tegenfeldt, E. C. Cox, J. C. Sturm, R. H. Austin, and H. Craighead. 2002. Role of molecular size in ratchet fractionation. *Phys. Rev. Lett.* 89:178301(1)–178301(4).
211. Astumian, R. D., and P. Hanggi. 2002. Brownian motors. *Phys. Today* 55:33–39.
212. Chou, C. F., O. Bakajin, S. W. P. Turner, T. A. J. Duke, S. S. Chan, E. C. Cox, H. G. Craighead, and R. H. Austin. 1999. Sorting by diffusion: an asymmetric obstacle course for continuous molecular separation. *Proc. Natl. Acad. Sci. USA* 96:13762–13765.
213. van Oudenaarden, A., and S. G. Boxer. 1999. Brownian ratchets: molecular separations in lipid bilayers supported on patterned arrays. *Science* 285:1046–1048.
214. Han, J. Y., J. P. Fu, and R. B. Schoch. 2008. Molecular sieving using nanofilters: past, present and future. *Lab Chip* 8:23–33.
215. Chun, K. Y., and P. Stroeve. 2002. Protein transport in nanoporous membranes modified with self-assembled monolayers of functionalized thiols. *Langmuir* 18:4653–4658.
216. Ku, J. R., and P. Stroeve. 2004. Protein diffusion in charged nanotubes: "On-Off" behavior of molecular transport. *Langmuir* 20:2030–2032.
217. Saksena, S., and A. L. Zydney. 1994. Effect of solution pH and ionic-strength on the separation of albumin from immunoglobulins (Igg) by selective filtration. *Biotechnol. Bioeng.* 43:960–968.
218. Pujar, N. S., and A. L. Zydney. 1998. Electrostatic effects on protein partitioning in size-exclusion chromatography and membrane ultrafiltration. *J. Chromatogr. A* 796:229–238.
219. Burns, D. B., and A. L. Zydney. 2000. Buffer effects on the zeta potential of ultrafiltration membranes. *J. Membrane Sci.* 172:39–48.
220. Ramirez, P., A. Alcaraz, and S. Mafe. 2003. Uphill transport of amino acids through fixed charged membranes. In *Encyclopedia of surface and colloid science*. A. Hubbard, editor. Marcel Dekker, New York, pp. 1–12.
221. Ku, J. R., S. M. Lai, N. Ileri, P. Ramirez, S. Mafe, and P. Stroeve. 2007. pH and ionic strength effects on aminmino acid transport through Au-nanotubule membranes charged with self-assembled monolayers. *J. Phys. Chem. C* 111:2965–2973.
222. Savariar, E. N., K. Krishnamoorthy, and S. Thayumanavan. 2008. Molecular discrimination inside polymer nanotubules. *Nat. Nanotechnol.* 3:112–117.
223. Mitchell, D. T., S. B. Lee, L. Trofin, N. C. Li, T. K. Nevanen, H. Soderlund, and C. R. Martin. 2002. Smart nanotubes for bioseparations and biocatalysis. *J. Am. Chem. Soc.* 124:11864–11865.
224. Volkmuth, W. D., and R. H. Austin. 1992. DNA electrophoresis in microlithographic arrays. *Nature* 358:600–602.
225. Striemer, C. C., T. R. Gaborski, J. L. McGrath, and P. M. Fauchet. 2007. Charge- and size-based separation of macromolecules using ultrathin silicon membranes. *Nature* 445:749–753.
226. Yu, S. F., S. B. Lee, M. Kang, and C. R. Martin. 2001. Size-based protein separations in poly(ethylene glycol)-derivatized gold nanotubule membranes. *Nano Lett.* 1:495–498.

227. Biesheuvel, P. M., P. Stroeve, and P. A. Barneveld. 2004. Effect of protein adsorption and ionic strength on the equilibrium partition coefficient of ionizable macromolecules in charged nanopores. *J. Phys. Chem. B* 108:17660–17665.
228. Schoch, R. B., and P. Renaud. 2005. Ion transport through nanoslits dominated by the effective surface charge. *Appl. Phys. Lett.* 86:25311(1)–253111(4).
229. Lakshmi, B. B., and C. R. Martin. 1997. Enantioseparation using apoenzymes immobilized in a porous polymeric membrane. *Nature* 388:758–760.
230. Kuo, T. C., D. M. Cannon, Y. N. Chen, J. J. Tulock, M. A. Shannon, J. V. Sweedler, and P. W. Bohn. 2003. Gateable nanofluidic interconnects for multilayered microfluidic separation systems. *Anal. Chem.* 75: 1861–1867.
231. Flachsbart, B. R., K. Wong, J. M. Iannacone, E. N. Abante, R. L. Vlach, P. A. Rauchfuss, P. W. Bohn, J. V. Sweedler, and M. A. Shannon. 2006. Design and fabrication of a multilayered polymer microfluidic chip with nanofluidic interconnects via adhesive contact printing. *Lab Chip* 6:667–674.
232. Beck, R. E., and J. S. Schultz. 1970. Hindered diffusion in microporous membranes with known pore geometry. *Science* 170:1302–1305.
233. Steinle, E. D., D. T. Mitchell, M. Wirtz, S. B. Lee, V. Y. Young, and C. R. Martin. 2002. Ion channel mimetic micropore and nanotube membrane sensors. *Anal. Chem.* 74:2416–2422.
234. Lee, S. B., and C. R. Martin. 2002. Electromodulated molecular transport in gold-nanotube membranes. *J. Am. Chem. Soc.* 124:11850–11851.
235. Martin, C. R. 1994. Nanomaterials: a membrane-based synthetic approach. *Science* 266:1961–1966.
236. Martin, C. R., M. Nishizawa, K. Jirage, and M. Kang. 2001. Investigations of the transport properties of gold nanotubule membranes. *J. Phys. Chem. B* 105:1925–1934.
237. Kang, M. S., and C. R. Martin. 2001. Investigations of potential-dependent fluxes of ionic permeates in gold nanotubule membranes prepared via the template method. *Langmuir* 17:2753–2759.
238. Lee, S. B., and C. R. Martin. 2001. pH-switchable, ion-permselective gold nanotubule membrane based on chemisorbed cysteine. *Anal. Chem.* 73:768–775.
239. Kittilsland, G., G. Stemme, and B. Norden. 1990. A submicron particle filter in silicon. *Sensor. Actuat. A Phys.* 23:904–907.
240. Desai, T. A., D. Hansford, and M. Ferrari. 1999. Characterization of micromachined silicon membranes for immunoisolation and bioseparation applications. *J. Membrane Sci.* 159:221–231.
241. Nishizawa, M., V. P. Menon, and C. R. Martin. 1995. Metal nanotubule membranes with electrochemically switchable ion-transport selectivity. *Science.* 268:700–702
242. Plecis, A., R. B. Schoch, and P. Renaud. 2005. Ionic transport phenomena in nanofluidics: Experimental and theoretical study of the exclusion-enrichment effect on a chip. *Nano Lett.* 5:1147–1155.
243. Stein, D., M. Kruithof, and C. Dekker. 2004. Surface-charge-governed ion transport in nanofluidic channels. *Phys. Rev. Lett.* 93:035901(1)–035901(4).
244. Gu, L. Q., M. Dalla Serra, J. B. Vincent, G. Vigh, S. Cheley, O. Braha, and H. Bayley. 2000. Reversal of charge selectivity in transmembrane protein pores by using noncovalent molecular adapters. *Proc. Natl. Acad. Sci. USA* 97:3959–3964.

245a. Alcaraz, A., E. M. Nestorovich, M. Aguilella-Arzo, V. M. Aguilella, and S. M. Bezrukov. 2004. Salting out the ionic selectivity of a wide channel: the asymmetry of OmpF. *Biophys. J.* 87:943–957.

245b. Schoch, R. B., J. Han, and P. Renaud. 2008. Transport phenomena in nanofluidics. *Rev. Mod. Phys.* 80: 839–883.

246. Liu, Y. L., M. Q. Zhao, D. E. Bergbreiter, and R. M. Crooks. 1997. pH-switchable, ultrathin permselective membranes prepared from multilayer polymer composites. *J. Am. Chem. Soc.* 119: 8720–8721.

247a. van der Heyden, F. H. J., D. Stein, K. Besteman, S. G. Lemay, and C. Dekker. 2006. Charge inversion at high ionic strength studied by streaming currents. *Phys. Rev. Lett.* 96:224502–224505.

247b. Alcaraz, A., E. M. Nestorovich, M. L. Lopez, E. Garcia-Gimenez, S. M. Bezrukov, and V. M. Aguilella. 2009. Diffusion, exclusion, and specific binding in a large channel: A study of OmpF selectivity inversion. *Biophys. J.* 96:56–66.

247c. He, Y., D. Gillespie, I. Boda, I. Vlassiouk, R. S. Eisenberg, and Z. Siwy. 2009. Tuning transport properties of nanofluidic devices with local charge inversion. *J. Am. Chem. Soc.* 131:5194–5202.

248. Shklovskii, B. I. 1999. Screening of a macroion by multivalent ions: correlation-induced inversion of charge. *Phys. Rev. E* 60:5802–5811.
249. Daiguji, H., Y. Oka, and K. Shirono. 2005. Nanofluidic diode and bipolar transistor. *Nano Lett.* 5: 2274–2280.

250. Alcaraz, A., P. Ramirez, E. Garcia-Gimenez, M. L. Lopez, A. Andrio, and V. M. Aguilella. 2006. A pH-tunable nanofluidic diode: electrochemical rectification in a reconstituted single ion channel. *J. Phys. Chem. B* 110:21205–21209.
251. Schasfoort, R. B. M., S. Schlautmann, L. Hendrikse, and A. van den Berg. 1999. Field-effect flow control for microfabricated fluidic networks. *Science* 286:942–945.
252. Fan, R., M. Yue, R. Karnik, A. Majumdar, and P. D. Yang. 2005. Polarity switching and transient responses in single nanotube nanofluidic transistors. *Phys. Rev. Lett.* 95:086607(1)–086607(4).
253. Fan, R., Y. Y. Wu, D. Y. Li, M. Yue, A. Majumdar, and P. D. Yang. 2003. Fabrication of silica nanotube arrays from vertical silicon nanowire templates. *J. Am. Chem. Soc.* 125:5254–5255.
254. Karnik, R., R. Fan, M. Yue, D. Y. Li, P. D. Yang, and A. Majumdar. 2005. Electrostatic control of ions and molecules in nanofluidic transistors. *Nano Lett.* 5:943–948.
255a. Karnik, R., K. Castelino, and A. Majumdar. 2006. Field-effect control of protein transport in a nanofluidic transistor circuit. *Appl. Phys. Lett.* 88:123114(1)–123114(3).
255b. Kalman, E. B., O. Sudre, I. Vlassiouk, and Z. Siwy. 2009. Control of ionic transport through gated single conical nanopores. *Anal Bioanal Chem.* DOI: 10.1007/s00216-008-2545-3
256. Kalman, E. B., I. Vlassiouk, and Z. S. Siwy. 2008. Nanofluidic bipolar transistors. *Adv. Mater.* 20:293–297.
257. Ito, T., L. Sun, and R. M. Crooks. 2003. Simultaneous determination of the size and surface charge of individual nanoparticles using a carbon nanotube-based coulter counter. *Anal. Chem.* 75:2399–2406.
258. Hinds, B. J., N. Chopra, T. Rantell, R. Andrews, V. Gavalas, and L. G. Bachas. 2004. Aligned multiwalled carbon nanotube membranes. *Science* 303:62–65.
259. Holt, J. K., H. G. Park, Y. Wang, M. Stadermann, A. B. Artyukhin, C. P. Grigoropoulos, A. Noy, and O. Bakajin. 2005. Fast mass transport through sub-2-nanometer carbon nanotubes. *Science* 312:1034–1037.
260. Wan, R., J. Li, H. Lu, and H. Fang. 2005. Controllable water channel gating of nanometer dimensions. *J. Am. Chem. Soc.* 127:7166–7170.
261. Rasaiah, J. C., S. Garde, and G. Hummer. 2008. Water in nonpolar confinement: from nanotubes to proteins and beyond. *Annu. Rev. Phys. Chem.* 59:713–740.
262. Fornasiero, F., H. G. Park, J. K. Holt, M. Stadermann, C. P. Grigoropoulos, A. Noy, and O. Bakajin. 2008. Ion exclusion by sub-2-nm carbon nanotube pores. *Proc. Natl. Acad. Sci. USA* 105:17250–17255.
263. Yeh, I. C., and G. Hummer. 2004. Nucleic acid transport through carbon nanotube membranes. *Proc. Natl. Acad. Sci. USA* 101:12177–12182.
264. Blow, N. 2008. DNA sequencing: generation next-next. *Nat. Methods* 5:267–272.
265. Hemmler, R., G. Bose, R. Wagner, and R. Peters. 2005. Nanopore unitary permeability measured by electrochemical and optical single transporter recording. *Biophys. J.* 88:4000–4007.
266. Bruckbauer, A., P. James, D. Zhou, J. W. Yoon, D. Excell, Y. Korchev, R. Jones, and D. Klenerman. 2007. Nanopipette delivery of individual molecules to cellular compartments for single-molecule fluorescence tracking. *Biophys. J.* 93:3120–3131.
267. Zwolak, M., and M. Di Ventra. 2005. Electronic signature of DNA nucleotides via transverse transport. *Nano Lett.* 5:421–424.
268. Lagerqvist, J., M. Zwolak, and M. Di Ventra. 2007. Influence of the environment and probes on rapid DNA sequencing via transverse electronic transport. *Biophys. J.* 93:2384–2390.
269. Sigalov, G., J. Comer, G. Timp, and A. Aksmentiev. 2008. Detection of DNA sequences using an alternating electric field in a nanopore capacitor. *Nano Lett.* 8:56–63.
270a. Liang, X. G., and S. Y. Chou. 2008. Nanogap detector inside nanofluidic channel for fast real-time label-free DNA analysis. *Nano Lett.* 8:1472–1476.
270b. Fischbein, M. D., and M. Drndic. 2007. Sub-10 nm device fabrication in a transmission electron microscope. *Nano Lett.* 7:1329–1337.
270c. Tsutsui, M., M. Taniguchi, and T. Kawai. 2009. Transverse field effects on DNA-sized particle dynamics. *Nano Lett.* 9:1659–1662.
271. Su, X. D., and A. A. Berlin. 2006. Method and apparatus for nucleic acid sequencing and identification. U.S. Patent No. 2006019247.
272. Smeets, R. M., U. F. Keyser, N. H. Dekker, and C. Dekker. 2008. Noise in solid-state nanopores. *Proc. Natl. Acad. Sci. USA* 105:417–421.
273. Uram, J. D., K. Ke, and M. Mayer. 2008. Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano* 2:857–872.
274. Vlassiouk, I., S. Smirnov, and Z. Siwy. 2008. Nanofluidic ionic diodes. comparison of analytical and numerical solutions. *ACS Nano.* 2:1589–1602.

12

Single-Molecule Manipulation Using Optical Traps

Michael T. Woodside and Megan T. Valentine

Abstract One of the most sensitive tools for manipulating single molecules and measuring their properties is the optical trap, also known as optical tweezers. Consisting essentially of a strongly focused light beam, optical traps were first developed and demonstrated in the 1970s and 1980s by Arthur Ashkin and colleagues (Ashkin et al. 1986). These early pioneers showed that micron-sized dielectric particles could be held and manipulated in solution by using optical forces to create a stable, three-dimensional potential well. Since then, optical trapping instrumentation has been refined and developed such that piconewton forces are now routinely applied, while at the same time measuring the resultant displacements to nanometre or even angström resolution. As a result of these advances, optical traps have been applied widely, from cytometry to the study of mesoscopic colloids and polymers and of course the properties of single biological macromolecules. This chapter begins with a description of the theory and design of optical traps, followed by an illustrative discussion of applications to the study of structure formation and molecular motors, a description of typical "tricks of the trade" for using optical traps, and a brief look at techniques for extending the capabilities of traps.

12.1. Theory and Design of Optical Traps

12.1.1. Theory

The operating principles of optical traps can be understood by considering the ray optics picture in Figure 12.1 of a spherical bead near the focus of a laser beam. A small dielectric sphere of radius r in the Mie regime ($r >> \lambda$, where λ is the wavelength of the trapping laser) acts like a lens, changing the direction and/or focus of the laser beam. The momentum transfers associated with the deflections of the laser beam exert forces on the bead. If the sphere moves off the optical axis, the refracted light exerts a radial force pushing the sphere back on axis; if the sphere moves away from the focal point of the laser beam, the refracted light exerts an axial force toward the focal point. The sphere is thus held in a three-dimensional trap made of light, with an equilibrium position on the optical axis. There is a slight equilibrium axial deflection from the focal point in the direction of beam propagation due to the scattering force

M.T. Woodside • National Institute for Nanotechnology, National Research Council of Canada, and Department of Physics, University of Alberta, Alberta, Canada
M.T. Valentine • Department of Mechanical Engineering, University of California, Santa Barbara, CA, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_12,

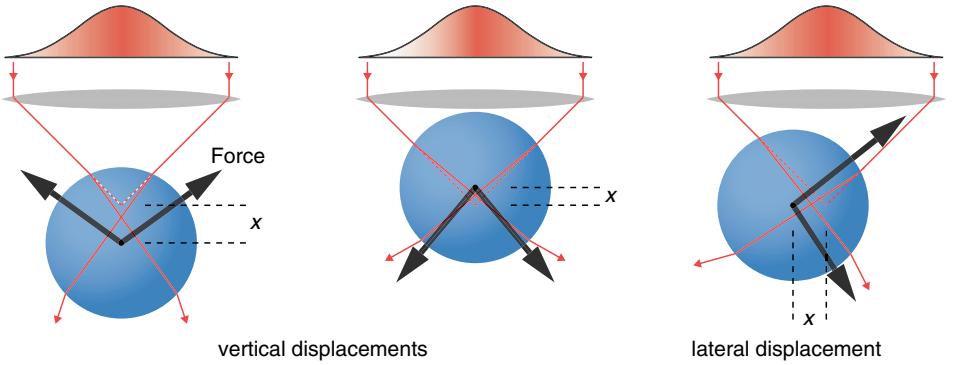


Figure 12.1. Ray optics schematic of optical trapping. Displacement of a bead away from the focal point of the beam deflects the beam, changing its momentum. The momentum transfer to the bead results in a restoring force in a direction opposite the bead displacement.

exerted on the sphere by the reflected rays. This picture gives a qualitative understanding of several features of optical traps: The forces are expected to vary directly with the beam intensity, the scattering force arises principally from the axial rays, and the trapping force arises principally from the extremal rays. High light intensities and high numerical aperture (NA) are thus essential to an effective trap.

A more quantitative picture of trapping can be gleaned from the Rayleigh limit ($r << \lambda$), in which the sphere is treated as a point dipole induced by the field of the laser beam. In this case, the time-averaged axial scattering force can be calculated from the energy absorbed and reradiated by the induced dipole:

$$F_{\text{scat}} = \frac{n_{\text{m}} C}{c} I \tag{12.1}$$

where I is the laser beam intensity, c is the speed of light in vacuum, $C = (8/3)\pi k^4 r^6 [(m^2 - 1)/(m^2 + 2)]$ is the scattering cross section of the sphere, $k = 2\pi/\lambda$ is the wavenumber, and $m = n_{\text{p}}/n_{\text{m}}$ is the ratio of the indices of refraction of the sphere (n_{p}) and the medium (n_{m}) (Harada and Asakura 1996). The gradient force can be calculated from the time average of the Lorenz force on the induced dipole, using the polarizability of a sphere, $\alpha = 4\pi {n_{\text{m}}}^2 r^3 [(m^2 - 1)/(m^2 + 2)]$: in MKS units,

$$\left\langle \vec{F}_{\text{grad}} \right\rangle = \left\langle (\vec{p} \cdot \vec{\nabla}) \vec{E} \right\rangle = \frac{2\pi n_{\text{m}}}{c} \left(\frac{m^2 - 1}{m^2 + 2} \right) r^3 \vec{\nabla} I \tag{12.2}$$

For a beam with a Gaussian profile, the intensity gradient is approximately linear near the centre of the beam. The trap thus acts as a Hookean spring, with force linearly dependent on displacement.

From the foregoing expressions, the scattering force increases with sphere radius faster than the gradient force. The bead cannot be too small, however, because the trapping potential must be significantly larger than the thermal energy, $k_{\text{B}}T$. These considerations place practical limits on the size of spheres that can be trapped: diameters usually range from ~200 to 5,000 nm. Thus the sphere size is of the same order as typical trapping wavelengths (500–1,100 nm), and neither of the limits described is truly representative. For proper

calculations of the forces, more complex theories are necessary: for example, extensions of Rayleigh theory to include higher-order scattering terms (Rohrbach and Stelzer 2001) or descriptions based on generalized Lorenz–Mie scattering theory (Xu et al. 2007). Because these methods do not shed additional insight into the physics of trapping, they are not explored further here.

12.1.2. Design of Optical Traps

Optical traps can be built based on a great variety of designs, each with its advantages and disadvantages, but all share common features: (1) one or more laser beams, (2) focusing optics, (3) elements for beam steering and/or mechanical sample manipulation, and (4) detection optics. A typical trap setup is illustrated schematically in Figure 12.2A. Detailed discussion of the elements of trap design can be found in a number of review articles (Block 1998; Neuman and Block 2004; Moffitt et al. 2008). Here, we comment only briefly on the key features of the design.

Laser Beams

Trapping lasers usually are chosen with wavelengths in the near infrared (800–1,100 nm) due to the high beam powers available, the relative transparency of biological

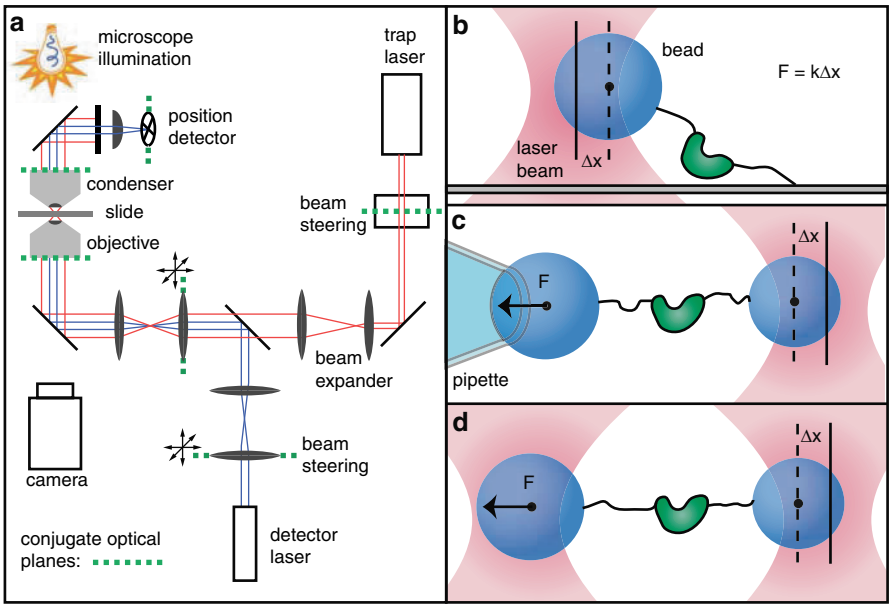


Figure 12.2. Schematic design of optical trap and trapping assays. **A.** Design elements include a trap laser, beam-steering elements, a beam expander, focusing optics, a position detector, and a camera; a separate detection laser may be used. Telescopes establish conjugate optical planes needed for beam steering. **B.** The surface-tethered assay: A single optical trap pulls on the molecule at one end while the other end is tethered to a microscope coverglass. **C.** The pipette-tethered dumbbell assay: The molecule is attached to two beads, one of which is held by an optical trap and the other by a pipette. **D.** The dual-trap dumbbell assay: The molecule is attached to two beads, both of which are held in optical traps.

materials, and the minimal damage to cells at these wavelengths (Neuman et al. 1999). There are several basic geometries for trapping beams. The simplest involves a single laser beam propagating in one direction, which is used to trap a bead bound to one end of the molecule of interest. The other end of the molecule is attached to a mechanical support such as a coverglass (Figure 12.2B) or a bead held on a micropipette (Figure 12.2C), creating a single-molecule tether that allows force to be applied to the molecule.

These two designs are versatile and straightforward to implement, but they have some disadvantages. First, the scattering force from a unidirectional beam shifts the trapping potential minimum away from the focal point, lowering the maximal force that can be applied (Neuman and Block 2004). This effect can be avoided (although at the cost of a more complex trap design) by including a second, counterpropagating beam (Smith et al. 2003), so that the scattering forces of the two beams are balanced. More important, the single-trap designs are very sensitive to motions of the mechanical support with respect to the laser beam. Such noise can be reduced by using a differential geometry tethering the molecule between two beads held in two orthogonally polarized traps generated from the same laser (Figure 12.2D). In this geometry, motions of the laser beams and surfaces are decoupled and no longer affect the measurement. If more than two traps are needed, they can be generated either by time-sharing a single laser beam between different trap locations or by using spatial light modulation to create so-called holographic optical traps (Grier and Roichman 2006). Such geometries allow for the simultaneous manipulation of multiple molecules for the study of macromolecular complexes or junctions (Dame et al. 2006).

Sample Manipulation

To manipulate the molecules freely and change the forces applied to them, it is useful to control the position of the sample stage (for surface- or pipette-bound molecules) and/or the trapping beam. The most sensitive means of controlling the stage position is provided by closed-loop piezoelectric flexure stages, which can provide subnanometer-scale resolution and repeatability in three dimensions. The laser beam position may be controlled using motorized mirrors (e.g., piezoelectric or galvanometer scanning mirrors) or beam deflectors like acousto-optic deflectors (AODs) or electro-optic deflectors (EODs). Beam steering elements should be placed at optical planes conjugate to the back-focal plane of the objective (Figure 12.2A), so that beam rotations become translations in the sample plane. For systems in which time response is critical, beam deflectors that lack moving parts provide the most rapid means of manipulating the molecules ($\sim$1 μs).

Focusing Optics

Optical traps are often built as add-ons to an optical microscope, which provides a ready-made platform for focusing the laser beams and imaging the samples. The focusing optics usually consist of a microscope objective because of the need for high NA to maximize the gradient force (NA > 1.2 is needed for a good trap). Oil-immersion objectives have the highest NA ($\sim$1.4–1.45), but they require operation within a few micrometers of the coverglass surface to avoid spherical aberration. Trapping deep in solution is possible with water-immersion objectives, but these have lower NA (typically $\sim$1.2). Fluorescence objectives are often good choices because they are designed to have reasonable transmission in the near-infrared.

Detection Optics

A variety of methods may be used to detect the position or motion of particles held in the trap(s). The simplest involve imaging the particle directly using video or other means; however, the greatest sensitivity and bandwidth are provided by laser-based detection. The laser that traps the particle can also be used to detect the relative position of the particle within the trap. This is most commonly done via back-focal plane detection (Visscher et al. 1996): Light passing through the sample is collected by a position-sensitive photodetector that is optically conjugate to the back focal plane of the condenser lens (this ensures that the detector response is insensitive to the location of the trap). When multiple traps are present or when the absolute position of the particle must be determined, a separate laser at a different wavelength can be used. These methods are used to detect the lateral position of the particle; others have been developed for detecting axial position but are not discussed here (Neuman and Block 2004).

12.1.3. Calibration of Optical Traps

Precise determination of position is essential for quantitative optical trapping studies because it is necessary to determine both the force acting on a molecule (via the displacement within the trap) and any motion of the molecule. While the details of the calibration depend on the detection method used, generally the detector is calibrated by recording the signal as a trapped particle is moved by a predetermined amount. Whenever possible, position calibration is best done individually on each particle that is being measured. This eliminates errors arising from differences between the particles being calibrated and those being measured, such as size or shape dispersion. For more details, readers are referred to the literature (Neuman and Block 2004).

The force being applied by the optical trap must also be determined quantitatively. For counterpropagating beam traps, it is possible to calculate the momentum transfer directly by collecting all of the light scattered by the trapped particle (Smith et al. 2003). However, for single-beam gradient traps, this method is difficult to implement, hence other force calibration methods are used. The three most commonly applied techniques rely on determining the stiffness of the trap within a harmonic approximation. These use the variance in the position of a thermally fluctuating trapped particle, the power spectrum of the fluctuations, and the displacement induced by moving the fluid surrounding the particle at a fixed speed (Svoboda and Block 1994a).

The variance method relies on the equipartition theorem to relate the thermal fluctuations of the particle in a harmonic potential of stiffness α to the thermal energy: $\frac{1}{2}\alpha\langle x^2\rangle = \frac{1}{2}k_B T$. This is a simple method that does not depend explicitly on anything other than the stiffness of the trap and the temperature, and it is thus insensitive in principle to the particle shape, proximity to the surface, and viscosity. However, it is a biased estimator (the variance is always positive) and hence instrumental noise and filtering can strongly affect the calibration.

Rather than look at the variance, which is just the integral over frequency of the fluctuation power spectrum, the power spectrum itself can be measured directly. The power spectrum is Lorentzian, characterized by constant power at low frequency and inverse-square behaviour above a roll-off frequency f_0:

$$P(f) = \frac{k_B T}{\pi^2 \beta (f_0^2 + f^2)} \tag{12.3}$$

where β is the drag coefficient for the particle, and the roll-off frequency is related to the stiffness by $f_0 = \alpha/(2\pi\beta)$. The calibration depends on knowledge of β, which in turn depends on the particle size and shape, the fluid viscosity, and the proximity of any surfaces (Svoboda and Block 1994a). For very accurate determinations of α from the power spectrum, the frequency dependence of β can also be included, as well as instrumental effects from sampling and filtering (Berg-Sørensen and Flyvbjerg 2004). It should be noted that in addition to measuring the stiffness, the power spectrum provides a sensitive diagnostic of the quality of the trap: misalignment, electronic noise, and mechanical vibrations all produce non-Lorentzian spectra.

The most direct method for calibrating trap stiffness measures the displacement of the trapped particle when subjected to a viscous drag force produced by moving the particle through fluid: $F = \beta v = \alpha x$, where v is the fluid velocity. The fluid motion is usually generated by periodic motion of the sample stage. Transients influenced by the stage response should thus be ignored in the calibration. An important advantage of this method is that by varying the magnitude of the drag force, the stiffness can be measured at different displacements from the centre of the trap. This provides a measurement of the region over which the trap is harmonic, in contrast to the other two methods, which measure the stiffness only near the trap centre. A calibration of the distance-dependent stiffness is essential because measurements are typically made ~50–100 nm from the trap centre to maximize the displacement signal. Indeed, careful measurements of the force–displacement curve of an optical trap using a dual-trap, dual-detector instrument (Greenleaf et al. 2005) indicate that the effective stiffness at a displacement of 100 nm can deviate significantly from the value at the trap centre (Figure 12.3).

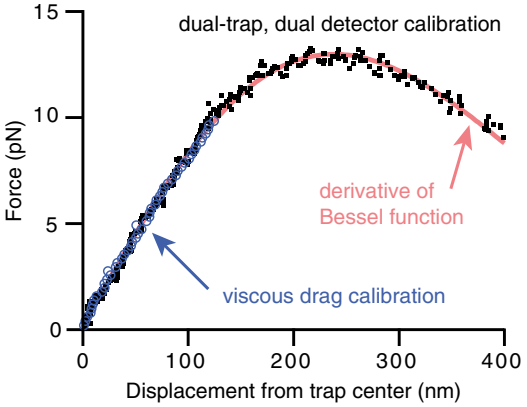


Figure 12.3. Displacement-dependent trap stiffness. The stiffness of an optical trap is determined by a viscous drag calibration (*blue*) and separately by using a dsDNA tether held in a dual-trap, dual-detector geometry to measure force and displacement simultaneously (*black*). The measurements are well fit by a derivative of a Bessel function (*red*). [Adapted from Greenleaf et al. (2005).]

Because each method has different systematic errors, it is usually best to use all three methods whenever possible. Several variations of these techniques have been developed with specific advantages. One that is particularly notable involves a simultaneous calibration of position and stiffness by measuring the power spectrum of a trapped particle while the sample stage is oscillated (Tolić-Nørrelykke et al. 2006), which does not require *a priori* knowledge of β.

12.1.4. Implementation of Single-Molecule Optical Trapping Assays

Single biological molecules are generally too small to be trapped directly. They are therefore usually attached to a micron-sized bead, and forces are applied via the bead (Figure 12.2B–D). These attachments are typically made by functionalizing the molecules and beads with ligand–receptor pairs, which allows for specific control of the molecule–bead link. Commonly used pairs include biotin/avidin (or streptavidin), digoxigenin/antidigoxigenin, fluorescein/antifluorescein, and pentahistidine/antipentahistidine tags. Complementary tags are typically covalently attached to the beads and molecules via functional carboxyl, amino, or sulfhydryl groups; the specific binding then produces a well-defined bead–molecule geometry, with bond strengths capable of withstanding tens to hundreds of piconewtons. The attachment chemistry is often dependent on the molecule being studied, and interested readers are referred to detailed descriptions of techniques for functionalizing molecules published elsewhere (Hermanson 1996).

A second important consideration when setting up an assay is the effect of surfaces. Each end of the molecule is attached to a surface (bead or coverglass), which can have important effects. Molecules can interact with the surface (e.g., adhering nonspecifically, denaturing, or reacting to surface charges), and the surface locally changes the viscosity of the fluid medium. Some of these effects can be mitigated by using blocking proteins (such as bovine serum albumen or casein) or engineered poly(ethylene glycol) and/or polyelectrolyte multilayers (Fordyce et al. 2007). Another common strategy is to separate the molecule from the surfaces by means of "handles" inserted between the attachment points on the molecule and the surfaces. The force applied by the trap to the beads is then transmitted to the molecule via the handles. Nucleic acid duplexes (DNA or DNA/RNA hybrids) are often used for handles because their mechanical properties are well known (see section Elastic Properties of DNA) and it is easy to control their length and functionalisation.

12.1.5. Technical Capabilities

Optical traps bring a number of characteristic capabilities to single-molecule biophysics that are distinct from those of other force probe techniques. They can apply moderate forces, in the range ~0.01–100 pN, on the same order as magnetic tweezers but considerably less than atomic force microscopes (1–1,000 pN). This is a direct result of the low stiffness of optical traps, typically ~0.01–1 pN/nm. They are distinguished by their ability to measure processes on time scales ranging over seven orders or magnitude, from ~100 μs to 1 hr (Woodside et al. 2006a). The low end of this range is limited by the intrinsic roll-off frequency arising from the stiffness of the bead–molecule system (Section 12.1.3); the high end is limited by the long-term stability of the trapping apparatus. Optical traps can also achieve superlative spatial resolution, ~1 Å/$Hz^{1/2}$ or less under optimal conditions. The combination of high temporal bandwidth and spatial resolution makes optical tweezers an excellent tool for studying processes like molecular folding or the motion of mechanoenzymes, which require precise time and distance measurements of small, rapid motions under controlled load. In-depth comparisons of optical traps with other methods can be found in a number of reviews (Greenleaf et al. 2007; Neuman and Nagy 2008).

The full capabilities of optical traps can only be achieved with careful design and construction of the instrument. To maximize time resolution, fast detectors should be used and experiments designed with smaller beads and stiffer molecules or handles wherever possible. Good spatial resolution requires attention to many details: Lasers should have low pointing

fluctuations or be fibre coupled, the instrument should be isolated vibrationally and acoustically from the environment, differential thermal expansion in the instrument should be minimized by controlling the temperature, and disruptive air currents should be blocked by enclosing the optical paths in a box; helium gas can be used in the box to reduce beam fluctuations further (Abbondanzieri et al. 2005). The geometry of the assay also plays a role. Where contact of the molecule with a surface is not needed, dual-trap dumbbell geometries (Figure 12.2D) with differential position detection provide the highest resolution (Greenleaf et al. 2005; Moffitt et al. 2006). This geometry decouples relative motion of the laser beam and the sample stage and allows differential measurement. Assays involving attachments to mechanically supported surfaces (Figures. 12.2B, C) are more sensitive to mechanical noise, but this can be compensated by measuring the position of a fiducial mark on the surface, effectively creating a differential measurement (Nugent-Glandorf and Perkins 2004). Finally, fluctuations in the laser intensity can be minimized by using feedback loops (Carter et al. 2007).

For many types of measurements, it is useful to maintain a constant load on the molecule being studied. Constant force removes the need to apply corrections to displacement signals due the series elastic compliance of the trap and molecule. It also avoids the complications that motion against a changing load introduces to the potential energy landscape of the molecule. Typically, force is kept constant using an active feedback loop that maintains a constant displacement of the trapped particle from the centre of the trap. Such active force clamps are straightforward to implement but suffer from reduced temporal resolution due to the finite feedback response time. This limitation can be overcome by creating one-dimensional line traps (Nambiar et al. 2004) or by trapping the particle not in the harmonic part of the potential close to the trap centre, but in the anharmonic part near the maximum of the force–displacement curve (Figure 12.3), where force is locally constant as a function of displacement (Greenleaf et al. 2005).

Optical traps are versatile instruments; they are capable of highly controlled, precise measurements, but at the price of complexity. Given the time, effort, and expense involved in building an optical trap, it is best to match the instrumental capabilities to the needs of the experimental system being investigated. The size of the forces to be applied, the ability to move the trap dynamically, the temporal and spatial resolution, the accuracy of absolute calibrations, and so on, all depend on the molecules under study. As a simple example, when measuring slow processes the emphasis should be on long-term stability rather than low noise, whereas for fast processes, low noise and maximal temporal resolution are paramount.

12.2. Applications to Studying Single Molecules

Most of the work on biological macromolecules using optical traps can be divided into two broad classes: studies of structural and mechanical properties (including folding and binding reactions) and studies of molecular motors (motion-producing enzymes). The practical use of optical traps in single-molecule biophysics is illustrated here with some specific examples in each of these classes. Examples have been chosen to highlight some of the important measurement methods used with optical traps, showing how the instrumental capabilities can be adapted to the needs of particular experiments, what can be learned from optical trapping data, and how they can be analyzed. These methods can often be applied quite generally to a wide range of systems.

12.2.1. Studies of Structural and Mechanical Properties

Because of their ability to apply force to a single molecule and to measure sensitively the resulting motion, optical traps are ideal for studying conformational changes and the mechanical properties of biological molecules. Structural changes resulting in a change of molecular extension can be observed directly, kinetic characteristics can be measured, binding or folding energies can be determined, and elastic properties can be characterized. There is also a significant body of theoretical work that helps to interpret and model such measurements. The basic techniques for studying structural and/or mechanical properties are discussed using three examples: the elastic properties of DNA, folding reactions, and binding reactions.

Elastic Properties of DNA

Occurring in the cell mainly as a double helix, DNA is a model polymer whose elastic properties have been probed extensively by optical trapping measurements (Smith et al. 1996; Wang et al. 1997). A single, double-stranded DNA (dsDNA) molecule subjected to force in an optical trap shows a characteristic relationship between force and extension, illustrated by the force–extension curve (FEC) in Figure 12.4. FECs, which plot the force on the molecule as the extension of the molecule is increased by moving the attachment points apart, are one of the most basic types of measurements using optical traps. For dsDNA, the force rises slowly up to ∼2 pN, at which point it begins to rise very quickly. This behaviour has been modelled by a number of different polymer theories, but fits best to a worm-like chain model treating the dsDNA as an elastic rod (Marko and Siggia 1995; Wang et al. 1997). A useful interpolation formula for worm-like chains has been developed for fitting FECs:

$$F(x) = \frac{k_B T}{L_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K}\right] \tag{12.4}$$

where L_p is the persistence length, L_c the contour length, and K the elastic modulus of the dsDNA. Entropic elasticity of the polymer chain dominates at low forces, whereas enthalpic elasticity of the bonds in the backbone dominates at high forces.

The persistence length of dsDNA has been measured under various solvent conditions and is typically ∼40–50 nm, and the elastic modulus is ∼1,000–1,200 pN (Wang et al. 1997). The apparent persistence length depends on the length of the polymer and becomes smaller for short molecules (Seol et al. 2007). Because L_p is well known for dsDNA, it can be used to determine whether measurements are being made on a single molecule or on several molecules at once. As a result, dsDNA is commonly used as a "handle" for attaching the molecule of interest to the surfaces and/or beads.

The elasticity of single-stranded DNA (ssDNA) has also been measured (Figure 12.4) and shows quite different behaviour (Smith et al. 1996): It is more resistant to force at low forces but more compliant at high forces. Simple polymer models like worm-like chains or freely jointed chains can be used, producing characteristic lengths of ∼1 nm, but they fail to capture the full complexity of ssDNA behaviour. Instead, more-complex models taking into account self-avoidance and electrostatics are needed (Dessinges et al. 2002). The difference between the elasticity of dsDNA and ssDNA can be used to probe the conversion of ssDNA into dsDNA or vice versa (Wuite et al. 2000).

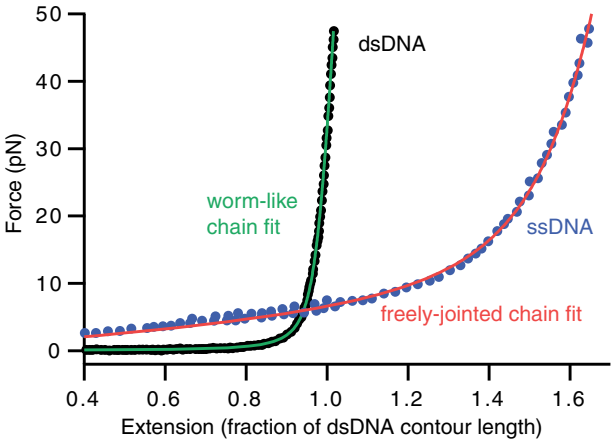


Figure 12.4. Force–extension curves (FECs) of double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA). The dsDNA FEC (*black*) is well fit by a worm-like chain model, Eq. (12.4) (*green*), whereas the ssDNA FEC (*blue*) is fit by a freely jointed chain model (*red*). [Adapted from Wang et al. (1997) for dsDNA and Smith et al. (1996) (ssDNA).]

Folding Studies

The "folding" process, by which macromolecules like nucleic acids and proteins take on their specific three-dimensional structures, is one of the central topics in biology because of the critical relationship between structure and function. Folding is usually studied by denaturing the molecular structure (e.g., using chemical denaturants, temperature, or pH) and then monitoring the structural changes using various forms of spectroscopy (e.g., fluorescence, circular dichroism, nuclear magnetic resonance). The force applied by an optical trap can also be used to denature molecules, and indeed single-molecule force spectroscopy is an increasingly common tool used to probe folding in proteins (Borgia et al. 2008) and nucleic acids (Woodside et al. 2008).

Force spectroscopy provides a number of advantageous features. Because the denaturant (force) is applied only to a single molecule, measurements on molecules of widely varying stabilities can be compared under identical buffer conditions. The ability to measure individual folding trajectories allows rare or transient states (such as partially folded intermediates) to be observed directly and subpopulations with different folding behaviour to be distinguished. Because unfolded states are fully stretched out under load, the unfolded state is simplified from an ensemble of high-entropy states to the low-entropy, fully extended configuration. Both initial and final (unfolded/folded) states are thus well defined, which can help to simplify interpretation. The molecular extension that is measured is also a natural coordinate for describing the course of the reaction, and changes in extension can often be interpreted in terms of specific structural elements.

One important aspect of force spectroscopy is that the vectorial nature of force imposes a preferential axis for the folding reaction. This raises the important question of how folding pathways of molecules denatured by force differ from those denatured in other ways. Evidence from both experimental and theoretical studies (Williams et al. 2003; Best et al. 2008; Hyeon and Thirumalai 2008) indicates that differences may indeed exist. While this does not make force spectroscopy a less valuable tool, these differences should be taken into account

when interpreting the data. The vectorial aspect of force can also be turned to advantage; for example, by changing the locations to which force is applied on the molecule, the choice of pathway can be changed (Dietz et al. 2006).

Because most folding studies using optical traps have concentrated on nucleic acids rather than proteins, we illustrate methods for measurement and analysis with examples from nucleic acid folding. The same approaches can also be applied to study protein folding (Cecconi et al. 2005). The typical experimental geometry has duplex handles (DNA or DNA/RNA hybrid) attached to the ends of the nucleic acid being studied and bound specifically to beads held by two optical traps (or one trap and a mechanical support). Measurements usually take one of two forms: FECs or extension trajectories under constant load. In both forms, the molecular extension increases/decreases when the molecule unfolds/folds. Similar types of information can be extracted from each measurement, but the detailed analysis differs.

Force–Extension Curves

In FECs, extension changes are accompanied by changes in force because of the displacement of the bead in the trap that results from the folding. This produces a characteristic sawtooth pattern of rips for each structure that unfolds, as seen in Figure 12.5A for an adenine riboswitch aptamer (a small adenine-binding RNA). The slope of the rip is determined by the stiffness of the trap and is thus instrument dependent, whereas the length of the rip is determined by the mechanical properties of the molecule. The change in contour length can be measured quite precisely, especially by averaging many FECs (Figure 12.5B); this allows the number of nucleotides unfolded to be determined and correlated with structural features of the molecule.

The unfolding force also reveals important information. Molecules unfold in FECs not at a single force value, but with a distribution (Figure 12.5C) whose shape depends on such factors as the rate at which the force is changed, the unfolding rate at zero force, and the shape of the energy landscape (which is what fundamentally determines the folding behaviour). Because force tilts the energy landscape (Figure 12.5D), the probability to unfold rises with the force. Taking into account the change in the location of the energy barrier with force, we can express the unfolding force distribution as (Dudko et al. 2006):

$$p(F) \propto \frac{k(F)}{r} \exp\left\{ \frac{k_{off}}{\Delta x^{\ddagger} r} - \frac{k(F)}{\Delta x^{\ddagger} r} \left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}} \nu\right)^{1-1/\nu} \right\} \tag{12.5}$$

where

$$k(F) = k_{off} \left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}} \nu\right)^{1/\nu - 1} \exp\left\{ \Delta G^{\ddagger} \left[1 - \left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}} \nu\right)^{1/\nu} \right] \right\}$$

k_{off} is the unfolding rate at zero force, $\Delta x^{\ddagger}$ is the distance to the transition state (the top of the energy barrier) from the folded state, $\Delta G^{\ddagger}$ is the height of the barrier, and ν describes the shape of the barrier: $\nu = 1/2$ for a sharp, cusp-like barrier, and $\nu = 2/3$ for a soft, cubic potential. A simpler expression equivalent to $\nu = 1$ is often used that neglects the force

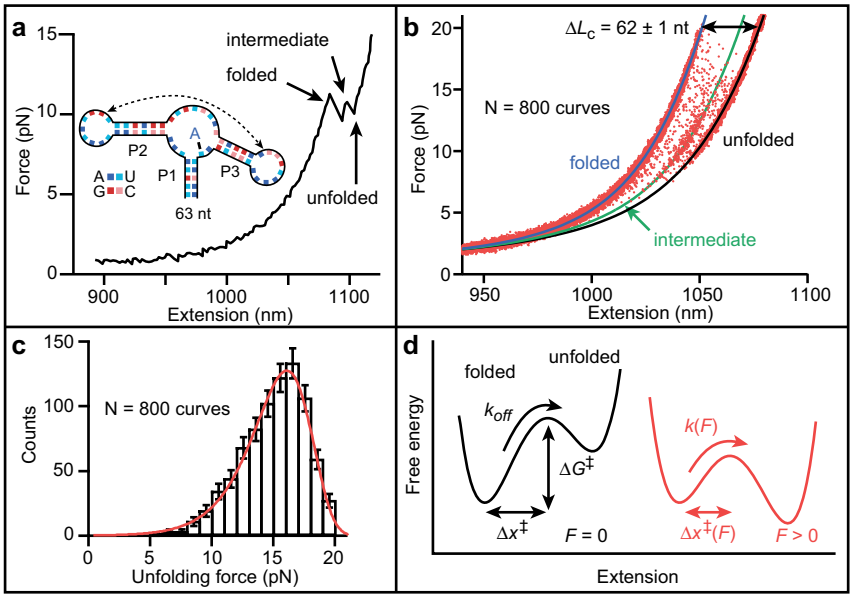


Figure 12.5. Unfolding force–extension curves (FECs) of an adenine riboswitch aptamer. **A.** The aptamer FEC shows two unfolding events where the length increases. **Inset:** Aptamer secondary structure. **B.** Multiple FECs provide a high-precision measure of contour length changes between states. **C.** Unfolding force distribution can be fit to kinetic theories, here Eq. (12.5) (*red*), to extract parameters describing the energy landscape. **D.** Force tilts the energy landscape, changing the barrier height and distance to the transition state. [Panels A–C adapted from Greenleaf et al. (2008).]

dependence of the barrier location (Evans and Ritchie 1997), but it tends to overestimate k_{off} considerably.

Fitting the measured force distribution thus yields important kinetic information, as well as parameters describing the energy landscape. $\Delta x^{\ddagger}$ is particularly useful because the known force–extension relation for nucleic acids can be used to convert $\Delta x^{\ddagger}$ into the number of nucleotides involved in the transition state, providing strong clues as to its structure. For example, in the case of the riboswitch aptamer, the transition state for unfolding the adenine-bound state was found to involve disrupting the first three base pairs in helix P1 (Figure 12.5A, inset).

The integral of the FEC gives the work done on the system, yielding information about the free energy of folding. However, this analysis is complicated by several factors. Because the unfolded state is extended by the load, the energy to stretch out the unfolded molecule must be subtracted before the energy can be compared to the results of other methods. The energy stored in the handles must also be taken into account. The work done to unfold the molecule is thus given not simply by the area under the unfolding rip in the sawtooth pattern, but rather by the difference between the integral of the measured FEC and the integral of the fit to the FEC for the fully unfolded molecule (Figure 12.6A).

Work may be dissipated in these measurements if the pulling rate is too fast to permit the system to equilibrate (evident experimentally as hysteresis in the unfolding and refolding curves). In this case, equilibrium free energies may be recovered from FECs using one of the recent fluctuation theorems, such as Jarzynski's equality (Jarzynski 1997) or Crooks'

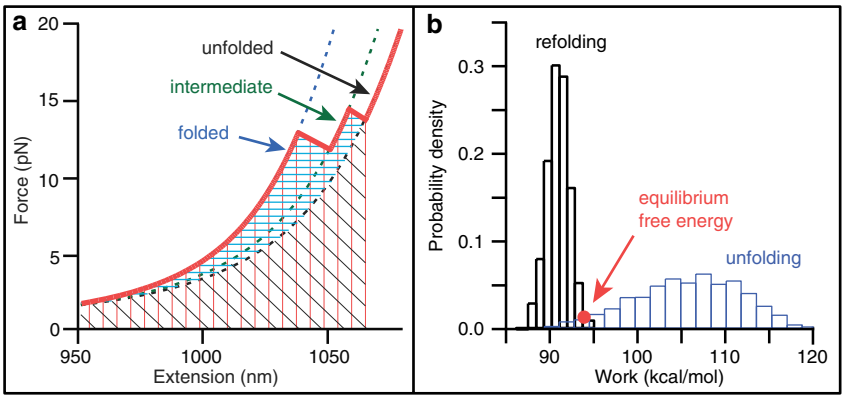


Figure 12.6. Free energy measurements. **A.** The work done to unfold the molecule is given by the area between the measured unfolding force–extension curve (FEC) and the FEC for the unfolded state (*horizontal hatching*). **B.** For measurements that are out of equilibrium, the Crooks fluctuation theorem gives the equilibrium free energy as the energy at which the work distributions for folding (*black*) and unfolding (*blue*) intersect (*red*). [Adapted from Collin et al. (2005).]

theorem (Crooks 1999). Jarzynski's equality relates the equilibrium free energy, ΔG_{eqm}, to an exponential average of the measured nonequilibrium work, W:

$$\Delta G_{\text{eqm}} = -k_{\text{B}}T \ln \left\langle \exp\left(-\frac{W}{k_{\text{B}}T}\right)\right\rangle \tag{12.6}$$

Crooks' theorem relates the work distributions for the unfolding and refolding reactions, $P_{\text{U}}(W)$ and $P_{\text{R}}(-W)$, respectively, to the energy dissipated:

$$\frac{P_{\text{U}}(W)}{P_{\text{R}}(-W)} = \exp\left(\frac{W - \Delta G_{\text{eqm}}}{k_{\text{B}}T}\right) \tag{12.7}$$

This leads to the very simple result that the equilibrium free energy is given by the crossing point of $P_{\text{U}}(W)$ and $P_{\text{R}}(-W)$ (Figure 12.6b). There is still some debate over how generally valid these theorems are (Vilar and Rubi 2008), and there are practical difficulties with implementing them far from equilibrium (Gore et al. 2003). Nevertheless, these theorems provide a very useful tool for folding studies because it is often difficult to study folding transitions at equilibrium due to intrinsically slow kinetics.

Constant-Force Extension Trajectories

The second type of measurement records the molecular extension as a function of time while the molecule is held under constant tension. The extension jumps in value as structural features fold or unfold, as illustrated in Figure 12.7A for the folding of the same riboswitch aptamer discussed previously. Different values of extension correspond to different structural states of the molecule. The extension change for each folding transition can be measured as the distance between peaks in histogram of the records (Figure 12.7B). Assuming the elastic properties of the polymer are known, the extension change at a given force may then be

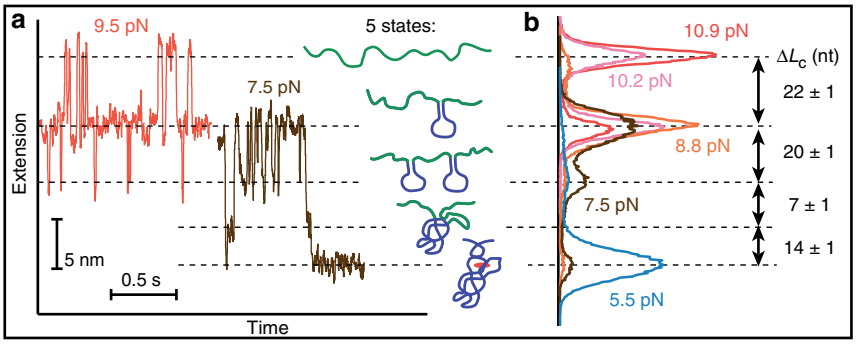


Figure 12.7. Constant-force measurements of folding in an adenine riboswitch aptamer. **A.** Extension records at two different forces reveal five different states. **B.** Histograms of the extension records at various forces, scaled by the force-dependent extension per nucleotide, reveal the number of nucleotides involved in each transition and allow the structural elements involved to be identified. [Adapted from Greenleaf et al. (2008).]

converted into a contour length change and hence the number of monomers (e.g., nucleotides or amino acids) involved in the folding determined. The folding transitions may thus be correlated with possible structures (Figure 12.7).

Quantities other than extension changes can also be determined. The interpretation of the data is often more direct than for FECs because the force is kept constant during the measurement and the system remains in equilibrium. The equilibrium free energy change between two states is simply the product of the extension change and the applied force at the force at which each state is equally likely to be populated (this includes the energy for stretching the unfolded molecule). The distribution of lifetimes of the states can also be determined directly. The rates for the highest-force transition shown in Figure 12.7A are plotted as a function of force in Figure 12.8. The force dependence of the rates is particularly informative, since

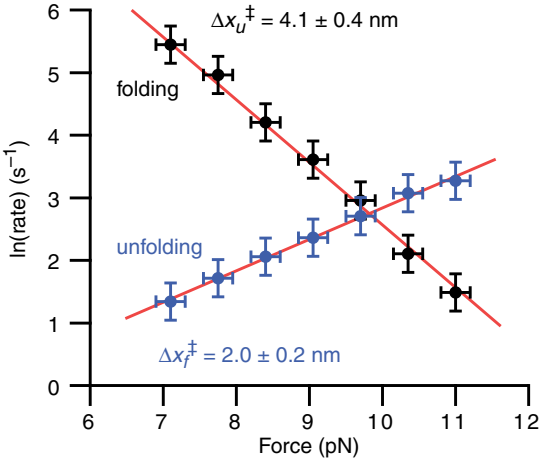


Figure 12.8. Force-dependent kinetics from constant-force measurements. The rates for folding (*black*) and unfolding (*blue*) of the highest-force, two-state folding transition in Figure 12.7A (folding of hairpin P2, inset Figure 12.5A) vary exponentially with force. The slopes of the fits yield the distances to the transition state.

the kinetics can be extrapolated to zero force (allowing comparisons to other experimental results), and the variation of the lifetimes with force yields the location of the transition state.

These interpretations are, however, model dependent. The simplest interpretation, due originally to Bell (1978), assumes that the transition-state location does not depend on force, and takes the rates for a two-state reaction as

$$k_u(F) = k_{0,u} \exp\left(\frac{\Delta x_u^{\ddagger}}{k_B T} F\right), \; k_f(F) = k_{0,f} \exp\left(-\frac{\Delta x_f^{\ddagger}}{k_B T} F\right), \tag{12.8}$$

where $k_{0,u}$ and $k_{0,f}$ are the rates for unfolding and folding at zero force, respectively, and $\Delta x^{\ddagger}{}_u$ and $\Delta x^{\ddagger}{}_f$ are the distances to the transition state from the unfolded and folded states, respectively. This approximation is reasonable in the limit of a small force range and a sharp barrier because the transition state will shift by $\sim \Delta F/\kappa$ (where ΔF is the range of forces probed and κ is the curvature of the barrier). In Figure 12.8, $\ln(k)$ varies linearly with F, as expected from Eq. (12.8); curvature in this plot would indicate that a more complex model is needed to include the change of transition-state location with force (Mañosas et al. 2006). The sum of $\Delta x^{\ddagger}{}_u$ and $\Delta x^{\ddagger}{}_f$ should be the same as the total extension change for the transition, otherwise the assumption of a two-state transition is likely invalid.

Constant-force measurements have additional advantages over FECs. States that are particularly unstable or have very short lifetimes are much easier to observe because not only is the time and spatial resolution typically higher in constant-force measurements, but also there is less of a tendency for unusually stable structures or interactions to protect less stable ones from unfolding, as happens in FECs. This aspect is nicely illustrated by Figures 12.6, 12.7 and 12.8: Only one partially folded intermediate state is observed in the FECs of the aptamer, but three are seen at constant force. Finally, constant-force extension measurements allow a direct determination of the shape of the folding energy landscape all along the reaction coordinate. By carefully measuring the probability of occupancy for every value of the extension, $P(x)$, on can find the free energy profile $G(x)$ from an inverse Boltzmann transform: $G(x) = -k_B T \ln[P(x)]$ (Woodside et al. 2006b). An indirect method for determining free energy profiles from FECs using Jarzynksi's equality has been proposed (Hummer and Szabo 2005) but not yet tested quantitatively.

It may not always be possible to measure folding at constant force, however, due to kinetics that are either slower than the accessible measuring times or faster than the force-clamp feedback loop. In these cases, variants of the constant-force extension measurement can be helpful. The "force-jump" technique allows force-dependent rates to be measured for a particular direction of the reaction (folding or unfolding) when the rates are very slow. The force is changed abruptly, up (to unfold the molecule) or down (to refold it), to a preset value, where it is clamped. The lifetime for the molecule to unfold/refold is then measured directly at that force. If the molecule folds too quickly for an active feedback loop to be used and a passive force clamp cannot be implemented, then "unclamped" measurements are a practical alternative. Here, the extension is measured as a function of time, but the force is allowed to change in tandem with the extension, as in FECs. Because a given state has a constant extension, the force applied to each state is a constant for that state (although different for each state). Hence, lifetimes can be still determined at constant force, but the force acting on each state must be monitored separately. No difference in force-dependent kinetics has been found for constant-force, force-jump, or unclamped measurements (Woodside et al. 2006a; Li et al. 2006).

One essential limitation of force spectroscopy of folding relates to the issue of how to assign a particular structure to a given folding transition. The end-to-end distance provides strong clues, through the number of monomers involved in the transition, but several structures may be degenerate in extension. This is especially problematic in large, complex molecules with many structural elements, but it can occur even in simple structures. When this is a problem, it may be necessary to use additional probes to distinguish the possible structures. Examples include mutagenesis or ligand binding (e.g., antisense oligomers) to block the formation of specific structural elements or interactions or the incorporation of fluorescent probes.

Binding Reactions

Another type of measurement of structural and mechanical properties that can be made with optical traps focuses on the interactions between different molecules. Examples include the binding of ligands to RNA aptamers (Greenleaf et al. 2008), the binding of endonucleases to DNA (Koch et al. 2002), the disassembly of nucleosomes (Brower-Towland et al. 2002), and the binding of myosin to actin (Nishizaka et al. 1995) or kinesin to microtubules (Kawaguchi and Ishiwata 2001). The energies and kinetics for binding/unbinding reactions can be probed very similarly to those for folding transitions. FECs may be used to measure the distribution of unbinding forces (and hence the energetics and kinetics), or kinetics may be measured at constant force by force jumps. Unlike folding measurements, unbinding experiments are single shot; hence, to achieve good statistics, it is helpful to design the assay in such a way that the complex can be reformed easily or without too much delay.

12.2.2. Studies of Molecular Motors

Optical traps have also proven to be very useful in the study of motor proteins—mechanoenzymes that convert the energy from chemical reactions (such as adenosine triphosphate [ATP] hydrolysis) into motion and mechanical work. Frequently, motor proteins bind to and move along linear, polymeric "tracks" made of components such as DNA, RNA, or the cytoskeletal elements actin or microtubules. The motion of individual motor proteins can be observed directly in optical traps and the effects of physiologically relevant forces probed. Such measurements provide unique insights into key motor properties, including the effects of force on motor velocity, the maximum force a motor can overcome (stall force, F_{stall}), the step size of the motor, and the number of mechanical steps a motor completes before dissociation (run length, L). By directly observing the individual translocations of processive motor proteins, it is also possible to distinguish between forward and backward stepping, to discriminate between periods of pausing and processive motion, and to probe the heterogeneity of enzymatic response.

From a practical perspective, how measurements are made depends strongly on the typical run length, or processivity, of the motor. Nonprocessive motors produce only one mechanical stroke per binding event ($L = 1$), whereas processive motors take many sequential steps per binding event ($L >> 1$). Approaches for measuring and analyzing the motion of different types of motors are illustrated with examples from the kinesin family, the myosin family, and RNA polymerase.

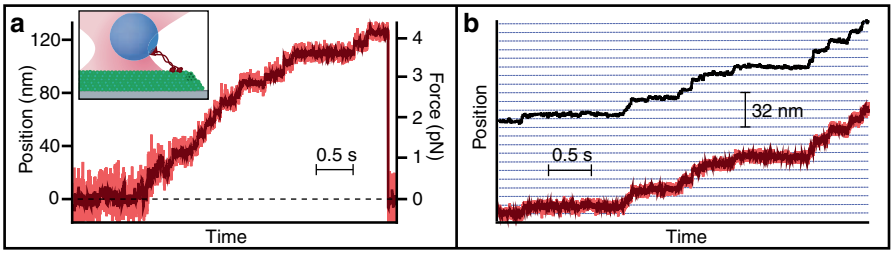


Figure 12.9. Stepping records of dimeric Eg5. **A.** A single Eg5 dimer (*red*) attached to a bead held in a fixed-position optical trap is placed over a microtubule (*green*). Individual steps are observed in the bead position (raw data, red; smoothed data, dark red) as the motor moves out of the trap. The force experienced by the motor increases with the bead displacement. **B.** A position record (*red*) with the force clamped reveals distinct 8-nm steps (*dotted blue lines*). The trap position (*black*) leads the bead position at a fixed offset, producing a constant assisting load. [Adapted from Valentine et al. (2006).]

Processive Mechanoenzymes

Single-Beam Optical Trap Measurements

Processive motor proteins are often measured with single motor protein bound to a bead held in a single-beam optical trap, which is placed next to the motor's "track" so that the motor can bind and move. The bead position is monitored as a function of time, giving an instant readout of motor displacement, as illustrated in Figure 12.9 for dimeric Eg5, a mitotic motor related to kinesin. Discrete steps separated by variable-length pauses can be clearly observed. This geometry is particularly useful for studying classical motor proteins, such as kinesins and dynein moving along microtubules or myosins moving along actin, as well as enzymes such as polymerases and helicases that move processively along nucleic acids (Rock et al. 2000; Kimura and Bianco 2006).

If the position of the optical trap is fixed, the force experienced by the motor increases incrementally with each step out of the trap (Figure 12.9A). The pause duration between steps increases as the motor works against a larger force, until the maximal force the motor can sustain is reached and the motor stalls or detaches from the track. The increased stiffness of the bead–motor assembly at higher forces also causes the noise on the signal to decrease as the bead moves out of the trap, allowing clean step transitions to be observed at high force. However, step sizes must be corrected for the compliance of the bead–motor linkage, which reduces the measured distances (Svoboda and Block 1994b).

Some of the most interesting results of such measurements are the dependence of pause duration or motor velocity on the force and buffer conditions (e.g., ATP concentration) because these give insight into the mechanochemical reactions that drive motion. However, the change in the load experienced by the motor at each step makes it challenging to extract force-dependence information using fixed-position traps. These difficulties can be avoided by implementing a force clamp (Visscher and Block 1998) to maintain a constant force during multiple stepping cycles (Figure 12.9B). Both the magnitude and the direction of the applied force can be controlled independently (Lang et al. 2002). The constant force eliminates any need for compliance corrections to the measured step sizes and allows good statistics to be collected for the force-dependent pause duration. Moreover, by using high trap stiffness the thermal fluctuations of the bead can be minimized, allowing the steps to be observed more clearly.

Kinetic Modelling of Force-Dependent Biochemistry

Force probes are particularly appropriate for studying mechanochemical processes like motor protein motion because force selectively affects the transitions involving motion. Characterizing these transitions is essential to understanding how motor proteins work, but they are difficult to probe using other methods. By measuring velocity as a function of force and ATP concentration (for an ATP-dependent motor), it is possible to determine the coupling of the motor's biochemical and mechanical cycles, to model kinetic pathways, and to construct energy landscapes (Bustamante et al. 2004).

Mechanical steps are commonly modelled as a transition over a free energy barrier, where the time to transition rises exponentially with the applied load (Schnitzer et al. 2000) following an Arrhenius–Boltzmann relation:

$$k(F) = k_{\mathrm{U}} + k_{\mathrm{F}} \exp\left(-\frac{F\delta}{k_{\mathrm{B}}T}\right) \tag{12.9}$$

where k_{U} includes the rates for all force-independent transitions, k_{F} is the scaling prefactor for the force-dependent rates, and δ is the distance along the reaction coordinate to the mechanical transition state. At zero force, the unloaded rate $k_0 = k_{\mathrm{U}} + k_{\mathrm{F}}$. In some cases, δ has been interpreted as an actual distance along the motor's track, but in general, the physical meaning of δ is not well defined. The load dependence of motor velocity is illustrated in Figure 12.10A for conventional kinesin (a cargo transporter), as determined from force-clamped records. When "hindering" loads (in the direction opposing forward stepping) are applied to kinesin, the velocity slows as the force-dependent translocation step becomes slow enough to be rate determining. When assisting loads are applied, however, the velocity does not increase, indicating that other biochemical transitions are rate determining in this regime.

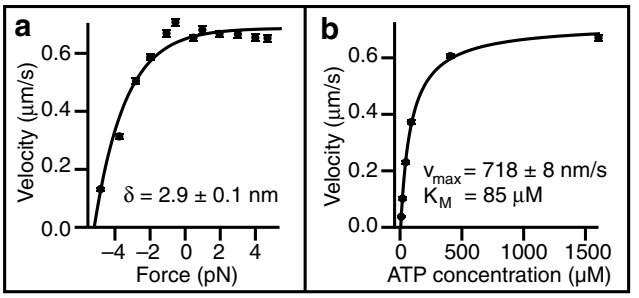


Figure 12.10. Force and adenosine triphosphate (ATP) dependence of conventional kinesin motion. **A.** The velocity as a function of force, obtained from force-clamped records at saturating ATP concentration, is fit by Eq. (12.9) to obtain kinetic parameters. **B.** The velocity as a function of ATP concentration under zero load is fit to the Michaelis–Menten model, Eq. (12.10). [Adapted from Block et al. (2003).]

The motor velocity also depends on the concentration of its enzymatic substrate (e.g., ATP, in the case of many motors). At any given applied force, the velocity typically displays Michaelis–Menten behavior:

$$v = \frac{v_{\max}[\mathrm{ATP}]}{[\mathrm{ATP}] + K_{\mathrm{M}}} \tag{12.10}$$

where v_{max} is the maximal velocity, [ATP] is the ATP concentration, and K_M is the Michaelis constant. Generically, the forward stepping rate is limited by ATP binding at low ATP concentrations, and hence increases proportionately with ATP concentration, whereas it asymptotes to v_{max} at saturating ATP levels. This is precisely the behavior seen for kinesin (Figure 12.10B).

For characterizing the mechanochemical cycle of the motor, v_{max} and K_M can be expressed more usefully in terms of the maximal catalysis rate of the protein, k_{cat}; the rate of productive ATP binding, k_b; and the step size of the motor, d: $v_{max} = dk_{cat}$ and $K_M = k_{cat}/k_b$. The term k_b incorporates the rates of all transitions up to and including the first irreversible transition, and k_{cat} includes all rates after ATP binding. The rates of some transitions may contribute to both k_b and k_{cat}. By determining the force- dependence of k_b and k_{cat}, one can find the location of the force-generating transition(s) in the biochemical cycle. In the example of kinesin subject to hindering loads, v_{max} decreases but K_M increases with increasing force (Visscher et al. 1999; Block et al. 2003), indicating that both k_{cat} and k_b are slowed under hindering load but k_b decreases more quickly at large forces. Hindering load thus has two effects on the mechanochemical cycle: It lowers the maximal stepping rate and lowers the rate of productive ATP binding. Loads applied transverse to the direction of motion also have an effect, but a much smaller one, lowering k_{cat} slightly without affecting k_b; a slight left–right asymmetry is also observed (Block et al. 2003). Taken together, these results suggest a model in which the primary working stroke occurs with or just after ATP binding and is aligned with the microtubule axis, whereas a smaller, asymmetric transverse motion occurs later in the cycle.

Complementary insight into the mechanochemistry can be obtained by analyzing the fluctuations in the enzyme activity (Shaevitz et al. 2005). Temporal irregularity in the motion can be characterized by the randomness, r, defined in terms of ensemble averages of the displacement $x(t)$ as

$$r = \lim_{t\to\infty} \frac{\langle x^2(t)\rangle - \langle x(t)\rangle^2}{d \cdot \langle x(t)\rangle} \tag{12.11}$$

where d is the motor step size. The randomness $r = 0$ for a perfectly clock-like enzyme, whereas $r = 1$ for a Poisson stepper with a single rate-limiting transition. In general, $r = 1/N$ for an enzyme with N slow, rate-determining transitions. Calculating the enzyme randomness under limiting and saturating ATP concentrations, as well as under various load conditions, provides insight into the minimal number of rate-determining transitions that should be included in modelling the motor mechanochemistry. Note that calculations of randomness do not require the observation of clear step transitions and are not sensitive to sources of stationary noise, although the presence of backward steps, inactivated states, or futile hydrolyses can increase r beyond unity.

Dual-Trap Measurements of Slow, Highly Processive Motors

Although single-beam optical trapping measurements are useful for measuring the mechanical and biochemical properties of many enzymes, the physical connection of the trapped bead to the coverglass surface via a protein or DNA linkage couples mechanical noise directly into the displacement record. This is particularly detrimental to measurements of slow, highly processive motors because data collection times may be many minutes,

substantially longer than the time over which drift occurs. For such motors, which include many nucleic acid enzymes, the more stable dual-trap dumbbell assay (Figure 12.2D), which eliminates mechanical connections to the microscope, is helpful.

Dual-trap assays have been particularly useful for studying transcriptional elongation by RNA polymerase (RNAP), a large protein complex that moves along a DNA template to synthesize RNA (Herbert et al. 2008). The resolution afforded by the dumbbell geometry has allowed direct observation of the very small (3.4 Å) steps made by RNAP (Abbondanzieri et al. 2005) and clear discrimination between slow elongation and transcriptional pausing, as shown in Figure 12.11. The reduction of mechanical drift also provides additional opportunities for analyzing the enzyme motion. Momentary reversals in direction have been identified and correlated with the correction of misincorporated nucleotides (Shaevitz et al. 2003), and the alignment of displacement records from multiple experiments has allowed the influence of the DNA sequence on pausing and elongation to be probed (Herbert et al. 2006). Moreover, the dumbbell assay can be designed to apply constant force to either the DNA template or the nascent RNA transcript, allowing the effects of RNA secondary structure on elongation and termination to be determined (Dalal et al. 2006; Larson et al. 2008).

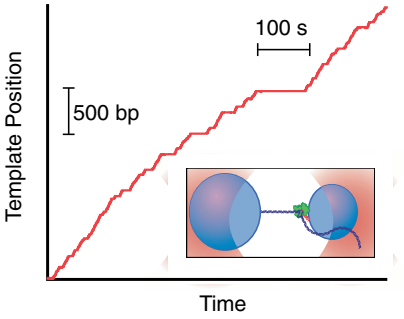


Figure 12.11. RNA polymerase (RNAP) transcription under constant load. A dumbbell is formed from a single RNAP molecule (*green*) loaded onto its DNA template (*blue*), and the length of the dumbbell is measured as the RNAP moves along the DNA, polymerizing the RNA transcript (*red*). Distinct periods of pausing are observed during transcription elongation. [Adapted from Herbert et al. (2006).]

Three-Bead Optical Trapping for Nonprocessive Motors

Nonprocessive mechanoenzymes, such as muscle myosin (myosin-II) or the kinesin-related protein ncd, move actin or microtubules, respectively, through the collective action of an ensemble. Measurements of the mechanical and biochemical properties of nonprocessive motors are often made using a different assay, such as the three-bead assay shown in Figure 12.12A. Here, the cytoskeletal filament is attached to two beads held in separate optical traps in a dumbbell geometry, whereas motor proteins are attached at low density to a platform (often a third particle) bonded to the coverglass surface. The suspended filament is held near the platform, allowing surface-attached motor proteins to interact with the filament transiently (Finer et al. 1994). Typically, the optical traps have low stiffness, allowing the suspended filament to undergo large thermal fluctuations. When the motor binds, the system stiffness increases significantly, leading to a decrease in the variance of bead position, as shown for a nonprocessive, monomeric myosin-VI motor in Figure 12.12B. Stochastic binding and detachment events can thus be accurately identified by monitoring bead variance as a function

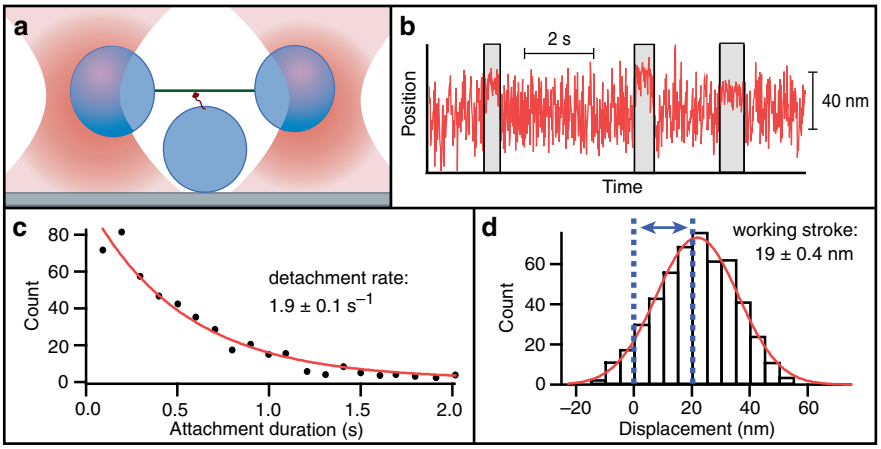


Figure 12.12. Three-bead assay for nonprocessive motion of myosin-VI monomers. **A.** A single actin filament (*green*) is tethered between two beads held in separate optical traps and placed near a third, surface-bound bead coated with low densities of motor protein (*red*). **B.** Motor binding events (*grey boxes*) are identified by the reduction in the Brownian fluctuations of the filament due to the increase in system stiffness on motor binding. **C.** The distribution of attachment times is fit by a single exponential. **D.** The distribution of displacements is Gaussian, offset from zero by the working stroke. [Adapted from Lister et al. (2004).]

of time (Molloy et al. 1995) and the distribution of attachment lifetimes (and its dependence on ATP concentration) analyzed. Biochemical rate constants can also be extracted: For example, detachment rates can be determined from the distribution of attachment times (Figure 12.12C).

The small distance the filament moves each time a motor binds (the "working stroke") may also be measured, but this measurement is complicated by the large thermal fluctuations of the filament. Typically, the motor-driven working stroke is determined from the distribution of time-averaged bead positions for all events, identifying events via the change in bead displacement variance as in Figure 12.12B (Molloy et al. 1995). On the assumption that that motor attachment occurs with equal probability over the entire range of motion of the filament, the distribution of bead displacements is Gaussian, with a width determined by the random thermal noise and a peak position offset from zero by the working stroke distance, as shown in Figure 12.12D. The measured stroke distance is frequently an underestimate due to the compliant nature of the bead and filament linkages and suboptimal orientation of the motor with respect to the filament axis.

Direct measurements of force-dependent kinetics are more challenging for nonprocessive motors than for processive enzymes like kinesin or RNAP because motor interactions are transient and there is very little time in which to identify a binding event, apply load, and measure the response. A clever approach to increasing the temporal resolution of event finding is to apply a small-amplitude carrier wave to the position of one optical trap. Motor binding may then be observed as a decrease in the mean amplitude of the oscillatory signal transmitted to the bead in the stationary trap (Veigel et al. 2005). For a ~1-kHz carrier wave, binding events can be identified and force applied to the motor within ~3 ms. With this approach, attachment lifetimes can be determined as a function of both ATP concentration and load, enabling more complete modelling of enzyme mechanochemistry.

In practice, it is easier to implement the three-bead assay for measurements of actin-based motors than for those that bind microtubules (deCastro et al. 2000). Because it is difficult to achieve end-specific attachments, filaments are attached to the beads via a side-binding interaction. The typical bead–bead separations are several microns, similar to the persistence length of an actin filament ($\sim$10 μm) but orders of magnitude smaller than the persistence length of a microtubule ($\sim$1 mm). An actin filament, because it is relatively flexible, remains roughly straight under tension, whereas a microtubule bends into a long arc between the two beads. This increased curvature may influence binding efficiencies and the observed value of the working stroke. In addition, the overall system stiffness is much higher in the case of microtubules, so traps with very low stiffness are required for variance-based discrimination between bound and unbound states.

12.3. Practical Experimental Considerations

State-of-the-art optical trapping measurements have become sufficiently complex that undertaking the study of a novel biomolecular system can seem daunting. Here we outline some of the practical considerations that may be helpful for designing optical trapping experiments and verifying the interpretation of results.

12.3.1. Ensuring "Real" Signals

There are many possible sources of spurious or misleading signals. For example, the optical mode in the trapping or detection lasers may change spontaneously, causing the trapped molecule to appear to move. Small particulate contaminants in solution may be trapped by the laser beam, causing what appear to be transient motions. Molecules may be damaged or their functionality impaired during preparation or measurement, they may interact nonspecifically with surfaces, or they may attach in groups (rather than singly) to the beads. Some of these effects may be mitigated through continuous monitoring of the trap and solution quality (e.g., measuring changes in the power spectrum or variance of the trap, filtering contaminants from buffers); others require separate controls.

Ensuring the Quality of the Molecules Measured

Controls should, of course, be run to test that the presence of surfaces and handle or bead attachments does not unduly change the behavior of the molecules being measured. Molecules may also be "damaged" during their production (e.g., by radiation, oxidation, or incorrect refolding), providing another motivation for controlled tests of their behavior. During optical trapping measurements, the laser light can damage molecules through oxidation due to oxygen radicals created by photodissociation. This effect can be minimized by using an oxygen-scavenging system to scour oxygen radicals from solution. One such system is a combination of glucose oxidase, catalase, and glucose, but others exist (Rasnik et al. 2006).

Working in the Single-Molecule Regime

For every measurement, it is critical to verify that only a single molecule is being probed at a time because the presence of multiple molecules can produce baffling results. For systems using nucleic acid tethers or handles, this is usually done with FEC measurements of the

persistence length (see section Elastic Properties of DNA), which is well known for single nucleic acids. Tethers consisting of multiple molecules tend to have a lower persistence length than expected; short contour lengths are an additional clue. For systems that do not have such tethers, as with many molecular motors, other methods must be used to ensure that measurements are in the single-molecule regime. One such way is to measure the fraction of beads that move along their substrate as a function of the motor concentration. The resultant distribution is well described by the Poisson distribution. The probability that the bead carries at least N motors is

$$P_N(c) = 1 - \sum_{n=0}^{N-1} (\lambda c)^n \frac{\exp(\lambda c)}{n!} \tag{12.12}$$

where c is the motor concentration and λ is a fitting parameter related to the binding efficiency of the motors to the beads, as shown in Figure 12.13 (Svoboda and Block 1994b). For processive enzymes, a single motor is sufficient to move a bead; for nonprocessive motors, a larger cooperative ensemble is needed, leading to a steeper transition. From such curves, the probability of engaging more than one motor can be determined quantitatively, given the percentage of moving beads. Ideally, one would work with vanishingly small numbers of motors. In practice, working in a regime where fewer than one third of all beads move or bind is a reasonable compromise, ensuring that a single motor is being measured ~90% of the time.

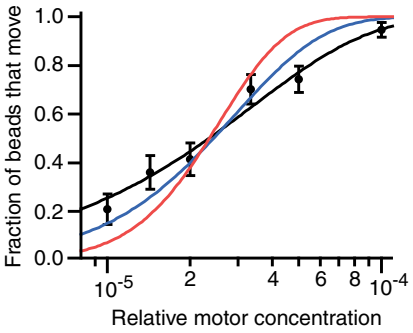


Figure 12.13. Poisson statistics used to determine processivity. The fraction of beads that move as a function of motor concentration for Eg5 dimers is fit well by the Poisson distribution (*black*) with $N = 1$ in Eq. (12.12), as expected for a processive motor. Curves for $N = 2$ (*blue*) and $N = 4$ (*red*), indicating cooperative motion, are steeper than actually observed. [Adapted from Valentine et al. (2006).]

Identifying Signals

To ensure that the signals being measured are truly due to the molecule of interest, controls should always be performed repeating the measurement in the absence of the molecule, in the absence of ligands or enzyme substrates, and so on, as appropriate. This is true even when the signals observed make good sense in terms of the expected behavior of the molecule. Repeating the assay with a molecule whose properties are well known (e.g., in a folding measurement, a hairpin of known length and rupture force) can also be helpful in establishing the soundness of the assay and the calibration of the measurement.

Distinguishing signals from noise is often difficult, especially when working with small motions or at low force, where Brownian motion of the beads can be significant. Binding events are particularly difficult to identify because it is often challenging to distinguish between specific and nonspecific binding, especially when surfaces are involved. For binding studies, it may be necessary to repeat the measurements with and without the complete unbinding complex and under a range of assay conditions (including buffers of different ionic strengths and various surface chemistries) to verify that functional binding can be resolved.

Processive motions are generally easier to identify because the motion is directional and thus cannot arise from diffusion alone. Individual steps may be distinguished by eye or by using various step-finding algorithms (Carter et al. 2008). Sometimes, the motion may be amplified by taking advantage of the assay geometry. For example, stiff lever arms (e.g., of dsDNA or actin) may be used to magnify motions. If motions of an enzyme are associated with structural changes in its substrate, these structural changes may make the motion easier to observe: for example, single-base-pair steps (0.34 nm) taken by a helicase unwinding a RNA duplex may be turned into ∼1-nm steps by holding a duplex hairpin under tension in the trap (Dumont et al. 2006); similarly, rotations of dsDNA may be observed through the large motions associated with plectoneme formation.

12.3.2. Sources of Error

Understanding the sources of error is important to take best advantage of the quantitative nature of optical trapping experiments. Statistical sampling errors are, as usual, estimated using Poisson or Gaussian statistics and minimized by repeated measurements. Systematic errors are more problematic, and there are several common sources of systematic error that are often overlooked. Some of these are systematic across measurements of a given molecule but can be converted to random errors by averaging over a sufficiently large number of molecules; others are systematic to all measurements. In both cases, however, these errors must be known to make quantitative comparisons of different measurements.

Errors from Bead-Size Polydispersity

Variation in the size of the beads used in the trapping experiments is a major source of systematic error when measuring a given molecule because the bead size strongly affects the force applied by the trap (Section 12.1.1) and the position calibration. Whenever possible, position and stiffness calibrations should be made on each bead that will be measured to minimize the errors that changes in bead size may introduce (typically at least 5%–10%). It is best to use beads with manufacturer-specified size variance and if possible to use beads from a single manufacturer lot for the entire course of experiments.

Errors in Stiffness Calibration

Even when multiple methods are used to calibrate the trap stiffness (Section 12.1.3), it is difficult to determine stiffness to better than ∼5%–10% (the typical discrepancy between different calibration methods). Different calibration methods have different systematic errors. The variance method underestimates the stiffness in the presence of vibrational, electronic, or optical noise (which all increase the apparent Brownian fluctuations) but overestimates it if fluctuations are filtered excessively (e.g., by the natural bandwidth of the detector or

amplifiers). At high stiffness, the latter effect dominates. The variance method also relies on a position calibration, which is affected by bead size.

The power spectrum method does not depend on position calibration, but it does depend on the drag coefficient, which varies with both bead size and height above the surface. It, too, is sensitive to mechanical noise, electronic noise, and instrumental filtering, which can all change the shape of the power spectrum. The power spectrum is also sensitive to misalignment of the optics and optical noise (e.g., pointing fluctuations). The practical effect of these noise sources is usually to reduce the apparent roll-off frequency of the Lorentzian fit, thereby underestimating the stiffness.

The viscous drag method tends to produce a less biased result than both other methods and provides essential information about the value of the stiffness as a function of the displacement from the centre of the trap. However, it is more time consuming, requiring a good position calibration, knowledge of the drag coefficient, and multiple measurements. It must also typically be performed in a separate experiment (e.g., very dilute solutions of beads must be used to avoid having multiple beads fall into the trap during the motion), and thus is less useful for calibration *in situ* before a measurement.

To minimize errors in the stiffness, calibrations should be repeated with all three methods on many different beads to average over bead size variations and should be redone frequently to track changes in the properties of the trap.

Errors from Trap Potential Anharmonicity

The trap potential is only approximately harmonic for small bead displacements. The size of the harmonic region depends on the details of the instrument geometry and the bead size, but is usually considered to extend ∼50–100 nm from the trap centre. It can be measured using the viscous drag method for stiffness calibration as the region over which force is linear with displacement (Figure 12.3). Trap potential anharmonicity is a significant error that is often underappreciated or even ignored. For small displacements (<50 nm), the trap stiffness determined as before can be used with reasonable accuracy. For larger displacements, however, a displacement-dependent stiffness is usually required for quantitative measurements: At a displacement of 100 nm, the effective stiffness can be reduced by ∼25% from the nominal stiffness at the trap centre (Greenleaf et al. 2005).

Errors in Determining the Bead Height above a Surface

The height of a bead above a nearby surface such as a coverglass affects optical trapping measurements in several ways. Position calibrations are sensitive to the height of the bead through changes in the light scattering, and the drag coefficient β of the bead (necessary for stiffness calibrations) is affected by boundary conditions at the surface (Svoboda and Block 1994a). The stiffness of the trap also depends on the height of the bead above the surface when using oil-immersion lenses with short working distances.

Measuring the axial position in the trap is more difficult than the lateral position and generally less sensitive, but it can be done—for example, by monitoring the interference between the light scattered by the bead and the unscattered light. The bead height can thus be found by using a motorized stage to bring the surface of the coverglass up to the trapped particle and determining the vertical position at which the bead is displaced from the trap centre (Neuman and Block 2004). Such measurements are accurate to only ∼10–100 nm;

this uncertainty must be considered in overall error determination, especially if the position and/or stiffness calibrations are made at a very different height from a surface than the actual measurement (as is common in motor protein measurements). To minimize these problems, it is best to measure and calibrate at the same height, whenever possible. An additional complication in the determination of bead height is that the vertical motion of the laser focal point is not the same as the motion of the stage, due to index mismatches (Neuman et al. 2005).

12.3.3. Comparing Optical Trapping Results to Results from Other Methods

One of the most challenging aspects of interpreting optical trapping measurements is correlating the results with those obtained using other techniques. Some of these challenges result from the particular details of the optical trapping assays. For example, the beads or other surfaces present may interact with the molecules, changing their behavior. Long handles used to attach to the molecules may introduce compliance corrections for distances, elastic corrections for energies, and hydrodynamic drag corrections for rates. Finite sampling rates, filtering effects, and slow feedback-loop closure times may also affect measurements of rates (Mañosas et al. 2007). Determining the quantitative impact of these sorts of effects is nontrivial but generally straightforward.

The deeper challenges relate more generally to the differences between the effects on the molecule of force, as compared to the effects of other, more traditional, probes. Different methods are also often sensitive to different aspects of the behavior. In folding studies, for example, the pathways probed with force-induced denaturation are not necessarily the same as those probed with chemical- or heat-induced denaturation (see section Folding Studies). Although folding rates often do agree among different types of measurement (see, e.g., Greenleaf et al. 2008), disagreement is not necessarily an indication of problems with the measurements: it may instead indicate a different (and possibly interesting) class of behavior when the molecule is held under tension.

Measurements of motor proteins pose similar challenges. Optical trapping measurements probe the effect of force on enzyme motion but are generally insensitive to the underlying biochemical reactions, whereas many biochemical studies probe the kinetics of substrate binding, hydrolysis, and product release via stopped flow but are insensitive to the effects of force (Gilbert and Mackey 2000). Moreover, whereas optical traps can resolve multiple steps in processive motion, it is often difficult to register the first step because a single translocation event is hard to distinguish from the Brownian noise. In contrast, the signals from bulk biochemical probes generally dephase after one or two translocations due to asynchronous motion in ensembles of motors. Optical trapping measurements thus often provide important complementary information, but side-by-side comparisons of data are not always possible or meaningful.

12.4. Extending the Capabilities of Optical Traps

Optical traps are powerful tools for studying single biological molecules, but there are many properties that cannot be probed simply by using traditional traps. To further expand the capabilities of optical traps, novel instrumentation continues to be developed. One new technique allows the application of controlled torque in addition to force, an advance that can be very useful for systems having both linear and rotational components to their motion. This is achieved by trapping particles that, unlike the beads normally used, are optically birefringent

by virtue of either an anisotropic refractive index or shape (La Porta and Wang 2004). Such particles are trapped more strongly along one axis (the slow optical axis or the long axis of the particle) and hence will align with the polarization of the trapping beam. Rotation of the trap polarization then exerts torque on the particle, which can be detected from the change in angular momentum of the light scattering by the particle; constant torque may be maintained with a torque clamp, analogous to a force clamp.

Other development efforts have concentrated on combining optical traps with other useful tools. Optical traps have been combined with microfluidic systems (Enger et al. 2004), opening new avenues for dynamic control over buffer composition. Traps have also been combined with nanopores (Trapagnier et al. 2007), allowing new control over translocation measurements with applications to folding studies and sequencing efforts. There has been a particular focus on combining optical trapping and single-molecule fluorescence modalities because of the utility of fluorescence for probing reactions that are difficult or impossible to detect with an optical trap alone. Such reactions include any biochemical reactions that do not depend on force and conformational rearrangements or binding events that do not change the length of the molecule. The complementary capabilities of trapping and fluorescence have already been used to study the relative timing of ATPase activity and motion in myosin (Ishijima et al. 1998), as well as to probe conformational changes in simple DNA duplexes (Lang et al. 2004) and in Holliday junctions (Hohng et al. 2007). Combinations of force and fluorescence should prove especially useful as increasingly complex systems are explored, building up from single components to the macromolecular complexes that carry out essential cellular functions. These and other developments hold the promise to reveal new and exciting insights into biology.

Acknowledgments

We thank Cuauhtémoc García-García for critical reading of the manuscript and Charles Asbury, Joshua Shaevitz, and Kristina Herbert for providing the raw data used in Figures 12.10 and 12.11. MTV is supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund.

References

Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, et al. (2005). Direct observation of base-pair stepping by RNA polymerase. Nature 438:460–465.

Ashkin A, Dziedzic JM, Bjorkholm JE, Chu S (1986). Observation of a single-beam gradient force optical trap for dielectric particles. Optics Lett 11:288–290.

Bell GI (1978). Models for specific adhesion of cells to cells. Science 200:618–627.

Berg-Sørensen K, Flyvbjerg H (2004). Power spectrum analysis for optical tweezers. Rev Sci Instrum 75:594–612.

Best RB, Paci E, Hummer G, Dudko O (2008). Pulling direction as a reaction coordinate for the mechanical unfolding of single molecules. J Phys Chem B 112: 5968–5976.

Block SM (1998). Constructing optical tweezers. In: Spector DL, Goldman RD, Leinwand LA (eds.), *Cells: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Block SM, Asbury CL, Shaevitz JW, Lang MJ (2003). Probing the kinesin reaction cycle with a 2D optical force clamp. Proc Natl Acad Sci USA 100:2351–2356.

Borgia A, Williams PM, Clarke J (2008). Single-molecule studies of protein folding. Annu Rev Biochem 77:6.1–6.25.

Brower-Towland BD, Smith CL, et al. (2002). Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. Proc Natl Acad Sci USA 99:1960–1965.

Bustamante C, Chemla YR, Forde NR, Izhaky D (2004). Mechanical processes in biochemistry. Annu Rev Biochem 73:705–748.

Carter AR, King GM, Ulrich TA, et al. (2007). Stabilization of an optical microscope to 0.1 nm in three dimensions. Appl Opt 46:421–427.

Carter BC, Vershinin M, Gross SP (2008). A comparison of step-detection methods: How well can you do? Biophys J 94:306–319.

Cecconi C, Shank EA, Bustamante C, Marqusee S (2005). Direct observation of the three-state folding of a single protein molecule. Science 309:2057–2059.

Collin D, Ritort F, Jarzynski C, et al. (2005). Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. Nature 437:231–234.

Crooks GE (1999). Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys Rev E 60:2721–2726.

Dalal RV, Larson MH, Neuman KC, et al. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. Mol Cell 23:231–239.

Dame RT, Noom MC, Wuite GJ (2006). Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. Nature 444:387–390.

deCastro MJ, Fondecave RM, Clarke LA, et al. (2000). Working strokes by single molecules of the kinesin-related microtubule motor ncd. Nat Cell Biol 2:724–729.

Dessinges MN, Maier B, Zhang Y, et al. (2002). Stretching single stranded DNA, a model polyelectrolyte. Phys Rev Lett 89:248102.

Dietz H, Berkemeier F, Bertz M, Reif M (2006). Anisotropic deformation response of single protein molecules. Proc Natl Acad Sci USA 103:12724–12728.

Dudko O, Hummer G, Szabo A (2006). Intrinsic rates and activation free energies from single-molecule pulling experiments. Phys Rev Lett 96:108101.

Dumont S, Cheng W, Serebov V, et al. (2006). RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. Nature 439:105–108.

Enger J, Goksor M, Ramser K, et al. (2004). Optical tweezers applied to a microfluidic system. Lab Chip 4:196–200.

Evans E, Ritchie K (1997). Dynamic strength of molecular adhesion bonds. Biophys J 72:1541–1555.

Finer JT, Simmons RM, Spudich JA (1994). Single myosin mechanics: picoNewton forces and nanometre steps. Nature 368: 113–119.

Fordyce PM, Valentine MT, Block SM (2007). Advances in surface-based assays for single molecules. In: Selvin P, Ha T (eds.), *Single-molecule techniques: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 431–460.

Gilbert SP and Mackey AT (2000). Kinetics: A tool to study molecular motors. Methods 22:337–354.

Gore J, Ritort F, Bustamante C (2003). Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. Proc Natl Acad Sci USA 100:12564–12569.

Greenleaf WJ, Woodside MT, Abbondanzieri EA, Block SM (2005). Passive all-optical force clamp for high-resolution laser trapping. Phys Rev Lett 95:208102.

Greenleaf WJ, Woodside MT, Block SM (2007). High-resolution, single-molecule measurements of biomolecular motion. Annu Rev Biophys Biomol Struct 36:171–190.

Greenleaf WJ, Frieda KL, Foster DNA, et al. (2008). Direct observation of hierarchical folding in single riboswitch aptamers. Science 319:630–633.

Grier DG, Roichman Y (2006). Holographic optical trapping. Appl Opt 45:880–887.

Harada Y, Asakura T (1996). Radiation forces on a dielectric sphere in the Rayleigh scattering regime. Opt Commun 124:529–541.

Herbert KM, La Porta A, Wong BJ, et al. (2006). Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell 125:1084–1094.

Herbert KM, Greenleaf WJ, Block SM (2008). Single-molecule studies of RNA polymerase: motoring along. Annu Rev Biochem 77:149–176.

Hermanson GT (1996). *Bioconjugate techniques*. Academic Press, San Diego, CA.

Hohng S, Zhou R, Nahas MK, et al. (2007). Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the Holliday junction. Science 318:279–283.

Hummer G, Szabo A (2005). Free energy surfaces from single-molecule force spectroscopy. Acc Chem Res 38: 504–513.

Hyeon C, Thirumalai D (2008). Multiple probes are required to explore and control the rugged energy landscape of RNA hairpins. J Am Chem Soc 130:1538–1539.

Ishijima A, Kojima H, Funatsu T, et al. (1998). Simultaneous observation of individual ATPase and mechanical events by a single myosin molecule during interaction with actin. Cell 92:161–171.

Jarzynski C (1997). Non-equilibrium equality for free energy differences. Phys Rev Lett 78:2690–2693.
Kawaguchi K, Ishiwata S (2001). Nucleotide-dependent single- to double-headed binding of kinesin. Science 291:667–669.
Kimura Y, Bianco PR (2006). Single molecule studies of DNA binding proteins using optical tweezers. Analyst 131:868–874.
Koch SJ, Shundrovsky A, Jantzen BC, Wang MD (2002). Probing protein–DNA interactions by unzipping a single DNA double helix. Biophys J 83:1098–1105.
Lang MJ, Fordyce PM, Engh AM, et al. (2004). Simultaneous, coincident optical trapping and single molecule fluorescence. Nat Methods 1:133–139.
Lang MJ, Asbury CL, Shaevitz, JW, Block SM (2002). An automated two-dimensional optical force clamp for single molecule studies. Biophys J 83:491–501.
La Porta A, Wang MD (2004). Optical torque wrench: Angular trapping, rotation and torque detection using quartz microparticles. Phys Rev Lett 92: 190801.
Larson MH, Greenleaf WJ, Landick R, Block SM (2008). Applied force reveals mechanistic and energetic details of transcription termination. Cell 132:971–982.
Li PTX, Collin D, Smith SB, et al. (2006). Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. Biophys J 90:250–260.
Lister I, Schmitz S, Walker M, et al. (2004). A monomeric myosin VI with a large working stroke. EMBO J 23: 1729–1738.
Mañosas M, Collin D, Ritort F (2006). Force-dependent fragility in RNA hairpins. Phys Rev Lett 96:218301.
Mañosas M, Wen JD, Li PTX, et al. (2007). Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. Biophys J 92:3010–3021.
Marko JF, Siggia ED (1995). Stretching DNA. Macromolecules 28:8759–8770.
Moffitt JR, Chemla YR, Izhaky D, Bustamante C (2006). Differential detection of dual traps improves the spatial resolution of optical tweezers. Proc Natl Acad Sci USA 103:9006–9011.
Moffitt JR, Chemla YR, Smith SB, Bustamante C (2008). Recent advances in optical tweezers. Annu Rev Biochem. 77:19.1–19.24.
Molloy JE, Burns JE, Kendrick-Jones J, et al. (1995). Movement and force produced by a single myosin head. Nature 378:209–212.
Nambiar R, Gajraj A, Meiner JC (2004). All-optical constant-force laser tweezers. Biophys J 87:1972–1980.
Neuman KC, Chadd EH, Liou GF, et al. (1999). Characterization of photodamage to *Escherichia coli* in optical traps. Biophys J 77:2656–2863.
Neuman KC, Block SM (2004). Optical trapping. Rev Sci Instrum 75:2787–2809.
Neuman KC, Abbondanzieri EA, Block SM (2005). Measurement of the effective focal shift in an optical trap. Opt Lett 30:1318–1320.
Neuman KC, Nagy A (2008). Single-molecule force spectroscopy: Optical tweezers, magnetic tweezers and atomic force microscopy. Nat Methods 5:491–505.
Nishizaka T, Miyata H, Yoshikawa H, et al. (1995). Unbinding force of a single motor molecule of muscle using optical tweezers. Nature 337:251–254.
Nugent-Glandorf L, Perkins TT (2004). Measuring 0.1-nm motion in 1 ms in an optical microscope with differential back-focal-plane detection. Opt Lett 29:2611–2613.
Rasnik I, McKinney SA, Ha T (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. Nat Methods 3: 891–893.
Rock RS, Rief M, Mehta AD, Spudich JA (2000). *In vitro* assays of processive myosin motors. Methods 22:373–381.
Rohrbach A, Stelzer EHK (2001). Optical trapping of dielectric particles in arbitrary fields. J Opt Soc Am A 18: 839–853.
Schnitzer MJ, Visscher K, Block SM (2000). Mechanism of force production by single kinesin motors. Nat Cell Biol 2:718–723.
Seol Y, Li J, Nelson PC, et al. (2007). Elasticity of short DNA molecules: Theory and experiment for contour lengths of 0.6–7 micron. Biophys J 93:4360–4373.
Shaevitz JW, Block SM, Schnitzer MJ (2005). Statistical kinetics of macromolecular dynamics. Biophys J. 89: 2275–2285.
Shaevitz JW, Abbondanzieri EA, Landick R, Block SM (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature 426:684–687.
Smith SB, Cui Y, Bustamante C (1996). Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. Science 271: 795–799.
Smith SB, Cui Y, Bustamante C (2003). Optical-trap force transducer that operates by direct measurement of light momentum. Methods Enzymol 361:134–162.

Svoboda K, Block SM (1994a). Biological applications of optical forces. Annu Rev Biophys Biomol Struct 23: 247–285.

Svoboda K, Block SM (1994b). Force and velocity measured for single kinesin molecules. Cell 77: 773–784.

Tolić-Nørrelykke SF, Schäffer E, Howard J, et al. (2006). Calibration of optical tweezers with positional detection in the back focal plane. Rev Sci Instrum 77:103101.

Trapagnier EH, Radenovic A, Sivak D, et al. (2007). Controlling DNA capture and propagation through artificial nanopores. Nano Lett 7:2824–2830.

Valentine MT, Fordyce PM, Krzysiak TC, et al. (2006). Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads *in vitro*. Nat Cell Biol 8:470–476.

Veigel C, Schmitz S, Wang F, Sellers JR (2005). Load-dependent kinetics of myosin-V can explain its high processivity. Nat Cell Biol 7:861–869.

Vilar JMG, Rubi JM (2008). Failure of the work-Hamiltonian connection for free-energy calculations. Phys Rev Lett 100:020601.

Visscher K, Gross SP, Block SM (1996). Construction of multiple-beam optical traps with nanometer-resolution position sensing. IEEE J Sel Top Quant Electr 2:1066–1076.

Visscher K, Block SM (1998). Versatile optical traps with feedback control. Methods Enzymol 298: 460–489.

Visscher K, Schnitzer MJ, Block SM (1999). Single kinesin molecules studied with a molecular force clamp. Nature 400:184–189.

Wang MD, Yin H, Landick R, et al. (1997). Stretching DNA with optical tweezers. Biophys J 72:1335–1346.

Williams PM, Fowler SB, Best RB, et al. (2003). Hidden complexity on the mechanical properties of titin. Nature 422:446–449.

Woodside MT, Behnke-Parks WM, Larizadeh K, et al. (2006a). Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. Proc Natl Acad Sci USA 103:6190–6195.

Woodside MT, Anthony PC, Behnke-Parks WM, et al. (2006b). Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. Science 314:1001–1004.

Woodside MT, García-García C, Block SM (2008). Folding and unfolding single RNA molecules under tension. Curr Opin Chem Biol, vol. 12, pp. 640–646.

Wuite GJL, Smith SB, Young M, et al. (2000). Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. Nature 404:103–106.

Xu F, Ren K, Gouesbet G, et al. (2007). Generalized Lorenz-Mie theory for an arbitrarily oriented, located, and shaped beam scatted by a homogeneous spheroid. J Opt Soc Am A 24:119–131.

13

Magnetic Tweezers for Single-Molecule Experiments

I. D. Vilfan, J. Lipfert, D. A. Koster, S. G. Lemay, and N. H. Dekker

Abstract Over the last decade, single-molecule techniques have proven their wide applicability in the study of processive motor proteins and other enzymes, yielding insight into their kinetics and mechanochemistry. In the context of force spectroscopy of protein–nucleic acid interactions, optical tweezers, magnetic tweezers, and atomic force microscopy have made important contributions. Advantages of magnetic tweezers include particularly straightforward control of the supercoiled state of DNA, facile extension to parallel measurement of multiple molecules and to integration with fluorescence measurements, and the simplicity and robustness of the experimental configuration. This chapter reviews the principles behind magnetic tweezers and their experimental implementation and points out recent improvements. It also describes several types of experiments that can be performed using magnetic tweezers.

13.1. Introduction

Nucleic acids (i.e., DNA and RNA) can exist in different structural forms that perform various functions *in vivo*. For example, double-stranded DNA (dsDNA) is mostly responsible for the storage of genetic information. Alternatively, single-stranded DNA (ssDNA) acts as an intermediate in biologically relevant processes such as DNA replication and recombination, as well as in the storage of viral genetic information [1–3]. RNA, which is predominately present in a single-stranded form (ssRNA), takes part in processes as diverse as transcription, protein translation, and gene regulation, and also acts as a carrier of genetic information in ssRNA viruses [4]. Its double-stranded counterpart (dsRNA) is present as a biological intermediate (e.g., in RNA silencing) and as a carrier of genetic information in dsRNA viruses [5–7].

The different structural forms adopted by nucleic acids have a number of consequences. For instance, the predominately double-stranded nature of DNA leads to the structural complexity of supercoiling: If torsion is applied to dsDNA, the molecule, provided that it is torsionally constrained, can acquire additional twist or form looped structures. This additional

I. D. Vilfan, J. Lipfert, D. A. Koster, S.G. Lemay, and N. H. Dekker • Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1 2628 CJ Delft The Netherlands

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_13,

intertwining of the strands in dsDNA poses a number of topological problems during fundamental processes such as cell cycle progression and cell division [8]. In addition, the structural forms of nucleic acids are not necessarily static structures but are instead constantly being converted from one to the other by cellular signals such as protein binding or enzyme activity.

Of interest, the mechanical properties of the various structural forms of nucleic acids (e.g., persistence length) change in response to their interconversion or protein binding and enzyme activity. Magnetic tweezers (MT) can probe these mechanical properties very sensitively and thus enable real-time detection of changes in nucleic acid state [9]. In addition, MT provide a way to perform force-dependence measurements at forces as low as ∼10 fN, which is lower than what is possible with other single-molecule techniques. Furthermore, MTs have the ability to control the degree of supercoiling, making it a technique of choice for studying supercoiling-dependent processes such as prokaryotic DNA condensation and regulation of dsDNA supercoiling by topoisomerases.

Although MTs have predominantly been used as an experimental tool to investigate biological processes involving nucleic acids, they have also been successfully applied in studies of the mechanical properties of proteins [10–14], as well as of the mechanical response of cells [11,15–18]. This chapter discusses the MT experimental configuration, then describes current challenges from a nucleic acid perspective, and finally considers examples of MT studies of the mechanical properties of proteins.

13.2. Experimental Design of the Magnetic Tweezers

Figure 13.1 shows a schematic picture of a typical MT experimental configuration. At the heart of the experimental configuration is a flow cell: A nucleic acid is tethered in this flow cell between a glass surface and a paramagnetic bead (references to a "bead" in the remainder of the chapter imply this paramagnetic bead unless otherwise specified) by means of noncovalent bonds that can resist forces on the order of ∼100 pN [19–23]. Alternatively, the nucleic acid can also be suspended between a surface-immobilized nonmagnetic bead and a tethered bead [24,25]. Above the flow cell, a pair of permanent magnets or an electromagnet is suspended, exposing the flow cell to a magnetic field B [26]. The magnetic field exerts an upward stretching force F on the bead given by

$$\vec{\mathbf{F}} = \frac{1}{2}\vec{\nabla}\left(\vec{\mathbf{m}} \cdot \vec{\mathbf{B}}\right) \tag{13.1}$$

where $\vec{\mathbf{m}}$ is the induced magnetic moment of the bead in an external magnetic field, $\vec{\mathbf{B}}$. The bead and the nucleic acid experience a constant F because the length scale of the motion of the tethered bead (typically micrometers) is small compared to the length scale of changes in the magnetic field B (typically millimeters; Section 13.7). The force F can be decreased or increased by moving the permanent magnets either up or down as the local magnetic field gradient is decreased or increased, respectively. MTs can apply a wide range of forces between ∼10 fN and ∼100 pN. At forces less than ∼10 fN, the bead is very close to the surface, where the interactions with the surface become dominant over the applied force. On the other hand, at forces greater than ∼100 pN, the majority of bonds that are used to tether the molecule to the surface are not stable.

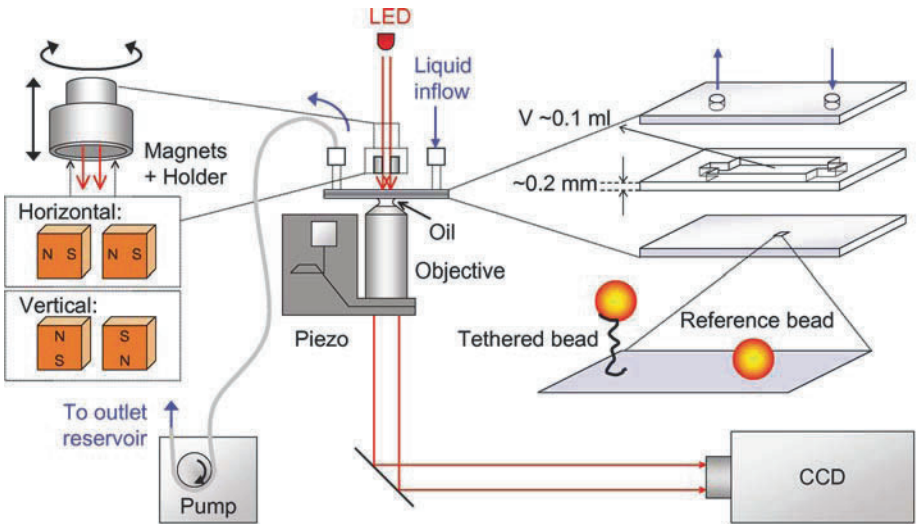


Figure 13.1. Schematic representation of a magnetic tweezers (MT) experimental configuration. A flow cell is built up from two microscope cover slips separated at the edges by a thin layer of parafilm approximately ~0.2 mm thin. In the top cover slip, two holes are drilled that are connected to an inlet and an outlet reservoir, allowing solutions to be flushed in and out of the cell using a pump. Polymers are attached to the bottom surface at one end and to a paramagnetic bead at the other end. A reference bead is located in the vicinity of the tethered paramagnetic bead and is fixed to the bottom cover slip surface. Above the cell, a parallel light source is located, as well as a pair of permanent magnets, that can be translated and rotated, making it possible to vary force and torque on the paramagnetic bead. There are two possible magnet configurations: horizontal and vertical. Below the bottom cover slip, an objective is located that is connected to a charge-coupled device (CCD) camera that records the image from the objective focal plane. A controller containing a piezoelectric crystal can adjust vertical position of the objective's focal plane. LED, light-emitting diode.

An external magnetic field will induce a major component of the magnetic moment to align in the direction of $\vec{\mathbf{B}}$. Due to a slight anisotropy in the magnetization of the bead, a minor component of the magnetic moment, $\vec{\mathbf{m}}_0$, is initially not aligned with $\vec{\mathbf{B}}$. This generates a torque, $\vec{\Gamma}$, on the bead:

$$\vec{\Gamma} = \vec{\mathbf{m}}_0 \times \vec{\mathbf{B}} \tag{13.2}$$

Rotating the magnets results in a rotation of the bead because the torque applied by the nucleic acid on the bead is negligible compared to the torque imposed by the magnetic field.

In the most common optical configuration, the bead is illuminated by parallel light placed above the flow cell and the magnet (Figure 13.1). The interference of the illuminating light with the light scattered by the bead produces concentric diffraction rings in the focal plane of the objective placed below the flow cell. The image of the diffraction pattern is recorded through an objective with a charge-coupled device (CCD) camera and analyzed with a computer (see Section 13.3). Furthermore, MT applications have been developed in which the nucleic acid is suspended between two beads rather than a bead and a glass surface [24,25]. In these applications, one bead is magnetic and the other is nonmagnetic and immobilized on a solid support, with the objective positioned perpendicular rather then parallel to the tethered nucleic acid [24,25].

13.3. Image Analysis

An MT experiment enables a real-time measurement of the vertical extension of the nucleic acid, l, as well as computation of the applied force, F. To extract these parameters from the acquired data, image analysis of the diffraction pattern recorded by the CCD is required. To monitor l, a method first designed by Gosse and Croquette is used, in which the objective's focal plane (OFP) is accurately shifted in the vertical direction while the diffraction rings of the bead attached to the nucleic acid are imaged (Figure 13.2A–D) [9]. In this manner, a calibration profile can be generated correlating the diffraction pattern of the bead to the distance between the bead and the OFP (Figure 13.2E). When this calibration profile is interpolated, the vertical displacements of the bead can be measured with an accuracy of up to ∼10 nm [9]. If one then measures a similar calibration profile for a reference bead fixed to the surface in the proximity of the tethered bead, a distance between the surface and OFP can be determined. Finally, the difference between the two measured distances yields l. The measurement of the calibration profile of the reference bead also allows one to perform a differential measurement between the tethered bead and the reference bead, minimizing the effect of drift between the objective and the sample holder (i.e., flow cell). To compute the applied force, both the mean vertical extension of the nucleic acid and the bead's in-plane fluctuations are required (see Section 13.4). The latter are computed via cross-correlation analysis of the intensity profiles displayed by a bead at subsequent time intervals [27,28]. As in the case of tracking the vertical position, in-plane tracking can be carried out at subpixel resolution to an accuracy of a few nanometers [9].

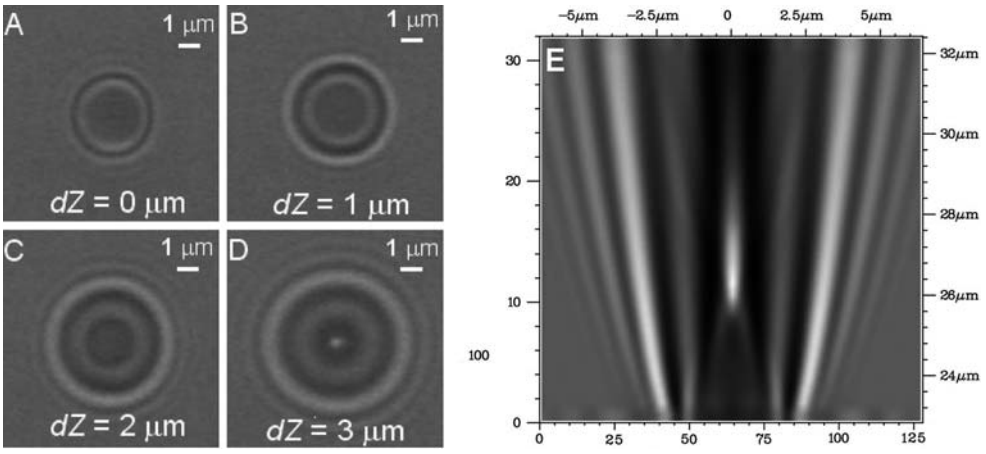


Figure 13.2. Generation of a calibration profile in the z direction for a bead in magnetic tweezers. A 10×10 μm image of the diffraction pattern of a reference bead with a diameter of ∼3 μm at different positions of the objective's focal plane (OFP). First, the OFP is positioned close to the bead **(A)**. Next, the OFP is moved away from the bead, and the images of the diffraction rings are recorded at successive positions of the OFP **(B–D)**. If the bead is illuminated with a parallel light source, the diffraction pattern of the bead will always be identical when the objective is focused at the same relative distance from the bead. By recording the size and intensity of the diffraction rings at different positions of the OFP a calibration profile can be made **(E)**, which can be used to measure vertical relative displacements of the bead with an accuracy of ∼10 nm, limited by the resolution of the piezo that controls the position of OFP.

13.4. Determination of the Applied Force

The applied force can be determined from the fluctuations of the in-plane position of the bead, provided that we know the extension of the nucleic acid, l (Figure 13.3) [9,29]. Two forces act on the bead: (1) the applied magnetic force (referred to as the "applied force" in the remainder of the chapter) and (2) the restoring force exerted by the nucleic acid. The total potential energy of the system E_p thus equals

$$E_p = E_{NA} + E_{mag} = A(l) - Fz \tag{13.3}$$

where $A(l)$ is the mechanical energy stored in the nucleic acid, which depends on the end-to-end extension of the nucleic acid l, and F is the applied force acting on the bead. In equilibrium, where the position of the bead $(x,y,z) = (0,0,l) = \vec{\mathbf{r_0}}$, all partial derivatives of the total potential energy are equal to zero. The potential gradient in the z direction can thus be expressed as

$$\frac{\partial E_p}{\partial z} = \frac{\partial A(l)}{\partial z} - F = \frac{\partial A(l)}{\partial l} - F = 0 \tag{13.4}$$

resulting in $F = \partial A/\partial l$. Brownian motion constantly forces the system out of its equilibrium position (indicated by the dotted outline of the bead in Figure 13.3). A Taylor expansion around the equilibrium position yields to leading order:

$$\frac{\partial^2 E_p}{\partial x^2} = \frac{\partial^2 E_p}{\partial y^2} = \frac{F}{l} \tag{13.5}$$

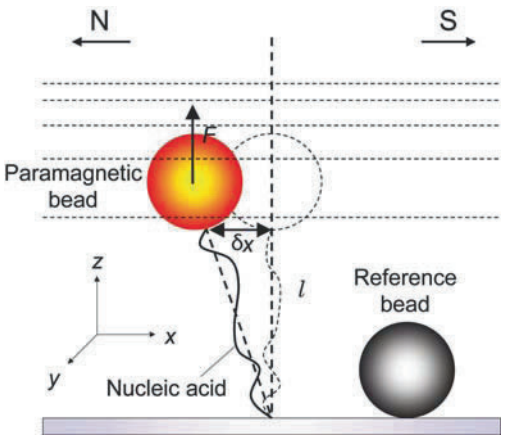


Figure 13.3. A schematic representation of a tethered bead with a reference bead fixed to the bottom cover slide. The paramagnetic bead (*orange*) is tethered by a nucleic acid polymer to a surface. In the magnetic field gradient the paramagnetic bead experiences a vertical stretching force F. The paramagnetic bead is constantly displaced from its equilibrium position (*vertical dashed line*) by Brownian motion of the solvent molecules. The displacements of the tethered paramagnetic bead δx are measured relative to a nonmagnetic reference bead that is fixed close to the surface (*gray*) to compensate for sample drift.

$$\frac{\partial^2 \mathrm{E_p}}{\partial z^2} = \frac{\partial F}{\partial l} \tag{13.6}$$

yielding an expression for the total potential energy around the equilibrium position:

$$\mathrm{E_p}(\overrightarrow{\mathbf{r}}) \approx \mathrm{E_p}(\overrightarrow{\mathbf{r_0}}) + \frac{1}{2}\left(\frac{F}{l}\right)\delta x^2 + \frac{1}{2}\left(\frac{F}{l}\right)\delta y^2 + \frac{1}{2}\left(\frac{\partial F}{\partial l}\right)\delta z^2 \tag{13.7}$$

Setting the potential energy in equilibrium position to zero, we can approximate the contribution of the displacement in the direction of the magnetic field (i.e., the x direction) to the average potential energy by the expression

$$\langle \mathrm{E_p} \rangle = \frac{1}{2}\frac{F}{l}\left\langle \delta x^2 \right\rangle \tag{13.8}$$

where F/l represents an effective trap stiffness in the x direction, k_x; and $\langle \delta x^2 \rangle$ is equal to the variance of the bead excursions in the x direction around its equilibrium position at $x_0 = 0$:

$$var\,\delta x \equiv \left\langle \delta x^2 \right\rangle - \langle \delta x \rangle^2 = \left\langle \delta x^2 \right\rangle \tag{13.9}$$

By the equipartition theorem, the energy of this degree of freedom equals $\frac{1}{2}k_\mathrm{B}T$. Rearranging terms, this yields a simple expression for the applied force,

$$F = \frac{k_\mathrm{B}Tl}{\langle \delta x^2 \rangle} \tag{13.10}$$

By measuring the extension of the molecule l and the variance of the bead excursions $\langle \delta x^2 \rangle$ in real space, one can thus evaluate the applied force from Eq. (13.10). Figure 13.4 shows an example of the determination of $\langle \delta x^2 \rangle$ and the subsequent calculation of the applied force using Eq. (13.10). A comparison of the bead excursions measured at two different forces confirms that the excursions of the bead center in the (x, y) plane indeed decrease with an increase in the applied force as predicted by Eq. (13.10) (Figure 13.4A). The actual value $\langle \delta x^2 \rangle$ is obtained, however, from the statistical analysis of x position of the bead as a function of time (Figure 13.4B). Subsequently, $\langle \delta x^2 \rangle$ and l can then be used to compute the applied force (Figure 13.4C). In practice, the use of Eq. (13.10) is limited to a certain force range. In particular, at forces greater than ~1 pN the measurement of $\langle \delta x^2 \rangle$ in real space shows a systematic error that can be corrected in Fourier space as described in the next section.

13.4.1. Calculation of the Applied Force—Analysis of the Brownian Motion of the Bead in Real and Fourier Space

Although it is possible to calculate the applied force using Eq. (13.10) in real space, the computation of $\langle \delta x^2 \rangle$ is typically performed in Fourier space because it allows for a more accurate determination. Namely, by decomposing the fluctuations in bead excursions into their spectral components, one can, for instance, compute corrections for the sample drift and the finite integration time of the CCD camera; these necessary corrections become larger asthe

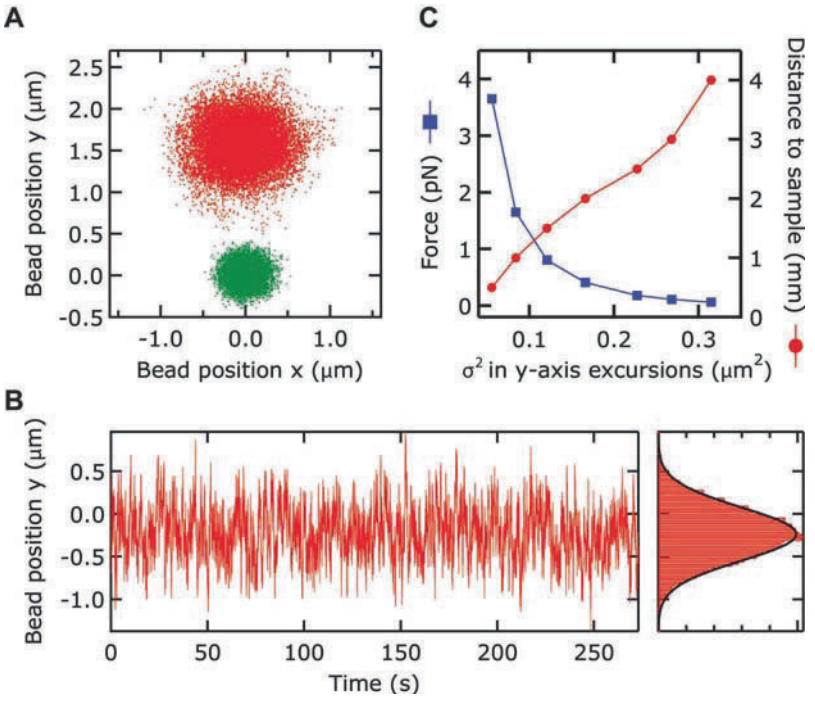


Figure 13.4. Measuring the applied force using the fluctuations in bead excursions in real space. **A.** The excursions of the bead center in the (*x, y*) plane for low force (*red*) and higher force (*green*). The data points for the low force have been offset for better representation. **B.** The bead excursion in the *x* dimension in time (the low-force measurement from panel A is shown), with a histogram on the right, including a Gaussian fit to the histogram. **C.** The computed variance of the bead excursions (not obtained from a fit to the Gaussian) for a variety of magnet positions (*red solid circles*), as well as the calculated corresponding forces (*blue solid squares*).

applied force increases (see the discussion on the correction for the camera integration time in Section 13.4.2).

In Fourier space, $\langle \delta x^2 \rangle$ can be calculated by integrating the power spectrum of the bead motion:

$$\left\langle \delta x^2 \right\rangle = \frac{1}{2\pi} \int \mathrm{P}(\omega) d\omega \tag{13.11}$$

where $\mathrm{P}(\omega)$ is the power spectrum [30]. To derive the theoretical power spectrum $\mathrm{P}(\omega)$ of the motion of the bead in the x direction, we consider the equation of motion of the bead:

$$m \frac{\partial^2 \mathrm{x}(t)}{\partial t^2} = -6\pi R\eta \frac{\partial \mathrm{x}(t)}{\partial t} - k_x \mathrm{x}(t) + \mathrm{f_L}(t) \tag{13.12}$$

where m is the mass of the bead, $\mathrm{x}(t)$ is the bead position in time, η is the viscosity of the buffer, R is the radius of the bead, $k_x, = F/l$ (Eq. 13.8), and $\mathrm{f_L}$ is a Langevin force at a given bandwidth caused by Brownian fluctuations of the bead. Taking the Fourier transform of Eq. (13.12) gives

$$X(\omega) = \frac{F_L(\omega)}{-m\omega^2 + 6\pi i\eta R\omega + k_x} \tag{13.13}$$

where $i = \sqrt{-1}$. Here $F_L(\omega) = 2\sqrt{6k_B T\pi\eta R}$ represents the square root of the power spectrum density of f_L [31]. With typical values for m and R of the bead equal to 10^{-15} kg and 10^{-6} m, respectively, along with $\eta = 10^{-3}$, $k_x = 10^{-9}$ N/m, and $\omega = (k_x/m)^{1/2} \sim 1$ kHz, the inertial term $m\omega^2$ is significant only at frequencies $\omega > \sim 6\pi\eta R/m \sim 30$ MHz. Because the sampling frequency of the camera is considerably smaller (typically ~100 Hz), the inertial term can be neglected in both Eqs. (13.12) and (13.13). Defining $\omega_c = k_x/6\pi\eta R$ and assuming $F_L(\omega)$ to be constant, we find that the square of Eq. (13.13) yields the theoretical power spectrum:

$$P(\omega) = \left|X^2(\omega)\right| = \frac{F^2(\omega)}{\left(1 + \frac{i\omega}{\omega_c}\right)^2 k_x^2} = \frac{2k_B T 6\pi\eta R}{\left(1 + \left(\frac{\omega}{\omega_c}\right)^2\right)\pi k_x^2} \tag{13.14}$$

Equation (13.14) is a Lorentzian, and by fitting the experimental power spectrum to Eq. (13.14) we obtain ω_c, which can be used to calculate k_x and subsequently the applied force F. Alternatively, the applied force can be calculated using $\langle \delta x^2 \rangle$ [Eq. (13.10)], which is determined by integration of the experimental power spectrum over all frequencies:

$$\left\langle \delta x^2 \right\rangle = \int_0^\infty \frac{24\pi k_B TR}{\pi k_x^2} \frac{1}{1 + (\omega/\omega_c)^2} d\omega \tag{13.15}$$

Experimentally, the integral in Eq. (13.15) is determined within a finite signal bandwidth at the high-frequency limit at half of the sampling frequency f_s and at low frequency at the inverse of the data acquisition time $f_L = {T_a}^{-1}$. Equation (13.15) can thus be used, provided that $f_L << f_c << f_s/2$, where f_c is the cutoff frequency of the system and $f_c = \omega_c/2\pi$ [9]; otherwise one does not have a statistically accurate determination.

Let us now compare the evaluations of the applied force in the real space and in Fourier space. The experimentally derived power spectra for the same two forces as employed in Figure 13.4A are shown in Figure 13.5A. The integration of the experimental power spectra (Figure 13.5B) yields corresponding variances that are in turn used to calculate the applied forces (Figure 13.5C). It is clear from the Figure 13.5C that the forces calculated in real space (red triangles in Figure 13.5C) overestimate the applied force, particularly at high forces, where larger excursions due to rapid fluctuations tend to be averaged out.

13.4.2. Correction for the Camera Integration Time

The calculation of the variance of the bead excursions necessary for the determination of the applied force requires a measurement of an accurate position of the bead. The systems commonly used to detect the image of the bead (e.g., a CCD camera) record the image for a finite time interval to acquire sufficient signal. Consequently, the acquired image does not reflect an instantaneous position x but instead a "blurred" average of instantaneous positions x_m explored by the bead during the finite time interval [9,30]. The correction for the camera integration time is derived below and follows that by Wong and Halvorsen [30].

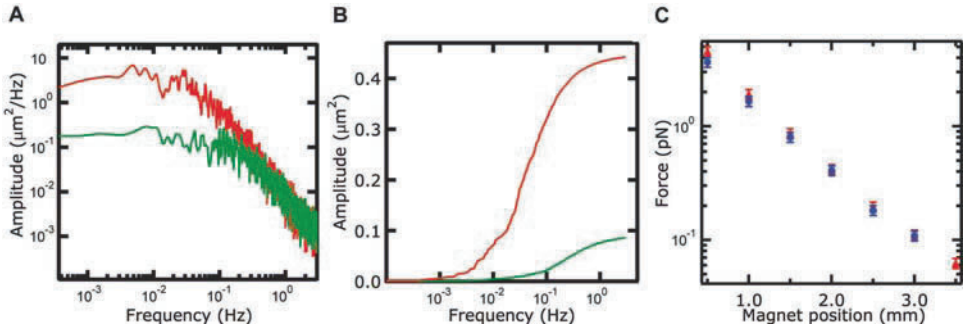


Figure 13.5. Measuring applied force in Fourier space. **A.** The spectra of the one-dimensional fluctuations for the same two forces as shown in Figure 13.4A. The spectrum is relatively flat until the cutoff frequency f_c (see text), after which they decay with $1/f^2$. Note qualitatively here that f_c increases with applied force. **B.** The integral of the spectra shown in panel A; the variance of the bead excursions can be obtained by integrating the spectra. Note that the variance decreases with increasing applied force. **C.** A comparison between the force measurement in real space (*red solid triangles*) and Fourier space (*blue solid circles*). In the low-force regime (less than ~1 pN), the two methods yield similar results. Only at higher forces, where the fluctuations of the bead are fast, does one run into the problem of underestimating the variance of the bead fluctuations, leading to an overestimation of the applied force.

In the case of a rectangular integration time window, the relation between the average x_m and x at time t can be obtained by convolving the true trajectory in x with a rectangular function [30]:

$$x_m(t) = \int_{t-W}^{t} x(t')\mathrm{H}(t-t')\,dt' \tag{13.16}$$

where H(t) is defined as

$$\mathrm{H}(t) = \begin{cases} \dfrac{1}{W} & 0 < t \leq W \\ 0 & \textit{Elsewhere} \end{cases} \tag{13.17}$$

where W is the integration time of the camera. Consequently, the true trajectory taken by the particle in the x direction cannot be extracted from a measurement with a finite integration time. However, assuming the bead is undergoing Brownian motion within a harmonic potential, variances in x_m and x can nonetheless be calculated, provided respective power spectra are known. The var(x) can be calculated using Eq. (13.15), and var(x_m) can be expressed using the definition of variance in Eqs. (13.9) and (13.16) as

$$\mathrm{v}ar(x_m) = \frac{1}{2\pi}\int \mathrm{P}(\omega)\mathrm{P}_\mathrm{H}(\omega)d\omega \tag{13.18}$$

If the power spectrum of the moving rectangular integration window is written as [30]

$$\mathrm{P}_\mathrm{H} = \left(\frac{\sin(\omega W/2)}{\omega W/2}\right)^2 \tag{13.19}$$

then Eq. (13.18) can be integrated to yield an expression for the variance in the measured position var(x_m):

$$\mathrm{v}ar(x_m) = \frac{2k_\mathrm{B}T}{k_x}\left[\frac{\tau}{W} - \frac{\tau^2}{W^2}\left(1 - e^{-W/\tau}\right)\right] \tag{13.20}$$

where $\tau = 1/\omega_c$. The var(x_m) can be expressed in terms of var(x):

$$\mathrm{v}ar(x_m) = \mathrm{v}ar(x)\left(\frac{2}{\alpha} - \frac{2}{\alpha^2}\left(1 - e^{-\alpha}\right)\right) \tag{13.21}$$

where α is a dimensionless exposure time W/τ. The expression in large parentheses on the right-hand side of Eq. (13.21) is referred to as the motion blur correction function [30]. This correction function has been experimentally verified over a range of parameters by Wong and Halvorsen [30].

The experimentally measured quantity var(x_m) is thus expressed in terms of the quantity one would like to know, var(x), and the parameter α, which is a function of ω_c. An iterative procedure can be used to determine the value of var(x). One starts with the experimental power spectrum $\mathrm{P_m}(\omega)$, undoes the effect of the camera windowing according to Eqs. (13.18) and (13.19), and finally fits the underlying spectrum $\mathrm{P}(\omega)$ according to Eq. (13.14) to obtain a value for ω_c. This procedure is repeated until the fitting error is below a minimally acceptable value, at which point var(x) can be correctly computed from the fitted $\mathrm{P}(\omega)$ using Eq. (13.15). This value for var(x) should comply with Eq. (13.21).

13.5. Nucleic Acids under Force and Torque

As described in Section 13.2, by controlling the distance between the magnets and the flow cell, one can control the force applied to the tethered bead. By rotating the magnets, one can apply a torque to the bead. The response of nucleic acids to such forces and torques can be analyzed using simple models of polymer physics. This section briefly reviews the behavior of DNA and RNA under applied force and torque.

If the bead in question is tethered to the surface by an unnicked, double-stranded nucleic acid molecule that is attached via multiple bonds at each end, the molecule is torsionally constrained and rotating the magnets will twist the DNA or RNA. In contrast, if the nucleic acid molecule that tethers the bead to the surface is attached only via a single bond at either end or if the molecule has a single or multiple nicks or is (partly) single-stranded, the molecule will swivel around the single bond(s) and remain torsionally unconstrained, even if the magnets are rotated.

The force–extension behavior of torsionally unconstrained double-stranded DNA molecules for forces less that 5–10 pN is well described by the worm-like chain (WLC) model (Figure 13.6A) [32–36]. The WLC model is a simple polymer model that treats the DNA as a homogeneous, isotropic polymer with a bending stiffness. It has two free parameters—the contour length L_c and the persistence length L_p. The length L_c is the total contour of the DNA, which is typically close to the crystallographic length, which is equal to 3.4 Å per base pair. L_p is the decay length of bending fluctuations, that is, the length scale over which the DNA undergoes significant bending [37]. Under typical experimental conditions (~10–100 mM monovalent salt), L_p is approximately equal to 50 nm, and its value decreases slightly with increasing ionic strength. The WLC model was shown to also provide an accurate description of the force–extension behavior of double-stranded RNA, with L_p of 64 nm [38]. For forces

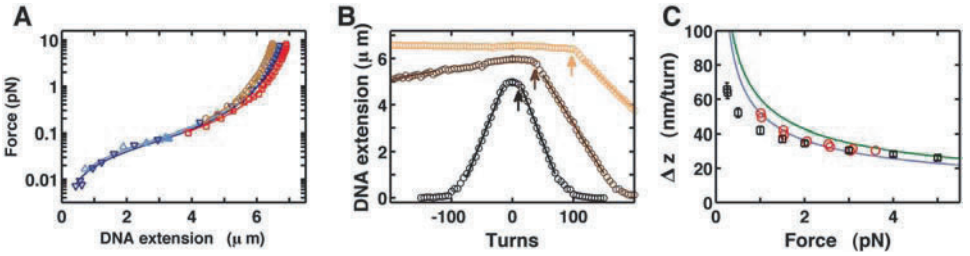


Figure 13.6. Behavior of double-stranded DNA under applied forces and torques. **A.** Force–extension curves of single 21-kbp DNA molecules. Symbols represent force measurements obtained from analysis of the spectrum of the bead fluctuations as described in Section 13.4. Solid lines are from fits of the worm-like chain (WLC) equation described in Section 13.5. The four different colors correspond to independent measurements with different magnet configurations, described in more detail in Section 13.7. The fits of the WLC equation gave contour lengths of 7.1 ± 0.2 μm and persistence lengths of 47 ± 3 nm (mean and standard deviation, respectively). **B.** DNA extension as a function of magnet rotations for a torsionally constrained 21-kbp DNA molecule at forces of 0.25 pN (*black*), 1 pN (*brown*), and 5 pN (*light brown*). At low forces (<1 pN) the formation of plectonemic supercoils is symmetric for positive and negative coiling. At higher forces, plectonemic supercoils are formed under positive torque, but the DNA denatures locally when a negative turn is applied, as can be seen from the asymmetry of the curves at 1 and 5 pN. The points at which the buckling transition into the plectonemic regime occurs at positive torques are denoted by errors. The slopes in the plectonemic regime (*solid lines*) give the change in extension per turn, Δz. **C.** Change in DNA extension per turn, Δz, deduced from the slopes of rotation curves (as in panel B). Black symbols show the results of magnetic tweezers measurements from our lab. Red symbols are the data of Forth et al. obtained using optical torque tweezers [42]. The green solid line shows the prediction of the simple model derived in the text. The blue solid line is the prediction of the Marko model, with a plectoneme helicity of 24 nm.

>5–10 pN, enthalpic stretching of the DNA becomes significant and can be incorporated into the WLC model by adding an elastic stretching term [34]. At forces greater than 50–60 pN, DNA undergoes an overstretching transition [39,40].

White showed that for torsionally constrained DNA the linking number *Lk* is a topological invariant [41]:

$$Lk = Tw + Wr = \text{const} \tag{13.22}$$

Tw is the twist of the double helix, defined as the number helical turns in the double helix of DNA. For torsionally relaxed DNA the natural twist Tw_0 of DNA is equal to the number of base pairs divided by the helical turn (10.5 bp per turn). The writhe *Wr* is the number of times the axis of the double helix crosses itself. In the absence of external constraints, the writhe is zero, $Wr_0 = 0$. A molecule is said to be supercoiled when its linking number is different from $Lk_0 = Tw_0 + Wr_0$. A useful quantity in this context is the supercoil density σ:

$$\sigma = \frac{Lk - Lk_0}{Lk_0} \tag{13.23}$$

Rotating the bead attached to a torsionally constrained nucleic acid changes *Lk*, as shown in Figure 13.6B. Starting from a torsionally relaxed molecule, the change in linking number is initially absorbed by elastic twist deformations and increases *Tw* of the molecule, whereas *Wr* remains unchanged. In this regime, the torque Γ increases linearly with the number of turns *N*:

$$\Gamma = C\frac{2\pi N}{L_c} \tag{13.24}$$

where C is the torsional modulus; $C \approx 90k_BT$ for DNA [43,44]. If one continues to rotate the magnets, the molecule undergoes a buckling transition when the energy stored in the elastic twisting deformation is reduced by forming a loop (see the arrows in Figure 13.6B). Further rotations past the buckling transition do not increase *Tw* but instead increase *Wr* as plectonemic supercoils are formed. A simple model for the buckling torque Γ_B and the size of the plectonemes can be derived by assuming a loop of size R [42,43]. At Γ_B, the work done by one more turn is $2\pi\Gamma_B$ and equals the work required to form a loop of size R:

$$E_R = 2\pi RF + \frac{\pi L_p k_B T}{R} \tag{13.25}$$

where the first term corresponds to the work done against the external (magnetic) force F and the second term is the work required to bend a stiff polymer into a loop of radius R. By minimizing E_R with respect to R, one finds for Γ_B and for the change in extension per superhelical turn (Δz) [43,44]

$$\Gamma_B = \sqrt{2L_p k_B TF} \text{ and } \Delta z = 2\pi R = \pi\sqrt{\frac{2L_p k_B T}{F}} \tag{13.26}$$

This simple model qualitatively yields the right trends. In particular it predicts an increase in Γ_B and a decrease in the shortening of the DNA per turn with increasing applied force. However, it fails to quantitatively match the experimental data, in particular at low forces (Figure 13.6C). To describe the elastic response of twisted DNA more accurately, a number of models have been developed. Recently, Marko derived an analytical model that includes an additional parameter—the twist stiffness of the plectonemic state [45]. The Marko model provides additional physical insight and improves the fit to the experimental data in the force range >1 pN. However, it asymptotically approaches the result of the simple model at low forces and fails to accurately match the experimental data below ~1 pN (Figure 13.6C). In addition, neither model explicitly accounts for the effect of ionic strength. Experimentally, increasing the ionic strength of the solution decreases Δz at constant force, which is qualitatively expected because increased electrostatic screening makes smaller plectonemic loops energetically less costly. Neukirch and coworkers derived a theoretical framework for obtaining the effective plectoneme radius from experimental data [46,47]. Another description of DNA under force and torque is the rod-like chain (RLC) model, an extension of the WLC model that considers twisting rigidity and twist–stretch coupling. The RLC model appears to yield an accurate description of twisted DNA under low forces; however, it can only be solved numerically [48–50].

13.6. Current Capabilities in Terms of Temporal and Spatial Resolution: Practical Limitations

The foregoing considerations described the experimental configuration of the MT and explained how it can be used to apply forces and torques to nucleic acids. Particularly in applications to nucleic acid–protein interactions, it is of interest to know with what resolution in both space and time one can monitor the molecule.

Assuming that one measure long enough (i.e., longer than $1/f_c$) to probe the full excursion probability distribution, one can obtain an unbiased and infinitely small spatial resolution by measuring an infinite number of time points. For a measurement of N data points, the accuracy with which one measures the x, y, or z position is given by the standard error of the mean (S_E):

$$S_E = \frac{SD}{\sqrt{N}} \tag{13.27}$$

In other words, measuring faster will only lead to an increase in spatial resolution when the sampling takes place longer than $1/f_c$, and thus f_c sets the temporal resolution. For example, with an acquisition frequency of 60 Hz, one obtains a spatial resolution of ~30% in 200 ms, or 3% in 17 sec at 2 pN. According to Eq. (13.8), the MT trap stiffness decreases inversely with the length of the nucleic acid. A reduction in trap stiffness with nucleic acid length leads to a reduced f_c. MTs typically have a lower temporal (and in practice thus a lower spatial) resolution than other force-spectroscopy techniques such as atomic force microscopy and optical tweezers [26,51]. To a certain extent, this may be compensated by running experiments with short molecules at high forces, although the degree to which this is possible depends on the biological question being addressed. Nevertheless, the typical spatial resolution of MTs is between 2 and 10 nm, and their resolution is usually between 10^{-2} and 10^{-4} sec.

13.7. Optimization of the Magnet Geometry

Section 13.4 described how the applied force can be calibrated from the motion of the bead using the equipartition theorem. This section discusses how the applied force from a pair of permanent magnets can be calculated from basic physics principles, without requiring measurement of any bead fluctuations. Being able to compute the magnetic force from first principles provides a consistency check on the calibration from the fluctuations in the in-plane position of the bead. More importantly, it provides physical insight and allows the design of magnet geometries for particular experimental specifications without the need to build and test a large number of magnet configurations.

The magnetic force exerted on a bead is given by the negative gradient of the magnetic energy [Eq. (13.1)]. To compute the force exerted by a given magnet configuration on a bead, both $\vec{\mathbf{B}}$ and $\nabla\vec{\mathbf{B}}$ at the position of the bead, as well as the value of $\vec{\mathbf{m}}$ for a given external magnetic field, have to be determined. Two complementary strategies can be used to compute $\vec{\mathbf{B}}$ for a given magnet configuration [52]. For simple symmetric geometries and in the absence of an iron yoke or other complicating factors, it is possible to compute the magnetic field from first principles using the Biot–Savart law [53]. In this case, the magnets are replaced by current loops using the method of equivalent sources and integrating their contributions to the magnetic field at any point in space. This approach is semianalytical: Only the one-dimensional integral derived from the Biot–Savart law needs to be evaluated numerically. The only property of the magnets (apart from their geometry) required for the calculation is their remnant field, which is available from the manufacturer's specifications. Alternatively, the remanent field can be determined from a direct measurement of the surface field by using a Hall probe.

An alternative strategy to evaluating the magnetic field of arbitrary magnet geometries is to solve the magnetostatic problem using a three-dimensional (3D) finite-element partial

differential equation solver. The Multiphysics AC/DC module from COMSOL has been used to build complete 3D models of several magnet geometries and to solve the magnetostatics problem numerically [52]. The finite-element approach can be applied to more complicated geometries. In addition, material properties of components of the setup, for example, an iron yoke or iron tips, can be included in the model.

To quantitatively assess the accuracy of magnet simulations, the magnetic field of a pair of permanent magnets has been measured using a Hall probe [52]. Figure 13.7A shows the magnetic field as a function of distance to the center of the magnet pair for a gap size (i.e., the distance between the magnets) of 1 mm and vertical (red and brown) and horizontal (light and dark blue) magnet alignments (see Figure 13.1 for a definition of the magnet alignments). Data are shown for the absence (red and light blue) and presence (brown and dark blue) of an iron yoke (Figure 13.7A). The semianalytical calculations (black dashed lines) are in excellent agreement with the results of the 3D finite-element modeling (solid lines) for the geometries without added iron yokes (Figure 13.7A). For the case of added iron yokes, only the finite-element solution was obtained. In addition, the theoretical predictions are in good agreement with experimentally measured values for the magnetic field (symbols in Figure 13.7A). Similar agreement for other gap sizes has been obtained. It should be emphasized that the lines in Figure 13.7 are direct predictions from magnetostatic theory, without any free fitting parameters. The data suggest that it is possible to accurately model the magnetic field of a pair of magnets from first principles.

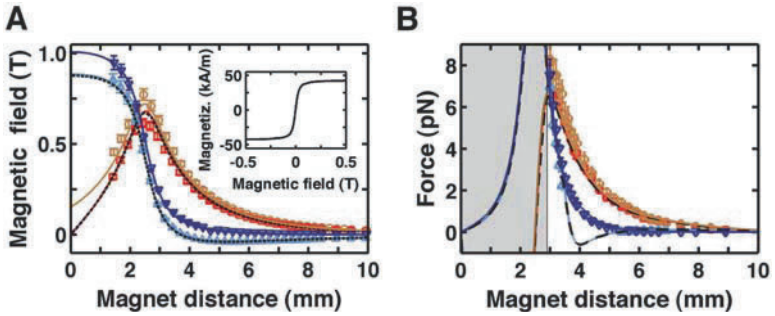


Figure 13.7. Calculations of magnetic fields and forces from first principles. **A.** Magnetic fields as a function of distances from the center of magnet pairs in the vertical (*brown, red*) and horizontal (*light blue, dark blue*) configuration. Data points are from measurements with a Hall probe in the absence (*red, light blue*) and presence (*brown, dark blue*) of an iron yoke. Solid lines show the results of corresponding calculations using a three-dimensional (3D) finite-element solver, and black dashed lines show the results of semianalytical calculations (Section 13.7). The inset shows the magnetization of MyOne beads from Invitrogen as a function of the external magnetic field as reported by the vendor. **B.** Magnet forces exerted on a 1-μm MyOne bead by the magnet configurations described in Figure 13.1 (with the same color code as in panel A). Symbols represent experimental results determined from the analysis of the bead's fluctuations. Lines are predictions from 3D finite-element simulations (*solid lines*) and the semianalytical theory (*black dashed lines*). The region that is experimentally inaccessible due to the finite thickness of the magnets and the flow cell is shown as a shaded region in the plot. Force–extension curves for these data are shown in Figure 13.6A, using the same color code as here.

To compute the force exerted on a bead, knowledge of m in response to an external B is required. For beads commonly used in MT experiments, m is approximately described by the Langevin function [26], and these data are available from the literature (for Dynabeads MyOne beads from Invitrogen see Ref. 54 and the manufacturer's Web site; for MagSense 1-μm beads see the technical support at the manufacturer's Web site). The inset of

Figure 13.7A shows m for MyOne beads as a function of the external B. Two regimes can be distinguished: For external B less than ~50 mT, m increases approximately linearly with the external B, and, as a result, F will grow as ∇B^2. For B greater than ~100 mT, m saturates and is approximately constant. In this regime, F will be proportional to ∇B. In typical MT configurations, B spans both the linear and saturation regimes, and, therefore, it is crucial to take into account the full magnetization behavior in the force calculation.

Figure 13.7B shows force measurements for the different magnet geometries considered in the foregoing using DNA-tethered MyOne beads. The theoretical predictions from semianalytical theory (dashed line) and from the 3D finite-element simulations (solid lines) are in good agreement with the direct measurements of the force from the observed fluctuations in the bead in-plane position (symbols; analyzed in Section 13.4). Several interesting observations can be made. For the horizontal magnet configuration and in the absence of an iron yoke (light blue lines), the force becomes negative, that is, the bead is pushed toward the surface. Experimentally, one cannot measure negative forces using the strategy outlined in Section 13.4; therefore only positive (i.e., pulling) forces are shown in Figure 13.7B. Adding an iron yoke to the setup suppresses the negative-force region for the horizontal configuration (dark blue curves). Of interest, addition of an iron yoke increases the maximum force that can be applied in either configuration only slightly.

The applied force decreases much faster with increasing distance to the magnets in the horizontal configuration than in the vertical configuration. For applications that require measurements at very low forces (<0.1 pN; Figure 13.6A), a horizontal magnet configuration is therefore desirable. For the horizontal configuration, reducing the gap size between the magnets does not significantly increase the maximum force that can be applied because the maximum of the force occurs at a position inside the gap between the magnets that is not experimentally accessible. In contrast, for the vertical magnet configuration, decreasing the gap size can significantly increase the maximum applicable force. For applications that benefit from large applied forces, a horizontal magnet configuration with a small gap size and a thin flow cell are therefore recommended. In addition, the choice of beads can be optimized for particular applications. For example, to maximize the forces that can be applied, large beads with a high magnetization per volume are desirable.

13.8. Flow Cells for Magnetic Tweezers

MT measurements require specially designed reaction chambers or flow cells. A flow cell must satisfy several criteria to be suitable for MT measurements. First, the walls of the flow cell have to be transparent for the light used to visualize the beads (e.g., quartz glass slides). Second, a flow cell requires an inlet and an outlet to enable an exchange of buffers during an experiment. Third, the inner surface of the flow cell opposite to the magnet must provide attachment points for the nucleic acid conjugated to the bead (Figure 13.1). Finally, the remainder of the inner surface not covered by the attachment points has to be passivated to prevent any nonspecific interactions that may interfere with the measurement of the extension of the nucleic acid or of the applied force (e.g., sticking of the nucleic acid and/or bead to the inner surface). The following paragraphs review the current strategies for nucleic acid labeling to suspend the nucleic acids between the bead and the glass surface, as well as ways to passivate the inner glass surface.

13.8.1. Strategies for Tethering Nucleic Acids to the Flow Cell and the Bead

To suspend a nucleic acid between the attachment point on the inner surface of the flow cell and the bead, the two extremities of the nucleic acid have to be specifically labeled with two different labels. The labels have to provide bonds that can resist the applied forces. The inner surface and bead are thus usually coated with either antidigoxigenin or streptavidin, whereas the corresponding extremities of the nucleic acid are labeled with their binding partners, digoxigenin or biotin, respectively [55]. The fluorescein–antifluorescein binding pair has also been used in MT and tethered particle motion (TPM) applications, with fluorescein placed on the nucleic acid end and antifluorescein adsorbed on the inner surface of the flow cell or attached to the bead [56,57]. Considering the stability of the fluorescein–antifluorescein binding pair against forces typically applied in MT studies, it is a viable alternative to the two standard binding pairs mentioned previously.

Labeling of nucleic acids can be carried out in several ways: (1) Generation of labeled DNA and RNA fragments by polymerase chain reaction (PCR) or *in vitro* run-off transcription, respectively, followed by their ligation to the corresponding nucleic acid terminus using DNA or RNA ligase [58,59]; (2) Chemical modifications of the nucleic acid ends [56,60–62]; (3) enzymatic extension of the nucleic acid ends with modified nucleotides using PolyA polymerase (pAP), terminal deoxynucleotidyl transferase (TdT), or T4 RNA ligase 1 (T4 RNL 1) [63–65]; or (4) hybridization of a labeled RNA or DNA strands that are complementary to the ends of the nucleic acid [38,58,59,66,67]. The choice of optimal labeling protocol for a particular application will primarily depend on the nature of the nucleic acid (i.e., DNA vs. RNA, and single- vs. double-stranded nucleic acid) [55]. Although ligation-based protocols have been applied with great success in DNA-related studies [55,68–70], the other protocols listed have been developed primarily for RNA-based applications due to the limited selection of molecular biology tools that can be used in the synthesis of RNA molecules. Finally, application of a torque in the MT requires a torsionally constrained double-stranded nucleic acid, which is most conveniently obtained by introducing multiple labels at each end of the nucleic acid using PCR and *in vitro* run-off transcription of corresponding DNA and RNA, respectively [38,58,68].

13.8.2. Inner Surface Passivation Techniques

Ideally, a nucleic acid molecule should be suspended between the inner surface and the bead in the absence of any interfering interaction among the nucleic acid, bead, and the inner surface. However, most of the reaction buffers used in *in vitro* biological reactions contain components that promote interaction of the nucleic acid and/or the bead with the flow cell surface (e.g., Mg^{2+}) [71,72]. These interactions affect the measured extension of the nucleic acid, as well as the fluctuations of the bead, which prevents an accurate determination of the applied force and mechanical properties of the system under investigation. Thus, to prevent these interactions, the inner surface of the flow cell has to be passivated after the adsorption of the attachment points for the nucleic acid (e.g., antidigoxigenin, biotin, or fluorescein).

The most common way to passivate a flow cell surface is by adsorbing inert proteins (e.g., bovine serum albumin [BSA]) or polypeptides (e.g., poly-L-glutamic acid [PGA]) to the part of the surface not occupied by the nucleic acid attachment points [38,68,73]. Although both the passivating material and the nucleic acid attachment points can be adsorbed to bare glass with reasonable efficiency, an intermediate polystyrene layer has been commonly used to enhance their adsorption [38,68]. Alternatively, a monolayer composed of poly(ethylene

glycol) (PEG) and biotinylated PEG has been successfully used to reduced nonspecific interactions between the inner surface and the components of the system [74] (for the properties of the PEGylated glass surface see Ref. 75). In addition, the authors' lab has used nitrocellulose-based passivation of the flow cell surface [76], which is characterized by its ease of preparation and high density of nucleic acid tethers.

13.8.3. Considerations When Working with RNA

There are several reasons why MT experiments on RNA have been few and far between. First, during every step of MT experiments, RNA systems require special considerations not necessary in the case of DNA. For example, due to the lack of a molecular biology toolbox comparable to that for DNA, a combination of molecular biology and chemical techniques is required to assemble a viable RNA construct (for a comprehensive review of available techniques and protocols, see Ref. 55). Furthermore, a major obstacle to a wider application of MT to RNA-related research is the instability of RNA under *in vitro* conditions. As is widely acknowledged, there are two principal reasons why RNA is less stable than its deoxyribose counterpart, DNA: (1) the abundance of RNA-degrading enzymes—RNases—in the environment [77], and (2) RNA's chemical instability due to the presence of 2′-hydoxyl groups, which can serve as nucleophiles to attack their proximal phosphorous atoms, leading to a nucleophilic substitution and resulting in a cleavage of the phosphodiester bond [78]. Nevertheless, there are measures that can enhance RNA stability and enable a successful RNA MT study [38].

Whereas RNases are active under a wide range of conditions, the RNA phosphodiester bond cleavage by 2′-hydoxyl group requires elevated temperatures and the presence of divalent ions as catalysts [79–82]. RNase nuclease activity in the MT can be inhibited by the use of both RNase-free reagents and commercially available RNase inhibitors. Alternatively, the cleavage of the phosphodiester bond by the 2′-hydorxyl group can be considerably inhibited by the use of chelating agents that bind divalent ions (e.g., ethylenediaminetetraacetic acid or sodium citrate). In addition, the use of inert surfaces such as PEGylated glass slides further increases the stability of RNA constructs in the MT.

13.9. Use of Magnetic Tweezers in Biological Experiments: Examples

Having detailed the MT experimental configuration in the preceding sections, this section considers three different examples of experiments in which DNA–protein interactions have been monitored using MT. The first two examples depend on the unique ability of MT to easily control and monitor the degree of supercoiling in a single molecule of DNA. The ease with which this can be done has allowed researchers to probe the mechanical properties of supercoiled DNA and to study the activity of topoisomerases (Example 1) (Figure 13.8). It will also be shown how MT can be used to probe restriction enzymes, RNA polymerases, and translocases, whose primary function is not in supercoil regulation but whose activities can be monitored indirectly via the degree of supercoiling induced (Example 2) (Figure 13.9). In addition, it is shown how MT can be used to study processive translocating enzymes such as DNA helicases (Example 3) (Figure 13.10). Finally, recent MT studies on purely protein-based systems are reviewed (Example 4).

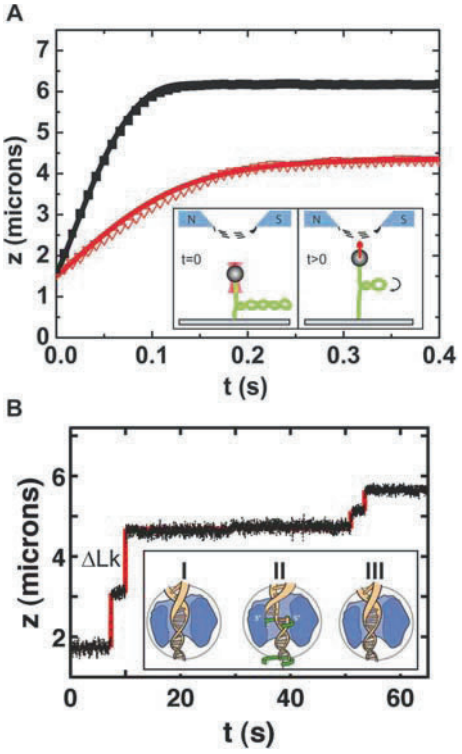


Figure 13.8. Examples of magnetic tweezers studies of supercoil dynamics and removal. **A.** The dynamics of a supercoiled DNA molecule bound to a bead can be addressed in two ways. First, combined magnetic and optical tweezers were used as shown in the inset. The laser beam used to trap the bead is represented in pink. Alternatively, supercoils were removed in the presence of a nicking enzyme. The traces shown here were obtained with the combined tweezers setup for two different values of the magnetic force: 2.2 pN (*inverted red triangles*) and 4.1 pN (*black squares*). These traces monitor the end-to-end extension of the DNA molecule, which increases after release of the optical trap at $t = 0$ until it reaches its steady-state value under the magnetic force. The solid lines represent the quasistatic model that was used by the researchers to successfully describe the dynamics. **B.** Supercoil removal by topoisomerase IB. Topoisomerase IB first binds double-stranded DNA (dsDNA) (step I in the inset), cleaves one strand of the dsDNA, which allows the release of the supercoils (step II in the inset), and finally ligates the cleaved strand to yield back the intact dsDNA (step III in the inset). On supercoil release, the measured extension, z, increases.

13.9.1. Example 1: Supercoils Dynamics, and Supercoil Removal

During cell division, the genetic material is copied, and the enzymes responsible for this must be able to access the base sequences. This is only possible if the portion of DNA to be replicated is unwound. Such DNA unwinding, which takes place during both replication and transcription, gives rise to supercoils in the DNA. The degree of supercoiling must be carefully regulated to avoid impeding the motion of the molecular machinery because supercoils can give rise to torsional forces in the DNA, the magnitude of which increases the more the DNA is unwound. In the context of DNA replication, such forces can delay the process of cell division and under certain conditions even arrest it [83].

A number of processes, both mechanical and biochemical, are involved in supercoil regulation. In this context, MT can be used to probe the dynamics of supercoiled DNA to

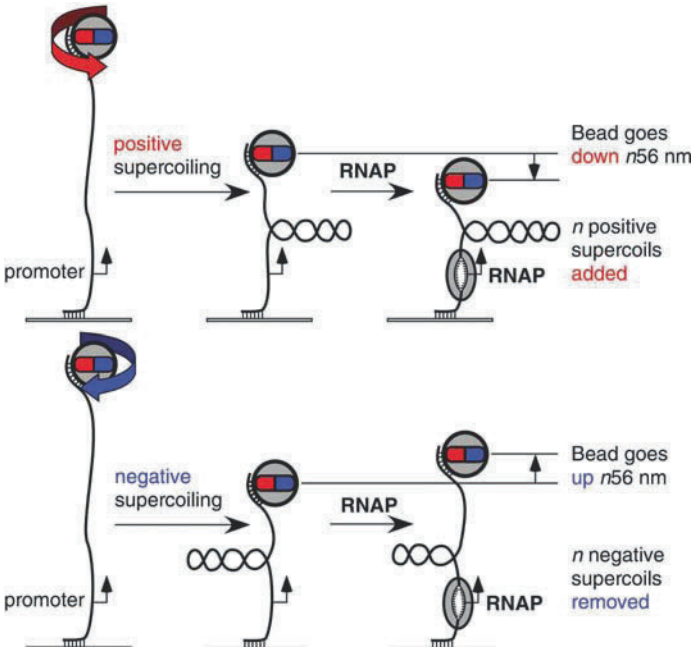


Figure 13.9. The experimental strategy employed to observe DNA untwisting by RNA polymerase (RNAP). A torsionally constrained molecule is held in molecular tweezers. Unwinding of DNA by RNAP will induce negative twist, which in a torsionally constrained molecule must be compensated by positive writhe. If the DNA prior to the unwinding by RNAP contains positive writhe, the activity of the RNAP will serve to further increase the writhe. This will lead to an observable decrease in the end-to-end extension of the DNA molecule **(top)**. If the DNA prior to the unwinding by RNAP contains negative writhe, the activity of the RNAP will serve to make the overall degree of writhe less negative. Consequently, this will lead to an observable increase in the end-to-end extension of the DNA molecule, as indicated **(bottom)**. (From Revyakin et al. [86]. Reprinted with permission from AAAS. American Association for the Advancement of Science.]

assess its precise role in such regulation. For instance, diffusing supercoils of opposite sign might annihilate on a circular plasmid, provided their drag is low enough to allow their rapid diffusion. MT can also be used to probe enzymes called topoisomerases that control DNA supercoiling. These enzymes serve to reduce the torsional forces and to unknot DNA. They are grouped into two classes, types I and II. Type I topoisomerases are characterized by their ability to cleave a single strand of duplex DNA, whereas type II topoisomerases cleave both strands of duplex DNA.

Two distinct methods have been used to examine the dynamics of supercoiled DNA. In the first method, the force on a supercoiled DNA molecule was rapidly increased, leading to a conversion of the plectonemic supercoils to DNA twist (Figure 13.8A, inset). To achieve rapid force switching in practice, MTs were supplemented with an optical trap to keep DNA under low tension while the magnets were translated to increase the force. Subsequently shutting off the laser trap led to motion of the bead back to its equilibrium position under the high magnetic force established. In the second method, supercoil removal dynamics were probed by incubating the supercoiled DNA in the presence of a nicking enzyme that relaxes the torsional constraint by nicking the DNA [84]. Both experiments on plectoneme dynamics could be described in a satisfying way with a simple theoretical model assuming fast internal

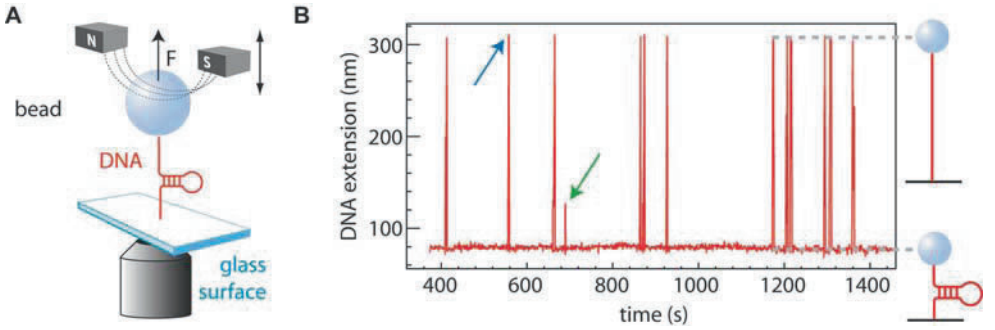


Figure 13.10. Unwinding of a DNA helix by bacteriophage T4 helicase (gp41). **A.** Experimental configuration for studying gp41 unwinding of a DNA duplex, where the DNA hairpin is suspended between the glass surface and the bead. **B.** In the presence of gp41, the DNA hairpin is unwound, resulting in the lengthening of the DNA, followed by shortening as a consequence of rehybridization of the complementary strands back into a DNA hairpin. The initial increase and subsequent decrease in DNA extension results in the observed spikes. The majority of spikes correspond to the unwinding of the entire DNA hairpin (an example is indicated with a *blue arrow*). Occasionally, partial unwinding events can be observed as indicated by lower spikes (*green arrow*). (Adapted from Lionnet et al. [90]. Copyright 2007 National Academy of Sciences, USA.)

DNA dynamics (data for the force switching experiment are shown in Figure 13.8A). Surprisingly, the relaxation of plectonemic structures that occurred during these experiments did not significantly slow down DNA dynamics [84]. This study of the dynamics of bare DNA has provided a solid baseline to which the dynamics of supercoiled DNA in the presence of proteins should be compared.

To illustrate the use of MT in probing topoisomerase activity, consider the enzyme topoisomerase IB, a eukaryotic enzyme that releases the torsion built up in a supercoiled DNA. The enzyme releases the torsion from the DNA by surrounding the dsDNA like a clamp and temporarily cleaving one of the two strands. The accumulated torsional forces in the DNA are then spun out around the intact strand. After a number of turns, the topoisomerase IB again firmly grabs the spinning DNA and neatly ligates the broken strands back together (Figure 13.8B, inset). Using single-molecule techniques, Koster et al. could follow single topoisomerase IB enzymes over time as they acted on single DNA molecules [68]. Because the supercoils are removed from the DNA as soon as the topoisomerase cuts through one of the two DNA strands, they were able to determine the exact number of turns removed by the topoisomerase between "cutting" and "sealing" (Figure 13.8B). In addition, various parameters, such as the torque dependence of the average number of turns removed and the degree of friction of the rotating DNA in a cavity of the enzyme, could be measured. The experimental evidence for the existence of friction relied on a comparison with a control experiment in which supercoil relaxation was carried out by a nicking enzyme, as described previously, because such an enzyme does not impose noticeable friction following cleavage.

13.9.2. Example 2: DNA Scrunching by RNA Polymerase

In the previous example, changes in supercoil density were shown to yield kinetic information on enzymes whose primary role involves supercoil removal. However, the precise analysis of supercoils in the magnetic tweezers can also yield information on enzymes whose primary role is quite different. An example of this is the detection by Strick and colleagues of

DNA scrunching by RNA polymerase during abortive initiation [85,86]. In this experiment, a single DNA molecule that included a unique promoter for *Escherichia coli* RNA polymerase (RNAP) was prepared in the MT. As shown in Figure 13.9, the molecule can be prepared in either a positively or a negatively supercoiled state. RNAP will start unwinding the DNA, but as the DNA molecule is torsionally constrained, the linking number cannot be altered. Any negative twist induced through DNA unwinding by RNAP must therefore be compensated by positive writhe. If the DNA is initially positively supercoiled, the compensation by positive writhe will decrease the end-to-end extension of the DNA molecule, whereas if the DNA is initially negatively supercoiled, the compensation by positive writhe will increase the end-to-end extension of the DNA molecule (Figure 13.9). In this manner, the researchers could measure the degree of unwinding by the RNA polymerase in several states (open complex, initially transcribing complex, etc.). For instance, the ability to measure the small differences in the change of DNA length during abortive initiation in a positively versus a negatively supercoiled state, equal to $\Delta l_u = (\Delta l_{obs,neg} - \Delta l_{obs,pos})/2$, where $l_{obs,neg}$ is the length change of the DNA during abortive initiation at negative supercoiling and $\Delta l_{obs,pos}$ is the corresponding length change at positive supercoiling, proved the existence of DNA scrunching by RNAP in this phase of its cycle.

13.9.3. Example 3: DNA Helicase Activity

It has been shown how a knowledge of the behavior of supercoiled DNA in the tweezer can permit sensitive detection of enzymatic activity. MTs are by no means limited to the detection of enzymes that act on supercoils, either directly or indirectly. Indeed, they have also been used to monitor replication by DNA polymerases, unwinding by DNA helicases, and looping by GalR, among others [87–89]. This subsection briefly describes how MTs can be used to study DNA helicase activity.

As a helicase unwinds duplex DNA or RNA into two separated single strands, its motion can be monitored in several ways. One manner is to take advantage of different force–extension relations for single- and double-stranded nucleic acids (i.e., different persistence lengths and contour lengths per nucleotide). Under a fixed force, these varying stiffnesses result in altered end-to-end lengths monitored in the MT, depending on whether the nucleic acid is predominantly single or double stranded [40]. Such discrimination has been used to monitor the activity of DNA helicases [88]. An alternative technique to monitor helicase activity, first introduced by the Bustamante group, is to monitor DNA unzipping of a single hairpin [59]. Here, the signal-to-noise ratio is higher because it is not a length differential between single- and double-stranded nucleic acids that contributes to the signal, but rather the entire length revealed by the appearance of additional single-stranded DNA under tension. Unzipping of a 231-bp hairpin by bacteriophage T4 helicase (gp41), leading to lengthening of the DNA during unzipping and shortening during subsequent rehybridization, is shown in the context of MT in Figure 13.10 [90]. From such time traces, one can readily deduce the velocity and the processivity distribution, as well as the force dependence of helicase activity, very precisely, which places stringent constraints on potential helicase mechanisms.

13.9.4. Example 4: MT Applications in Protein Science

Recently MTs have been applied to nucleic acids–free protein systems [10–14]. Lee et al. developed an immunoassay using MTs that enabled antigen detection with two-orders-of-magnitude-higher sensitivity than conventional solid surface phase immunoassay techniques

due to the increase in the signal-to-noise ratio in the case of MTs [14]. In addition, MTs have been used to measure the lengths and dissociation rates of bonds between phage-displayed peptides and their protein targets [13]. This work was later expanded with studies of antigen–antibody interactions [12]. Finally, MTs were used to measure the stiffness of protein assemblies, as well as the bond strengths stabilizing these assemblies [10, 11].

13.10. Outlook

Single-molecule MT techniques have been applied to many biological problems with considerable success. Although they were applied initially to study DNA supercoiling and its biological regulation, MTs have also been used to study other important systems involving nucleic acids and proteins, such as nucleic acid helicase and polymerases. There are several technical and biological challenges that remain to be solved. First, the accuracy of spatial tracking should continue to be improved, and the MT should be extended to directly measure the applied torque. Second, developing the instrumentation and software to enable high-speed parallel measurements of several beads (i.e., several single molecules) will prove beneficial in studying biological processes with low initiation rates [91]. Furthermore, combing MT with other single-molecule techniques such as fluorescence resonance energy transfer or optical tweezers should expand the versatility of the technique, enabling studies of multicomponent processes. Finally, extending the abilities of current MTs to investigate nucleic acid–dependent metabolism in living cells, in which various proteins are bound to DNA under "truly" *in vivo* conditions, should further diversify the experimental questions that can be addressed by this technique.

Acknowledgments

We thank the many members of our laboratory who have contributed to the research effort using MTs over the last few years. We are particularly grateful to Jeroen Abels for a number of illustrations, to Xiaomin Hao for her help on magnet modeling, to Sven Klijnhout for carrying out the experiments necessary to test the elastic response of twisted DNA, to Zhuangxiong Huang and Gary M. Skinner for help with flow cell coating, and to Serge Lemay for pointing out the alternative derivation of the spring constant of the MT. We also thank Richard H. Ebright and Timothee Lionnet for providing the figures on RNA polymerase scrunching and DNA helicase unwinding, respectively. Funding is acknowledged from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) via its Vidi program (IDV and JL), from the Human Frontiers Science Program (DAK), and from the European Science Foundation via a EURYI grant (NHD).

References

1. Palker TJ (1992) Human T-cell lymphotropic viruses—Review and prospects for antiviral therapy. Antivir Chem Chemother 3: 127–139
2. Chase JW, Williams KR (1986) Single-stranded DNA binding proteins required for DNA replication. Annu Rev Biochem 55: 103–136
3. Kowalczykowski SC, Dixon DA, et al. (1994) Biochemistry of homologous recombination in Escherichia coli. Microbiol Rev 58: 401–465

4. Wickner RB (1992) Double-stranded and single-stranded RNA viruses of *Saccharomyces cerevisiae*. Annu Rev Microbiol 46: 347–375
5. Wickner RB (1996) Double-stranded RNA viruses of *Saccharomyces cerevisiae*. Microbiol Rev 60: 250–265
6. Tomari Y, Zamore PD (2005) Perspective: machines for RNAi. Genes Dev 19: 517–529
7. Voinnet O (2001) RNA silencing as a plant immune system against viruses. Trends Genet 17: 449–459
8. Lavelle C (2007) Transcription elongation through a chromatin template. Biochimie 89: 516–527
9. Gosse C, Croquette V (2002) Magnetic tweezers: micromanipulation and force measurement at the molecular level. Biophys J 82: 3314–3329
10. Ajjan R, Lim BCB, et al. (2008) Common variation in the C-terminal region of the fibrinogen beta-chain: Effects on fibrin structure, fibrinolysis and clot rigidity. Blood 111: 643–650
11. Mierke CT, Kollmannsberger P, et al. (2008) Mechano-coupling and regulation of contractility by the vinculin tail domain. Biophys J 94: 661–670
12. Shang H, Lee GU (2007) Magnetic tweezers measurement of the bond lifetime–force behavior of the IgG-protein a specific molecular interaction. J Am Chem Soc 129: 6640–6646
13. Shang H, Kirkham PM, et al. (2005) The application of magnetic force differentiation for the measurement of the affinity of peptide libraries. J Magn Magn Mater 293: 382–388
14. Lee GU, Metzger S, et al. (2000) Implementation of force differentiation in the immunoassay. Anal Biochem 287: 261–271
15. Smith AS, Sengupta K, et al. (2008) Force-induced growth of adhesion domains is controlled by receptor mobility. Proc Natl Acad Sci USA 105: 6906–6911
16. Kanger JS, Subramaniam V, et al. (2008) Intracellular manipulation of chromatin using magnetic nanoparticles. Chromosome Res 16: 511–522
17. Tanase M, Biais N, et al. (2007) Magnetic tweezers in cell biology. In *Cell Mechanics*, Methods is cell Biology 83: 473–493
18. Bausch AR, Moller W, et al. (1999) Measurement of local viscoelasticity and forces in living cells by magnetic tweezers. Biophys J 76: 573–579
19. Pincet F, Husson J (2005) The solution to the streptavidin–biotin paradox: The influence of history on the strength of single molecular bonds. Biophys J 89: 4374–4381
20. Merkel R, Nassoy P, et al. (1999) Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. Nature 397: 50–53
21. Dammer U, Hegner M, et al. (1996) Specific antigen/antibody interactions measured by force microscopy. Biophys J 70: 2437–2441
22. Moy VT, Florin EL, et al. (1994) Intermolecular forces and energies between ligands and receptors. Science 266: 257–259
23. Lee GU, Kidwell DA, et al. (1994) Sensing discrete streptavidin–biotin interactions with atomic-force microscopy. Langmuir 10: 354–357
24. Chiou CH, Huang YY, et al. (2006) New magnetic tweezers for investigation of the mechanical properties of single DNA molecules. Nanotechnology 17: 1217–1224
25. Todd BA, Rau DC (2008) Interplay of ion binding and attraction in DNA condensed by multivalent cations. Nucleic Acids Res 36: 501–510
26. Neuman KC, Lionnet T, et al. (2007) Single-molecule micromanipulation techniques. Ann Rev Mater Res 37: 33–67
27. Cheezum MK, Walker WF, et al. (2001) Quantitative comparison of algorithms for tracking single fluorescent particles. Biophys J 81: 2378–2388
28. Gelles J, Schnapp BJ, et al. (1988) Tracking kinesin-driven movements with nanometre-scale precision. Nature 331: 450–453
29. Charvin G, Allemand JF, et al. (2004) Twisting DNA: Single molecule studies. Contemp Phys 45: 383–403
30. Wong WP, Halvorsen K (2006) The effect of integration time on fluctuation measurements: calibrating an optical trap in the presence of motion blur. Opt Express 14: 12517–12531
31. de Grooth BG (1999)A simple model for Brownian motion leading to the Langevin equation. Am j phys 67: 1248–1252
32. Bustamante C, Marko JF, et al. (1994) Entropic elasticity of lambda-phage DNA. Science 265: 1599–1600
33. Smith SB, Finzi L, et al. (1992) Direct mechanical measurements of the elasticity of single DNA-molecules by using magnetic beads. Science 258: 1122–1126
34. Bouchiat C, Wang MD, et al. (1999) Estimating the persistence length of a worm-like chain molecule from force-extension measurements. Biophys J 76: 409–413

35. Marko JF, Siggia ED (1995) Stretching DNA. Macromolecules 28: 8759–8770
36. Vologodskii A (1994) DNA extension under the action of an external force. Macromolecules 27: 5623–5625
37. Nelson P (2003) *Biological Physics: Energy, Information, Life*. W. H. Freeman, New York
38. Abels JA, Moreno-Herrero F, et al. (2005) Single-molecule measurements of the persistence length of double-stranded RNA. Biophys J 88: 2737–2744
39. Leger JF, Romano G, et al. (1999) Structural transitions of a twisted and stretched DNA molecule. Phys Rev Lett 83: 1066–1069
40. Smith SB, Cui YJ, et al. (1996) Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. Science 271: 795–799
41. White JH (1969) Self-linking and Gauss-integral in higher dimensions. Am J Math 91: 693–728
42. Forth S, Deufel C, et al. (2008) Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. Phys Rev Lett 100: 4
43. Strick T, Dessinges M-N, Charvin G, Dekker NH, Allemand J-F, Bensimon D, Croquette V (2003) Stretching of macromolecules and proteins. Rep Prog Phys 66: 1–45
44. Strick TR, Allemand JF, et al. (2000) Stress-induced structural transitions in DNA and proteins. Annu Rev Biophys Biomolec Struct 29: 523–543
45. Marko JF (2007) Torque and dynamics of linking number relaxation in stretched supercoiled DNA. Phys Rev E 76: 021926
46. Clauvelin N, Audoly B, et al. (2008) Mechanical response of plectonemic DNA: An analytical solution. Macromolecules 41: 4479–4483
47. Neukirch S (2004) Extracting DNA twist rigidity from experimental supercoiling data. Phys Rev Lett 93: 4
48. Moroz JD, Nelson P (1997) Torsional directed walks, entropic elasticity, and DNA twist stiffness. Proc Natl Acad Sci USA 94: 14418–14422
49. Bouchiat C, Mezard M (1998) Elasticity model of a supercoiled DNA molecule. Phys Rev Lett 80: 1556–1559
50. Vologodskii AV, Marko JF (1997) Extension of torsionally stressed DNA by external force. Biophys J 73: 123–132
51. Neuman KC, Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nat Methods 5: 491–505
53. Lipfert J, Hao X, et al. (2009) Quantitative Modeling and Optimization of Magnetic Tweezers. Biophys j 96(12): in press
53. Griffiths DJ (1999) *Introduction to Electrodynamics*. Benjamin Cummings, Menlo Park, NJ
54. Derks RJS, Dietzel A, et al. (2007) Magnetic bead manipulation in a sub-microliter fluid volume applicable for biosensing. Microfluid Nanofluidics 3: 141–149
55. Vilfan ID, Kamping W, et al. (2007) An RNA toolbox for single-molecule force spectroscopy. Nucleic Acids Res 35: 6625–6639
56. Lambert MN, Vocker E, et al. (2006) Mg^{2+}-induced compaction of single RNA molecules monitored by tethered particle microscopy. Biophys J 90: 3672–3685
57. Gore J, Bryant Z, et al. (2006) DNA overwinds when stretched. Nature 442: 836–839
58. Dekker NH, Abels JA, et al. (2004) Joining of long double-stranded RNA molecules through controlled overhangs. Nucleic Acids Res 32: e140
59. Liphardt J, Onoa B, et al. (2001) Reversible unfolding of single RNA molecules by mechanical force. Science 292: 733–737
60. Bakin AV, Borisova OF, et al. (1991) Spatial-organization of template polynucleotides on the ribosome determined by fluorescence methods. J Mol Biol 221: 441–453
61. Hansske F, Cramer F (1979) Modification of the 3′ terminus of tRNA by periodate oxidation and subsequent reaction with hydrazides. In: K. Moldave and L. Grossman (eds.), *Methods in Enzymology*, ed. Academic Press, New York, Vol. 59, pp. 172–181.
62. Willkomm DK, Hartmann RK (2005) 3′-terminal attachment of fluorescent dyes and biotin. In: Hartmann RK, et al. (eds.), *Handbook of RNA Biochemistry*, 1st ed. Wiley-VCH, Weinheim, Germany pp. 86–94.
63. Kinoshita Y, Nishigaki K, et al. (1997) Fluorescence-, isotope- or biotin-labeling of the 5′-end of single-stranded DNA/RNA using T4 RNA ligase. Nucleic Acids Res 25: 3747–3748
64. Martin G, Keller W (1998) Tailing and 3′-end labeling of RNA with yeast poly(A) polymerase and various nucleotides. RNA Publ RNA Soc 4: 226–230
65. Rosemeyer V, Laubrock A, et al. (1995) Nonradioactive 3′-end-labeling of RNA molecules of different lengths by terminal deoxynucleotidyltransferase. Anal Biochem 224: 446–449
66. Davenport RJ, Wuite GJ, et al. (2000) Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase. Science 287: 2497–2500

67. Mangeol P, Cote D, et al. (2006) Probing DNA and RNA single molecules with a double optical tweezer. Eur Phys J E 19: 311–317
68. Koster DA, Croquette V, et al. (2005) Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. Nature 434: 671–674
69. Koster DA, Palle K, et al. (2007) Antitumor drugs impede DNA uncoiling by Topoisomerase I. Nature 448: 213–217
70. Koster DA, Wiggins CH, et al. (2006) Multiple events on single molecules: unbiased estimation in single-molecule biophysics. Proc Natl Acad Sci USA 103: 1750–1755
71. Bonin M, Oberstrass J, et al. (2000) Determination of preferential binding sites for anti-dsRNA antibodies on double-stranded RNA by scanning force microscopy. RNA Publ RNA Soc 6: 563–570
72. Thomson NH, Kasas S, et al. (1996) Reversible binding of DNA to mica for AFM imaging. Langmuir 12: 5905–5908
73. Fulconis R, Mine J, et al. (2006) Mechanism of RecA-mediated homologous recombination revisited by single molecule nanomanipulation. EMBO J 25: 4293–4304
74. Kruithof M, Chien F, et al. (2008) Subpiconewton dynamic force spectroscopy using magnetic tweezers. Biophys J 94: 2343–2348
75. Crut A, Koster DA, et al. (2008) Controlling the surface properties of nanostructures for studies of polmerases. Nanotechnology 19: 465301
76. Skinner GM, Baumann CG, et al. (2004) Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase—A single-molecule view of the transcription cycle. J Biol Chem 279: 3239–3244
77. Sorrentino S (1998) Human extracellular ribonucleases: Multiplicity, molecular diversity and catalytic properties of the major RNase types. Cell Mol Life Sci 54: 785–794
78. Blasko A, Bruice TC (1999) Recent studies of nucleophilic, general-acid, and metal ion catalysis of phosphate diester hydrolysis. Acc Chem Res 32: 475–484
79. Kaga E, Nakagomi O, et al. (1992) Thermal-degradation of RNA–RNA hybrids during hybridization in solution. Mol Cell Probes 6: 261–264
80. Tenhunen J (1989) Hydrolysis of single-stranded RNA in aqueous solutions— Effect on quantitative hybridizations. Mol Cell Probes 3: 391–396
81. Butzow JJ, Eichhorn GL (1975) Different susceptibility of DNA and RNA to cleavage by metal ions. Nature 254: 358–359
82. Eichhorn GL, Clark P, et al. (1969) Interaction of metal ions with polynucleotides and related compounds. 13. Effect of metal ions on enzymatic degradation of ribonucleic acid by bovine pancreatic ribonuclease and of deoxyribonucleic acid by bovine pancreatic deoxyribonuclease I. J Biol Chem 244: 937–942
83. Khodursky AB, Peter BJ, et al. (2000) Analysis of topoisomerase function in bacterial replication fork movement: use of DNA microarrays. Proc Natl Acad Sci USA 97: 9419–9424
84. Crut A, Koster DA, et al. (2006) Fast dynamics of supercoiled DNA revealed by single-molecule experiments. Proc Natl Acad Sci USA 104: 11957–11926
85. Revyakin A, Ebright RH, et al. (2004) Promoter unwinding and promoter clearance by RNA polymerase: Detection by single-molecule DNA nanomanipulation. Proc Natl Acad Sci USA 101: 4776–4780
86. Revyakin A, Liu C, et al. (2006) Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. Science 314: 1139–1143
87. Maier B, Bensimon D, et al. (2000) Replication by a single DNA polymerase of a stretched single-stranded DNA. Proc Natl Acad Sci USA 97: 12002–12007
88. Dessinges MN, Lionnet T, et al. (2004) Single-molecule assay reveals strand switching and enhanced processivity of UvrD. Proc Natl Acad Sci USA 101: 6439–6444
89. Henn A, Medalia O, et al. (2001) Visualization of unwinding activity of duplex RNA by DbpA, a DEAD box helicase, at single-molecule resolution by atomic force microscopy. Proc Natl Acad Sci USA 98: 5007–5012
90. Lionnet T, Spiering MM, et al. (2007) Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism. Proc Natl Acad Sci USA 104: 19790–19795
91. Ribeck N, Saleh OA (2008) Multiplex single-molecule measurements with magnetic tweezers. Rev Sci Instrum 79: 094301

14

Folding of Proteins under Mechanical Force

Michael Schlierf and Matthias Rief

Abstract Many proteins in the body are subject to mechanical forces in their natural context. Examples are muscle proteins or proteins of the cytoskeleton. Often a protein faces the challenge of refolding against a mechanical load. Although the unfolding of proteins under load has been extensively investigated, knowledge about refolding mechanics is still rare. This chapter develops a model that describes the effect of an external force on protein folding. The model can provide important help for the design of a single-molecule mechanical experiment. The chapter discusses how spacer length and elasticity, as well as probe spring constant, affect the observed results. In this context it also briefly discusses the difference between atomic force microscope and optical tweezers experiments.

14.1. A Model for Protein Folding under Force

When a protein is subject to mechanical force while it tries to fold its polypeptide chain into a compact structure, it is obvious that the force will strongly influence its folding kinetics. The influence of the force on the folding kinetics can be quantified by considering that, in addition to its natural folding barrier, an additional barrier will be imposed by load and will limit the kinetics of the folding process. In the following, we consider a model in which the barrier imposed by load and the natural folding barrier are additive. Figure 14.1 shows changes in force and energy of an unfolded polypeptide chain when part of the chain (the "folder," green) tries to fold into a compact structure while held by a cantilever of an atomic force microscope (AFM) while the rest of the chain ("spacer," blue) stays unfolded. The force–extension curve has been shown to follow closely the worm-like chain (WLC) model of polymer elasticity [1]. The red trace in Figure 14.1 shows the fully unfolded polypeptide, consisting of the unstructured folder and the spacer. We start with the fully extended chain at high loads. Now, the external force is reduced successively from higher to lower forces (red arrow). At point A, the force is so low that spontaneous folding of the protein occurs. Because folding of the folder is associated with a significant reduction in length, the total contour length is reduced by the length of the unfolded folder (green WLC curve), and the acting force on the cantilever is increased to point B. The final polypeptide elasticity can be described by the WLC for the spacer (blue). The top part of Figure 14.1 also illustrates

M. Schlierf, M. Rief • Physikdepartment E22, TU München, James-Franck-Strasse, 85748 Garching, Germany

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_14,

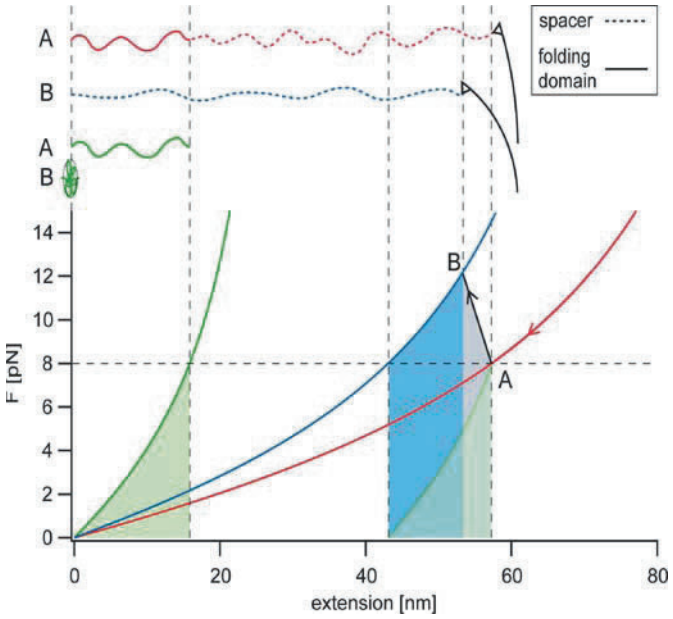


Figure 14.1. Additional folding barrier under force. **Top.** Prefolding (A) and postfolding (B) snapshots of the whole system. **Bottom.** When the external force is reduced, the folding domain and the spacer follow the prefolding trace (*red*) until the folding force is reached (A). After folding, the spacer dominates the postfolding trace (*blue*). The force–extension behavior of the folding domain is illustrated in green. The additional barrier consists of the blue shaded area and the gray shaded area reduced by the green shaded area.

prefolding (A) and postfolding (B) snapshots of the whole system. Before folding, the total contour length consists of the folding domain and the spacer. After the folding, the contour length is only the spacer. The elasticity of the folded protein is not considered because its spring constant is orders of magnitude higher than that of the polypeptide spacer. The total energy needed to fold at force F_A can be divided into three contributions: the energy loss due to the cantilever bending $E_\mathrm{L}(F_\mathrm{A})$ (grey area), the energy loss due to stretching the remaining spacer to higher forces $E_\mathrm{S}(F_\mathrm{A})$ (blue area), and the energy gain due to the collapse of the folder $E_\mathrm{domain}(F_\mathrm{A})$ (green area). In detail each contribution can be calculated by the following equations.

The energy needed to bend the cantilever can be calculated from

$$E_L(F_A) = 0.5[l_A(F_A) - l_B(F_B)](F_A + F_B) \tag{14.1}$$

where $l_\mathrm{A}(F_\mathrm{A})$ denotes the extension on the initial WLC at the initial force F_A (at this force, refolding would happen) and $l_\mathrm{B}(F_\mathrm{B})$ is the extension at the final force F_B on the final WLC curve reached after folding.

The energy needed to stretch the polypeptide spacer against its entropic elasticity is calculated from

$$E_\mathrm{S}(F_\mathrm{A}) = \int_{s_\mathrm{A}}^{s_\mathrm{B}} \frac{k_\mathrm{B}T}{p}\left(\frac{s}{L_s} + \frac{1}{4(1-s/L_s)^2} - \frac{1}{4}\right) ds \tag{14.2}$$

where $s_B = s_B(F_B)$ denotes the spacer's final extension on the polypeptide WLC curve, $s_A = s_A(F_A)$ is the spacer's initial extension on its WLC curve, $k_B T$ is the thermal energy, p is the persistence length of the polypeptide spacer, and L_S the contour length of the polypeptide spacer.

Accordingly, we can calculate the energy liberated when the folder collapses from the stretched to the folded state. We can calculate this entropic energy gain by

$$E_{\text{domain}}(F_A) = \int_{d_A}^{d_B} \frac{k_B T}{p} \left(\frac{d}{L_d} + \frac{1}{4(1 - d/L_d)^2} - \frac{1}{4} \right) dd \tag{14.3}$$

where d_B denotes the domain's final extension on its WLC curve, which in good approximation is zero because the size of the folded folder is much smaller than the unfolded folder. Here $d_A = d_A(F_A)$ is the initial extension and L_d is the contour length of the domain. It is important to note that the integral in Eq. (14.3) is negative because the integration is from a higher limit to a lower limit. This reflects the fact that this energy contribution actually favors folding. The total energy barrier at the initial force F_A can be calculated by summing up:

$$E_{\text{total}}(F_A) = E_L(F_A) + E_S(F_A) + E_{\text{domain}}(F_A) \tag{14.4}$$

This total barrier is illustrated in Figure 14.1 by the sum of the gray area and the blue area, subtracting the green area.

However, because the total barrier is calculated using WLC elasticity, its dependence on external force is not simple. Figure 14.2A shows the total energy barrier and its individual energetic contributions as a function of the external force under which folding occurs. For all calculations, if not stated otherwise, a persistence length $p = 0.5$ nm, a spacer contour length

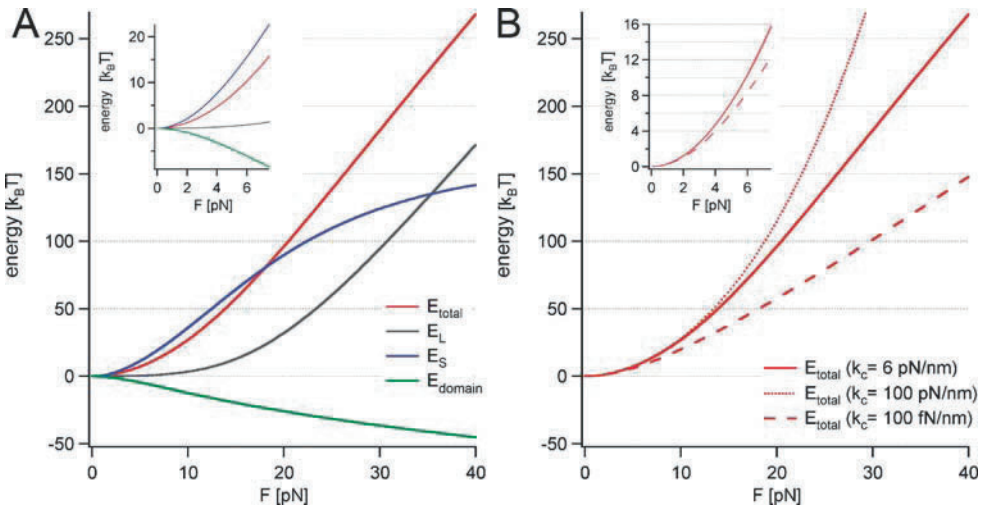


Figure 14.2. Energetic contributions at various external forces. **A.** The total energy barrier in a typical atomic force microscope (AFM) experiment is dominated by the spacer elasticity under low forces and the cantilever bending energy under high forces. **Inset.** In the experimental relevant force regime, the cantilever bending contribution is nearly negligible. **B.** Dependence of the total energy barrier on various AFM cantilever spring constants.

of $L_S = 100$ nm, a contour length of the folding domain $L_d = 25$ nm, and a cantilever spring constant $k_C = 6$ pN/nm are assumed.

Under low forces, the total energy is dominated by stretching the remaining spacer. At high external forces ($F > 35$ pN) the bending energy of the cantilever becomes the dominant contribution. Nevertheless, at such high forces, the additional barrier has already increased to more than $100k_BT$, and hence this force regime is likely not relevant for real proteins. The inset in Figure 14.2A shows a zoom to the low-force regime, where one can see that the total energy is strongly dependent on the entropic energy loss while stretching the spacer. A change of the cantilever spring constant $k_C = 0.1$, 6, and 100 pN/nm only has a strong effect at higher forces (Figure 14.2B). In the low-force regime (inset), no differences between a cantilever spring constant of 6 or 100 pN/nm can be observed. Even a spring constant of 0.1 pN/nm lowers the barrier by only $2k_BT$ at forces of 6 pN.

To transform the barrier height into an effective folding rate under a certain force, one can assume a simple Arrhenius equation:

$$k_f^{\text{eff}}(F_A) = k_f(F = 0) \ \exp\left(\frac{-E_{\text{total}}}{k_BT}\right) \tag{14.5}$$

where $k_f(F = 0)$ denotes the force-free folding rate of the protein of interest. A difference of the additional barrier in $2k_BT$ would result in an ~7.4 times increased effective rate. Hence, with softer probe springs, refolding events will be observed at slightly higher forces. In the following, we discuss experimental limitations for AFM and optical tweezers experiments on protein folding under force.

14.2. Protein Refolding at Constant Pulling Velocity

The preceding paragraph introduced the calculation of the additional energy barrier under force [Eq. (14.4)] and the resulting effective folding rate [Eq. (14.5)]. Figure 14.3A illustrates the effective folding rate under force. A force-free folding rate of $k_f(F = 0) = 100$ sec^{-1} and the previously used parameters $L_S = 100$ nm, $L_d = 25$ nm, $p = 0.5$ nm, and $k_C = 6$ pN/nm were assumed. If one imposes an external force of 5.5 pN on the folding protein, fewer than one transition per 60 sec will be observed. Real AFM experiments, however, often occur at constant pulling velocities rather than constant force. So far, we only have calculated refolding rates as a function of load. To model experiments in which the force applied constantly changes, we need to calculate probability densities for refolding as a function of force. The theoretical folding force distribution $dP_N/dF(F)$ giving the probability density of observing a folding event at a given force F can be derived as follows: For a two-state system $U \underset{k_u}{\overset{k_f}{\rightleftarrows}} N$, the probability of finding the protein in the native state N at a certain time t, $P_N(t)$, is calculated by solving the following differential equation:

$$\frac{dP_N}{dt}(t) = k_fP_U(t) - k_uP_N(t) = k_f - P_N(t)(k_f + k_u) \tag{14.6}$$

with $P_N(t) + P_U(t) = 1$. One possible solution of this equation is

$$P_N(t) = \left[P_N(0) - \frac{k_f}{k_f + k_u}\right] \exp(-k_ft - k_ut) + \frac{k_f}{k_f + k_u} \tag{14.7}$$

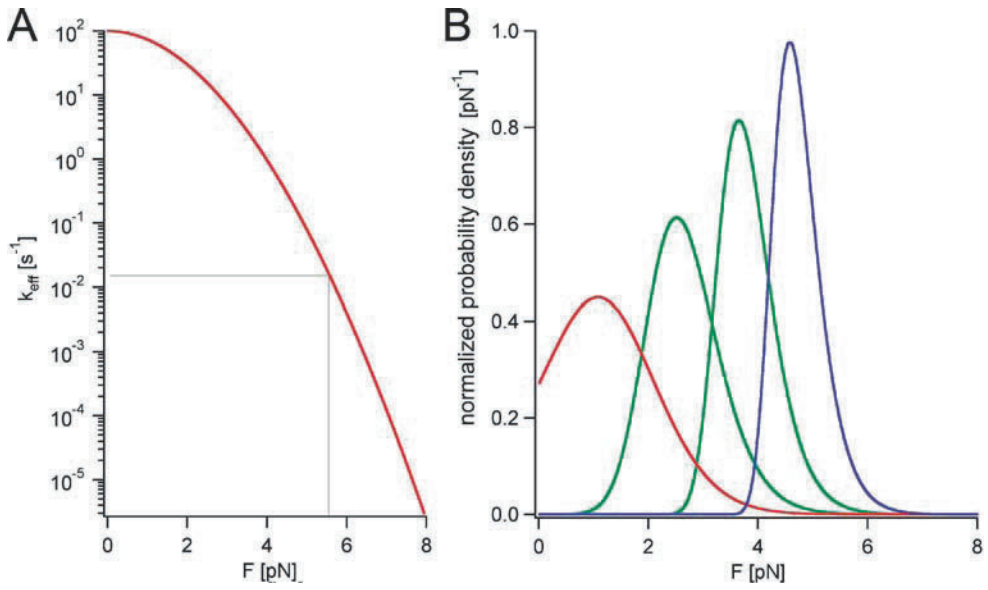


Figure 14.3. Folding rate and probability under force. **A.** The effective folding rate under external force drops more strongly than as a single exponential. At $F = 5.5$ pN fewer than one transition per 60 sec can be observed for a folder with a force-free folding rate $k_f(F = 0) = 100\ \text{sec}^{-1}$. **B.** Expected folding force probability distributions for various pulling velocities: $v_P = 0.1$ (*blue*), 1 (*dark green*), 10 (*green*), and 100 nm/sec (*red*).

Under low-force conditions, one can simplify this equation with the assumptions that the unfolding rate $k_u << k_f$, $P_N(0) = 0$, and under changing force conditions there is a time-dependent folding rate $k_f(t)$:

$$P_N(t) = 1 - \exp\left[-\int_0^t k_f(t')dt'\right] \tag{14.8}$$

Because Eq. (14.5) does not provide access to a time-dependent folding rate $k_f(t)$ but a force-dependent folding rate, a coordinate transformation $t \rightarrow F$ with $dt = \frac{dF}{\dot{F}}$ gives the force-dependent probability to be in the native state $P_N(F)$:

$$P_N(F) = 1 - \exp\left[-\int_F^0 k_f^0 \exp\left(\frac{-E_{total}(F')}{k_B T}\right) \cdot \frac{1}{\dot{F}'} dF'\right] \tag{14.9}$$

with the lower integration boundary F to be chosen such that $k_f(F) \approx 0$. One final differentiation results in the desired probability distribution $dP_N/dF(F)$:

$$\frac{dP_N}{dF}(F) = \frac{k_f^0}{\dot{F}} \exp\left(\frac{E_{total}(F)}{k_B T}\right) \exp\left[\int_0^F \frac{k_f^0}{\dot{F}'} \exp\left(\frac{-E_{total}(F')}{k_B T}\right) dF'\right] \tag{14.10}$$

The loading rate $\dot{F}$ depends on the pulling velocity, the polypeptide/spacer elasticity, and the cantilever spring constant. It is important to note that this loading rate is negative because

the force is lowered with time. In contrast to increasing unfolding force distributions with increasing pulling velocity, the folding force distribution decreases with increasing pulling velocity.

For the effective folding rate of Figure 14.3A, Figure 14.3B shows the folding force distributions at four different pulling velocities $v_P = 0.1$, 1, 10, and 100 nm/sec. These distributions illustrate the probability densities of the initial forces F_A. Because refolding forces increase with decreasing pulling velocity, low pulling velocities are typically desirable. Generally, the observable folding force depends on the following parameters:

- force-free folding rate k_f^0
- pulling velocity
- contour length of the folding domain L_d (number of amino acids)
- spring constant of the cantilever
- elasticity of the spacer (contour and persistence length)

First, we discuss the dependency of the folding force on the pulling velocity and contour length of the folding domain for various force-free folding rates. The next paragraph will then focus on the dependency on the spacer elasticity and cantilever spring constant. Figure 14.4A illustrates the maximum force of the refolding probability density $dP_N/dF(F)$ for four typical folding rates of proteins $k_f^0 = 1$, 10, 100, and 1000 sec^{-1} as a function of the pulling velocity. All previously assumed parameters $L_S = 100$ nm, $p = 0.5$ nm, $k_C = 6$ pN/nm, and $L_d = 25$ nm were kept constant. Clearly, the active folding force drops more strongly than as a single exponential and, for slow and medium-fast folders, at velocities of tenths of nm/sec it already reaches the typical experimental resolution limit of 1–2 pN.

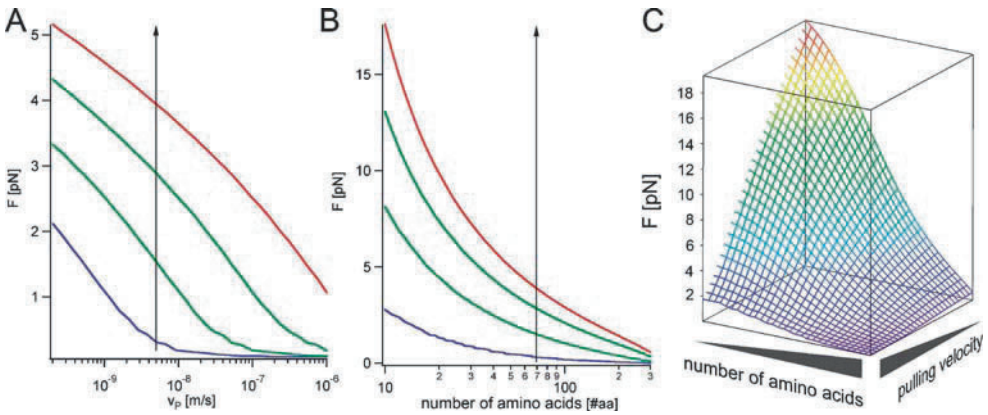


Figure 14.4. Expected folding forces as function of **(A)** the pulling velocity and **(B)** the number of folding amino acids for four typical force-free protein folding rates $k_f = 1$ (*blue*), 10 (*dark green*), 100 (*green*), and 1,000 sec^{-1} (*red*). (C) Illustration of the maximum folding force surface for a force-free folding rate $k_f = 100\ s^{-1}$. Both, increasing number of folding amino acids and increasing pulling velocity decreases the observable folding force rapidly.

The maximum of $dP_N/dF(F)$ as a function of the number of folding amino acids (aa) ($L_d = 0.365$ nm/aa $\times$ number of aa), keeping the pulling velocity constant at $v_P = 5$ nm/sec, is illustrated in Figure 14.4B. We realize that short folding units of only 10 amino acids can easily exceed 5- to 10-pN active folding forces. However, most small proteins contain at least 50 amino acids. As illustrated in Figure 14.4B, at 50−amino acids contraction length, this

model predicts folding forces hardly exceeding 5 pN if the folding rate is within the range of 1 (blue)−1,000 sec^{-1} (red). Both calculations of Figure 14.4A, B overlap at the upright arrows and span a folding force surface. Such a surface is illustrated in Figure 14.4C for a force-free folding rate $k_f^0 = 100\ sec^{-1}$. Both increasing the pulling velocity and increasing the number of folding amino acids decrease the folding force rapidly. The highest refolding forces greater than 18 pN (red part of the surface) are only attainable for pulling velocities less than 1 nm/sec and only 10 contracting amino acids ($L_d \approx 4$ nm). A folding rate of $k_f^0 = 100\ sec^{-1}$ can be considered already fast for most proteins. Therefore, the folding force surface of Figure 14.4C gives a realistic upper estimate for observable active protein folding forces (the calculated forces are the starting forces F_A and not the final forces F_B). Higher refolding forces are only to be expected for extremely fast folders, that is, *down-hill folders* [2].

14.3. Comparing AFM and Optical Tweezers Experiments

As an alternative to AFM experiments, optical tweezers also offer the possibility of observing protein folding under force [3,4]. In this section, differences in the experimental conditions, such as a different spacer contour and persistence length or a lower spring constant, and their effects on the mechanical folding barrier will be discussed. Typical dumbbell optical tweezers experiments attach the protein of interest with two ∼200-nm-long double-stranded DNA (dsDNA) spacers to the beads [5]. Therefore, the hybrid system consists of dsDNA with $L_S = 400$ nm and a polypeptide with a contour length of $L_d = 25$ nm. In this case, the higher persistence length of the dsDNA dominates the elasticity. Figure 14.5A shows several force–extension curves of a pure polypeptide system (blue curve) (AFM) and two hybrid systems (optical tweezers). The persistence length of dsDNA is commonly believed to be $p = 50$ nm (red curve). However, shorter dsDNA fragments attached to rotating beads are described as exhibiting a smaller persistence length [6,7]. A sample force–extension trace with $p = 8$ nm is illustrated in Figure 14.5A (green curve). In the following calculations we mostly assume a persistence length of $p = 8$ nm. After folding of the domain, the elasticity is truly spacer dependent (dashed lines in Figure 14.5A). Another significant difference between optical tweezers experiments and AFM experiments is the softer probe spring constant of $k_{o.t.} \approx 100$ fN/nm. Figure 14.5B shows the total additional barrier and its components for the described hybrid system. Unlike in the AFM case, the energy contribution from displacing the trapped beads already dominates the total barrier in the low-force regime. For a comparison, the energy barrier with a 400-nm-long polypeptide spacer and a typical AFM cantilever spring constant of $k_C = 6$ pN/nm is plotted as a black thin line. The top part of Figure 14.5B shows the difference of the barrier height between the two systems. Up to 10 pN both barriers are nearly identical. There is a small energetic advantage for the AFM polypeptide system between 3 and 8 pN. At higher forces, the optical tweezers setup clearly has the advantage (see inset in Figure 14.5B). A refolding experiment at $v_P = 5$ nm/sec would generate the probability distributions illustrated in Figure 14.5C. The small energetic advantage of the polypeptide-only system would result in a probability distribution (black) shifted by +0.5 pN in comparison to the optical tweezers experiments. The long polypeptide spacer obviously increases the observable refolding force in AFM experiments. However, the strong entropic elasticity of the polypeptide-only spacer will have a dampening effect on the magnitude of the observed force jump on folding. The force jump is the difference between the starting force F_A and the final force F_B: $\Delta F = F_B - F_A$. Moreover, mostly due to drift problems, the force resolution of AFM experiments is still fairly limited. Recently proposed lock-in detection

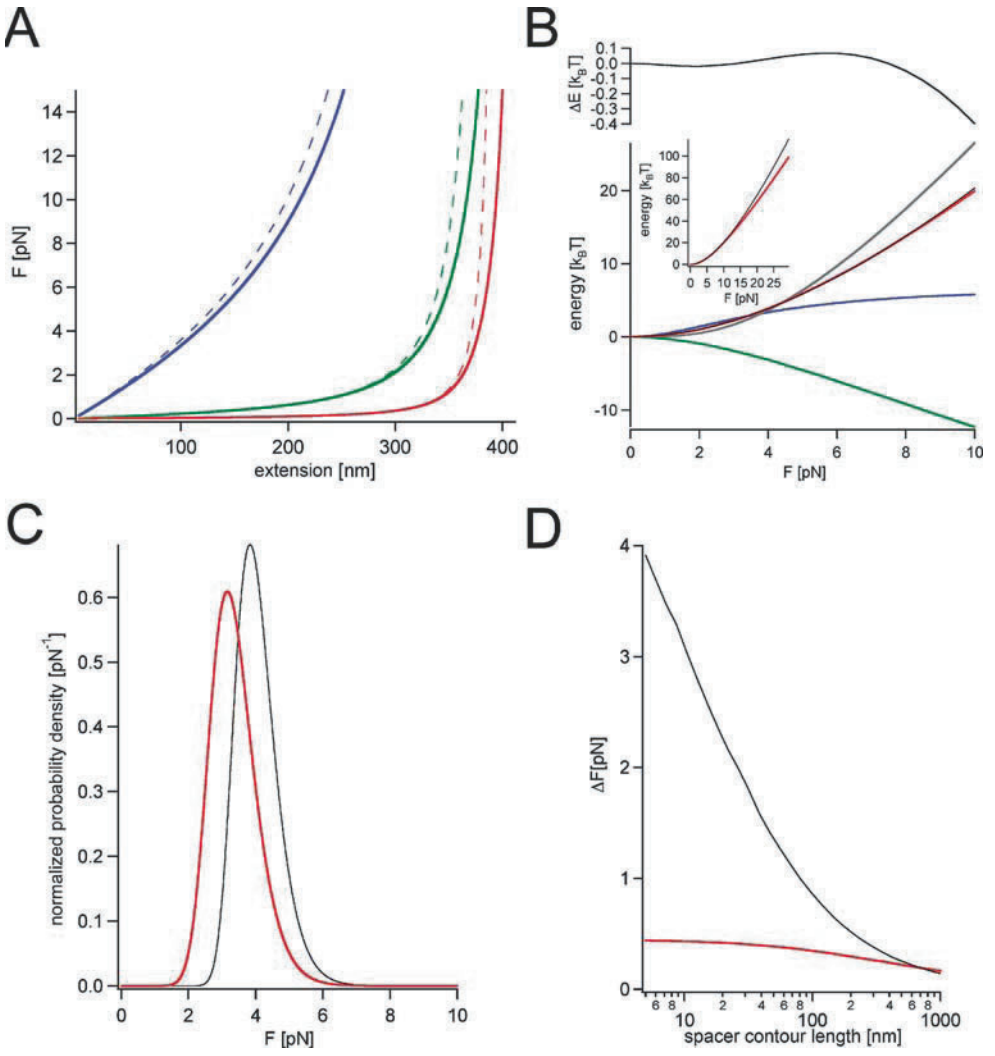


Figure 14.5. Comparison between pure polypeptide and hybrid systems. **A.** Force–extension curves for a pure polypeptide (*blue*) and two hybrid systems with $p_{\mathrm{dsDNA}} = 8$ nm (*green*) and $p_{\mathrm{dsDNA}} = 50$ nm. **B.** Energetic contributions of the probe displacement (*gray*), the entropic stretching of the spacer (*blue*), and the energy gain of the folder (*green*) to the total energy barrier (*red*) for a hybrid system with $p_{\mathrm{dsDNA}} = 8$ nm as a function of the external force. In comparison, the total energy barrier for a pure polypeptide system with the same contour length and an atomic force microscope typical spring constant is shown in black. The top graph shows the marginal energetic difference ΔE between a hybrid and a comparable polypeptide system. **Inset.** Only higher forces favor the hybrid system. **C.** Refolding force distribution of a hybrid system (*red*) and the compared pure polypeptide system (*black*). **D.** Force jump of the hybrid system (*red*) and a pure polypeptide system (*black*) as function of the spacer length.

schemes may offer a way around this problem [1]. Therefore, for an unequivocal identification of a refolding event in AFM experiments, a higher ΔF is desirable and shorter spacer lengths of 50–100 nm are preferred (Figure 14.5D). Due to their small friction, the force resolution of optical tweezers is often significantly better than in AFM experiments. Therefore, these

experiments do not require a high force jump. In optical tweezers experiments the force jump is less dependent on the spacer length due to its higher persistence length (Figure 14.5D).

14.4. Comparison to Experimental Data and Conclusion

So far, only few studies have reached the necessary resolution to measure refolding forces of proteins quantitatively. Those examples include thermodynamically extremely stable proteins such as membrane proteins, ankyrin, ubiquitin, and leucine zippers [8–12]. Active refolding of a topologically more complex protein has only recently been reported for RNase H using optical tweezers [3]. In a recent study, the refolding kinetics of ddFLN4 under external force was investigated [1]. Figure 14.6 shows the experimentally obtained refolding force distribution of the folding intermediate of ddFLN4. The expected refolding force distribution $dP_N/dF(F)$ calculated with the described model is plotted as black line. We assumed a folding of 60 amino acids with a force free folding rate of 55 sec^{-1} and a pulling velocity of 5 nm/sec. Due to an experimental uncertainty in the absolute force measurement, the theoretical distribution was convolved with a Gaussian with 0.5 pN width. In this case, we find excellent agreement between theory and experiment. The model presented here can provide an important basis for understanding and analyzing future experimental results on protein folding mechanics.

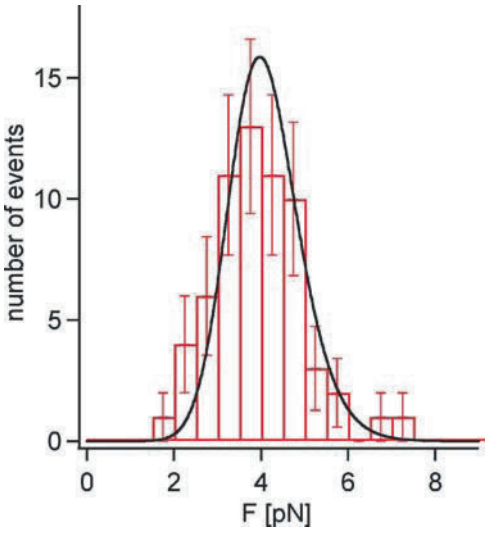


Figure 14.6. Folding force distribution of the folding intermediate of ddFLN4. The described model reproduces the experimental data with excellent agreement (*black line*).

References

1. Schlierf, M., F. Berkemeier, and M. Rief, Direct observation of active protein folding using lock-in force spectroscopy. Biophys J 2007;93(11): 3989–98.
2. Dyer, R. B., Ultrafast and downhill protein folding. Curr Opin Struct Biol 2007;17(1):38–47.

3. Cecconi, C., et al., Direct observation of the three-state folding of a single protein molecule. Science 2005;309(5743):2057–60.
4. Bechtluft, P., et al., Direct observation of chaperone-induced changes in a protein folding pathway. Science 2007;318(5855):1458–61.
5. Greenleaf, W. J., M. T. Woodside, and S. M. Block, High-resolution, single-molecule measurements of biomolecular motion. Annu Rev Biophys Biomol Struct 2007;36:171–90.
6. Seol, Y., et al., Elasticity of short DNA molecules: theory and experiment for contour lengths of 0.6–7 micron. Biophys J 2007;93(12):4360–73.
7. Wen, J.D., et al., Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. Biophys J 2007;92(9):2996–3009.
8. Kedrov, A., et al., Controlled unfolding and refolding of a single sodium-proton antiporter using atomic force microscopy. J Mol Biol 2004;340(5):1143–52.
9. Fernandez, J. M., and H. Li, Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science 2004;303(5664):1674–8.
10. Kessler, M., et al., Bacteriorhodopsin folds into the membrane against an external force. J Mol Biol 2006;357(2):644–54.
11. Lee, G., et al., Nanospring behaviour of ankyrin repeats. Nature 2006;440(7081):246–9.
12. Bornschlogl, T., and M. Rief, Single molecule unzipping of coiled coils: sequence resolved stability profiles. Phys Rev Lett 2006;96(11):118102.

15

Probing the Energy Landscape of Protein-Binding Reactions by Dynamic Force Spectroscopy

Andreas Ebner, Reinat Nevo, Christian Rankl, Johannes Preiner, Hermann Gruber, Ruti Kapon, Ziv Reich, and Peter Hinterdorfer

Abstract This chapter describes how the energy landscape that underlies protein-binding reactions can be revealed using dynamic force spectroscopy. The chapter begins with a detailed description of methodologies used and requirements of the experimental system, including tip and surface materials and their functionalization strategies. The next few sections discuss the fundamentals of measuring forces using the atomic force microscope, and the basics of performing force spectroscopy measurements from a practical point of view. Next, it presents an extensive account of methods for data analysis and current theoretical treatments. The remainder of the chapter illustrates the power of this methodology by several examples in which the location of energy barriers in a binding reaction pathway and their load-dependent dynamics are measured, the overall scale of roughness of the underlying energy surface is extracted, and alternative modes of protein activation are distinguished. Biological insight gained from these data is discussed. The intent is to provide the necessary theoretical and practical knowledge to begin force spectroscopy measurements on protein interactions.

15.1. Introduction

In this chapter we describe how the energy landscape that underlies a protein-binding reaction can be revealed using single-molecule force spectroscopy. The energy landscape of protein binding describes the way in which the free energy depends on structural parameters that describe the system, such as similarity to the bound state and the number of contacts made. Moreover, features of the energy surface define the dynamics of the reaction. The global minimum of such a surface represents the most stable state(s). Transition states

A. Ebner, C. Rankl, J. Preiner, H. Gruber, and P. Hinterdorfer • Institute for Biophysics, Johannes Kepler University of Linz, Altenbergerstr. 69, A-4040 Linz, Austria
R. Nevo, R. Kapon, and Z. Reich • Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel, 76100

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_15,

occur when local maxima are present, and intermediates are formed at local minima or saddle points. The reaction progresses along multiple routes taken by the reactants along the surface. Clearly, unravelling the shape of the energy surface that underlies a reaction is of major importance, and a great deal of effort has been placed on achieving this goal in recent years [1–8].

One inherent problem in unravelling an energy surface is the heterogeneity of both the ensemble of proteins that participate in a reaction and the environment in which the reaction proceeds. This heterogeneity necessitates the use of single-molecule methods when endeavouring to resolve more than the ensemble average of a particular quantity. Several methods have accordingly been used, including single-molecule optical methods such as fluorescence resonant energy transfer (FRET) and fluorescence correlation spectroscopy (FCS), which are described elsewhere in this book, and molecular dynamics simulations. The method that is the subject of this chapter, force spectroscopy, exploits the effects of exerting forces on protein pairs as a means of exploring the energy surface that underlies their association. As shall be explained in detail, when a force is applied to a protein complex, the energy surface is tilted along the direction of the force. By applying increasing forces, the energy surface is tilted further and further, effectively decreasing energy barriers, thereby making it easier for molecules to unbind through thermal fluctuations. The dependence of rupture forces on the rate at which force is applied is the force spectrum, which includes information on the number of barriers, their height, and their position along the reaction coordinate [9–12].

Force spectroscopy requires the application and detection of forces applied to single molecules or single molecular pairs. Several setups have been developed, namely, atomic force spectroscopy (AFM), biomembrane force probes (BFP), and optical tweezers. In this chapter we mostly limit ourselves to the use of AFM. In this technique the molecule of interest is sandwiched between a cantilevered tip and a surface that can be moved relative to each other in a precisely controlled manner, therefore exerting a force on the molecule or molecular pair. The movement is controlled by a piezoelectric transducer, and the position of the cantilever is detected with sub-nanometre accuracy by a photodiode detector.

A notable advantage of single-molecule force spectroscopy measurements over other single-molecule techniques is that the reaction coordinates reduce, to a large extent, to one dimension, with the path set primarily by the direction of the force to which it serves as a natural conjugate. Thus, a profile of the reaction can be reconstructed with a defined coordinate system, along which prominent features of the reduced energy surface can be allocated. In addition, force spectroscopy measurements provide a means of directly measuring bond lifetimes to probe rarely populated states and to extract mechanical properties by applying a ramp or a constant force. As discussed later in this chapter, mechanical properties play important roles in the function of certain proteins that are naturally subjected to external loads.

The structure of this chapter is as follows. We begin by providing detailed strategies and protocols for the biochemical functionalization of tips and surfaces (Section 15.2.1). Then we thoroughly discuss the experimental design and the nuts and bolts of measuring (Sections 15.2.2 and 15.2.3) and analyzing (Section 15.2.4) single intermolecular interaction forces. Following this we review theoretical considerations regarding the nature of these forces and their relation to kinetic chemical rate constants (Section 15.2.4), energy barriers (Section 15.2.5) and thermodynamic parameters (Section 15.2.6). In the next two sections (Sections 15.3.1 and 15.3.2), we discuss studies in which force spectroscopy was used to extract information on the location of activation energy barriers along the forced dissociation path, their load-dependent dynamics, and the overall scale of roughness of the underlying energy surface. The last section (Section 15.3.3) illustrates how this information can be used

to directly discriminate among alternative modes of protein activation that are otherwise very hard or impossible to distinguish. Even when limiting ourselves to these particular applications of force spectroscopy, we cannot provide an extensive coverage of the numerous excellent works that have been published on these topics. Instead we focus on a small number of studies, favouring cases in which several studies have been performed on the same type of interaction. In addition, the pioneering studies that were carried out at single loading rates are mentioned only in passing. We apologize to the many authors whose work could not be included in this chapter.

15.2. Dynamic Force Spectroscopy: Principles and Theory

15.2.1. Tip and Surface Immobilization

In the following section, established methods and reliable protocols of tip functionalization for molecular recognition force spectroscopy are described. In some cases simple physisorption is sufficient, although covalent chemistry is preferred in most cases. The direct functionalization of typical surfaces like silicon nitride or silicon is usually done in three steps: (1) aminofunctionalization, (2) unilateral attachment of a bifunctional cross-linker molecule, and (3) coupling of the ligand or receptor to the free end of the cross-linker. Alternatively, gold-coated tips may be used. The latter are usually functionalized by a coating with a self-assembling monolayer (SAM) containing ligands on the outer surface. Besides conventional tip functionalization, some rare approaches, such as tethering of whole cells to AFM tips [13,14] or using light tweezer-operated nanosphere sensors [15], also allow force spectroscopy experiments. In addition to tip functionalization with ligands, the corresponding receptor molecules must be mounted to the solid support, except when the receptors are embedded in native membranes of support-bound cells or organelles. Direct immobilization of isolated biomolecules on solid supports can be done in a similar way to tip functionalization chemistry. The following discussion presents an overview of tip and support functionalization with biomolecules.

Force spectroscopy experiments require a rather firm attachment of the ligand to the AFM tip and receptor molecules to the support (and vice versa) to ensure that rupture only occurs between ligand and receptor. Surprisingly, physisorption has been shown to fulfil this requirement under many circumstances. Nonspecific adsorption relies on noncovalent bonds such as hydrophobic interactions, stacking, hydrogen bonds, and electrostatic attraction. Although these interactions are rather weak in comparison to covalent bonds, biomolecules can form many such interactions, resulting in high overall adhesion forces. Typical tip materials such as silicon or silicon nitride are well suited for physisorptive binding of proteins. Because the tightness and irreversibility of the attachment depend significantly on the type of protein, methods allowing the binding of a well-suited protein in the first step are preferred. Avidin [16] and biotinylated bovine serum albumin (BSA) [17] are commonly used as addressable anchor proteins. Furthermore, cross-linking (e.g., with glutaraldehyde) improves the stability of the layer and allows higher pulling forces [18]. A layer of (strept)avidin on the tip allows for stable attachment of biotinylated protein such as concavalin A [17] or P-selectin [16]. Due to multiple biotinylation, these oligomeric proteins can safely be assumed to be simultaneously bound to adjacent (strept)avidin molecules, thereby forming a protein network with tight overall adhesion to the inorganic surface. These cross- or interlinking steps are important when the expected receptor–ligand unbinding force is in the range of the (strept)avidin–biotin interaction force (i.e., on the order of 20–400 pN) [19–23]. In most of the aforementioned studies the proteins were adsorbed directly to bare silicon or silicon nitride

AFM tips. Avidin and BSA are known to have positively charged clusters [24,25] that provide for tight attachment to the negatively charged deprotonated silanol groups of the oxide layer of silicon or silicon nitride due to high electrostatic interactions. Alternatively, tips can be precoated by using, for example, hydrophobic silanes [16,26] or positively charged aminosilane [27] to induce strong adsorption of virtually any kind of protein. By taking advantage of both hydrophobic interactions and electrostatic interactions as on aminofunctionalized surfaces, a stronger protein adsorption is obtainable [24] compared to bare silicon or silicon oxide tips. Similar approaches can be used for immobilization of the corresponding binding partners to flat support. In contrast to the typical tip materials glass, silicon, and silicon nitride, highly oriented pyrolytic graphite (HOPG) and mica are the most common materials used as supports because they are atomically flat and perfectly clean surfaces are easily generated by simple cleaving. HOPG is very hydrophobic and does not easily allow for covalent chemistry; thus, only physisorption of two-dimensional (2D) protein crystals [28] has been performed. In contrast, mica—probably the most important support material in AFM force spectroscopy—is hydrophilic and exhibits a negative surface potential, providing for firm binding of molecules exhibiting a positive net charge. Mica is a phyllosilicate, preferentially cleaved in the plane of K^+ ions, resulting in an atomically flat surface with a negative charge. This surface can be used for tight attachment of positively charged proteins [19,20]. The net negative charge of freshly cleaved mica can be inverted to a strongly positive surface potential if the mica-bound K^+ ions are exchanged for bivalent cations such as Mg^{2+} or Ni^{2+}, causing strong electrostatic adsorption of negatively charged molecules such as DNA [29,30] or whole viruses. In conclusion, physisorption of proteins can be a proper and easy method for binding molecules to the AFM tip or support. The major limitation here is the requirement that the receptor–ligand interaction has to be significantly weaker than the attachment of the anchor proteins. This limitation can easily be overcome by the use of covalent coupling chemistry.

Typical tip materials like silicon or silicon nitride are chemically rather inert. This inertness must be overcome by introducing more-reactive chemical functions. The most widely used method for this purpose is aminosilanization (Figure 15.1, middle and bottom parts) [19,20,22,31] or aminofunctionalization with ethanolamine-hydrochloride (Figure 15.1, top part) [19,20,32–35]. Only by using a proper silanization protocol it is possible to avoid having the silanized surface become sticky or rough, which would prevent its use in recognition force microscopy. Suitable aminosilanization protocols rely on vapour-phase deposition [23,36–43] or liquid-phase modification under rigorously dry conditions [20]. Most silanization reagents contain three methoxy or ethoxy functions (for binding to the inorganic support) and one specific side chain (e.g., aminopropyl in aminopropyl-triethoxysilane [APTES]). Premature hydrolysis of the methoxy/ethoxy function during storage yields large aggregates that have a much higher affinity for oxide surfaces than the intact monomers. For this reason, silanization reagents must be stored under rigorously dry conditions, and even then the storage time must not exceed a few months, after which time the silane must be replaced or redistilled. With vapour-phase deposition or a suitable protocol for liquid-phase reaction, a dense monolayer of silane monomers is formed on the oxide surface, followed by lateral condensation in the subsequent curing step. A well-established protocol for aminofunctionalization using APTES is the following [20]: The material is placed in a 6-L desiccator, which is filled with argon. Two small, open vials (e.g., the lid of 2-mL Eppendorf tubes), one with 60 μL of APTES and the other with 20 μL of triethylamine, are placed inside, the lid is closed, and the reaction is allowed to proceed for 2 hr. Subsequently, the vials have to be removed, and the desiccator is flooded extensively with argon. After a curing time of ∼2 days the samples are ready for use. Other coupling protocols with silanes, such as APTES,

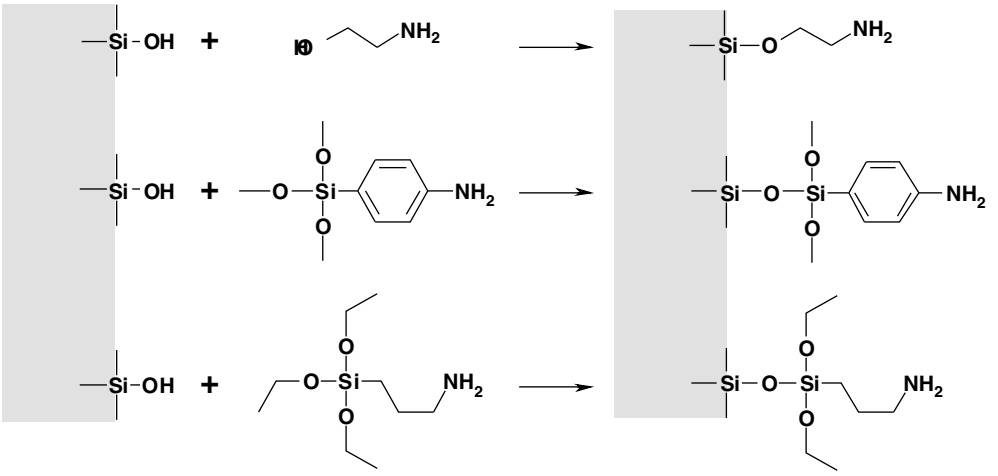


Figure 15.1. Schemes of aminofunctionalization chemistries using ethanolamine **(top)**, aminophenyltrimethoxysilane (APhS) **(middle)**, or 3-aminopropyltriethoxysilane (APTES) **(bottom)**. (From Ebner et al. [20].)

aminophenyltrimethoxysilane (APhS), or 3-aminopropyltriethoxysilane (APTES), using solvent procedures were also successfully used in molecular recognition force spectroscopy studies but require great care in handling to ensure dry conditions. A silane-free approach to tip aminofunctionalization uses ethanolamine, which is assumed to form a Si-O-C bond. Because the base form of ethanolamine is also known to catalyze the cleaving of such a bond, the salt ethanolamine-hydrochloride has been chosen, using dimethyl sulfoxide (DMSO) as solvent. Although the exact coupling mechanism is unclear, the aminofunctionalization using ethanolamine-hydrochloride is a very reliable and robust method for silicon, silicon nitride, mica, and glass. A well-established protocol for tip functionalization is the following: 6.6 g of ethanolamine hydrochloride is dissolved in 12 mL of DMSO by heating to ~90°C, molecular sieves are added, and after cooling to room temperature the mixture is degassed with aspirator vacuum for 30 min. The AFM cantilevers are placed in this solution overnight, rinsed with DMSO (3×), dried in a stream of nitrogen, rinsed in ethanol (3×), and again dried in a gentle stream of nitrogen. Aminofunctionalized tips should either be used within some days or stored under argon for up to several weeks to prevent oxidation of the primary amine groups. The described methods of aminofunctionalization can equally be used for supports such as mica, glass, silicon nitride, or silicon.

After aminofunctionalization a poly(ethylene glycol) (PEG) chain is usually attached via its amino-reactive terminus (in most cases an *N*-hydroxysuccinimide [NHS] ester). This reaction is mostly performed in chloroform, using 0.5% (v/v) triethylamine as catalyst. Typical PEG concentrations are 1 mg/mL (1 mM at molecular weight ~1,000 [44]), except when working with aldehyde-PEG-NHS [45]. Sometimes the PEG linker already carries the ligand molecules at its second end [19,40,41,46], thus functionalization is finished at this level. Accordingly, fluorescein-PEG-NHS has been used to study the mechanism of hapten–antibody interaction [40,41], and biotin-PEG-NHS (Figure 15.2A) [19,33] has been used to establish a test and start-up system for recognition force microscopy. In contrast to small ligands such as biotin or fluorescein, proteins or other sensitive ligands have to be bound in a multistep procedure. In this case heterobifunctional PEG linkers with one NHS ester function

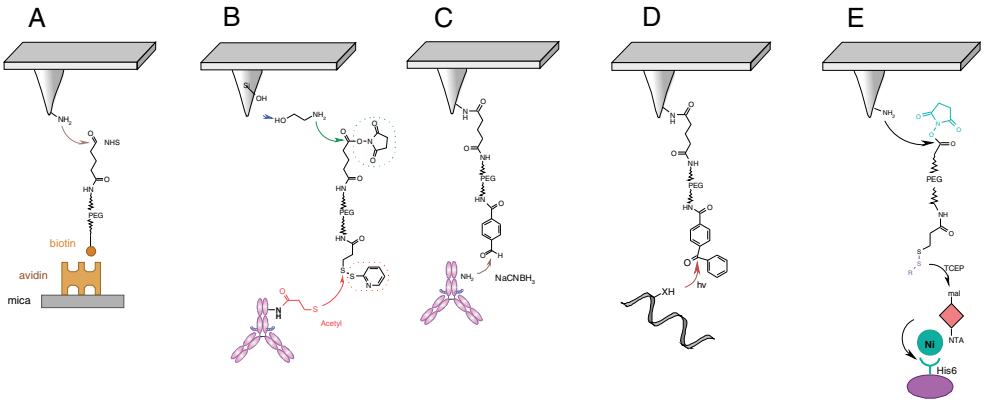


Figure 15.2. Common coupling schemes for tethering different ligands to aminofunctionalized tips. Functional end groups are: **A.** biotin, **B.** pyridyldithiopropionyl, **C.** aldehyde, **D.** benzophenon, **E.** nitrilotriacetic acid.

(for binding to the support) and another coupling function (for binding of the ligand) are used. In the following discussion the most prominent cross-linkers for aminofunctionalized tips are listed and their advantages or disadvantages are discussed.

Pyridyldithiopropionyl (PDP)-PEG-NHS (Figure 15.2B) is one of the most frequently used PEG linkers in force spectroscopy experiments [15,21,31,32,34,35,44,46–56]. After coupling this linker to the amine surface, the unreacted PDP function on the outer end can couple a thiolated protein via disulfide bond formation in phosphate buffer (pH 7.5). For protection of protein thiol from undesired oxidation, the inclusion of EDTA seems to be even more important than the use of an argon atmosphere. Proteins with endogenous thiols, such as half-antibodies [48] or cysteine peptides [52,53], can be bound directly, but unfortunately most ligands do not contain a free thiol group. This limitation can be overcome by prederivatization with *N*-succimidyl-3(acethylthio)-propionate (SATP) [44,57] as shown in Figure 15.2B. For the derivatization of sensitive antibodies, an advanced method is available that ensures conservation of the integrity of the binding epitope. The sugar residues of the antibody can be transformed into reactive thiols using periodate and pyridyldithiopropionic acid hydrazide (SPDP hydrazide) [58,59]. This allows the site-directed coupling of the PEG tether to the carbohydrate residue in the centre of the antibody. In contrast to the examples described earlier, PDP-PEG-NHS can also be used to create PEG-SH tentacles on the AFM tip, to which a thiol-reactive ligand molecule is coupled in the next step. This is exemplified by binding of iodoacetamodophlorizin, a potent blocker of the sodium glucose cotransporter, to an AFM tip [60]. Alternatively, small thiolated molecules can be coupled via vinylsulfone-PEG-NHS or maleimide-PEG-NHS. The vinylsulfone group is very stable in water, but it has a very low reactivity toward free thiols and needs high concentration of thiolated ligands. The maleimide function is somewhat sensitive to hydrolysis, but it excels with exceptional high reactivity toward thiol, and the resulting thioether bond can neither be cleaved by hydrolysis nor by reaction with other thiol molecules (in contrast to disulfide bridges [61]), and thus it is ideally suited for coupling of thiolated DNA to AFM tips.

Another approach for mounting ligands to AFM tips is the use of the noncovalent but rather strong hexahistidine (His_6)–Ni^{2+} ion–nitrilotriacetic acid (NTA) complex. The His_6 tag is very common in molecular biology, so that many proteins are available containing a His_6 tag. To bind His_6-tagged proteins, the aminofunctionalized tips must first been reacted

with SATP, generating a thiol-reactive surface. In the next step the heterobifunctional cross-linker PDP-PEG-NTA is bound via its PDP end, forming a covalent disulfide bridge as shown in Figure 15.2E. The NTA groups have been charged with Ni^{2+} ions for subsequent coupling of His_6-tagged, very low density lipoprotein receptor constructs [62], as well as of proteins involved in nuclear import [63].

The most widely applicable cross-linker for force spectroscopy experiments appears to be aldehyde-PEG-NHS (Figure 15.2C) [45]. Although this linker has two amino-reactive sites, the NHS ester is much more reactive and will bind preferentially to the amine groups on the tip surface when chloroform is used as solvent. The benzaldehyde residues on the free end remain active after coupling to the support on a time scale of hours and days [58,64,65] and can be further conjugated to the amino groups of the lysine residues of the protein that are to be coupled. Reaction of protein and aldehyde results in the formation of a Schiff base, which is subsequently fixed by reduction with $NaCNBH_3$ [45]. In spite of the high reactivity difference between NHS ester and aldehyde for amino groups, the tip-bound PEG-aldehyde tentacles have a significant probability of reacting with adjacent amino groups on the surface due to the close proximity between aldehyde and amine in this situation. Fortunately, the tendency for such loop formation on the surface can largely be suppressed by applying aldehyde-PEG-NHS at a concentration of 6.6 mg/mL rather than 1 mg/mL as with all other linkers [66].

In contrast to aldehyde-PEG-NHS, homobifunctional linkers like NHS-PEG-NHS show a significantly higher probability of loop formation. Nevertheless, NHS-PEG-NHS has successfully been used under suitable precautions, by adjusting a very low lateral density of amino groups on the tip surface [67]. A second problem is the high reactivity of the NHS groups, resulting in premature hydrolysis during the protein-coupling step [68,69]. In short linkers, the degree of hydrolysis is less critical because more-reactive sites are available in total compared to surfaces with long linkers like NHS-PEG_{3400}-NHS.

A very recently developed cross-linker is benzophenone-PEG-NHS (Figure 15.2D) [Wildling et al., manuscript in preparation] This linker does not need particular reactive functions like thiols or amino groups on the ligand, and it does not form loops on the surface. However, ultraviolet light is necessary for the coupling of protein or DNA to benzophenone. As a consequence, all coupling steps have to be done in dim light, so that it is only limited to experienced users. In contrast to other light-activated functions such as azides [35,70], benzophenone has the advantage that the photoactivated group can be excited to the reactive triplet state many times without loss of activity [71], thereby ensuring coupling efficiency.

A completely different strategy is the use of gold-coated tips (Figure 15.3). Such gold-coated tips are also common in force spectroscopy and can be obtained with low spring constants, allowing sensitive measurements. In the following paragraph the use of gold is mainly described for support functionalization because gold supports are much more widely used than gold tips. Gold supports have several advantageous features: (1) A perfectly clean and ultraflat gold surface can be generated by evaporation onto and subsequent stripping from mica [72–74]. (2) If the freshly stripped gold surface is immediately transferred into a thiol solution, a defined monolayer is formed and the surface stays protected from impurities. (3) Gold is chemically inert, allowing for harsh cleaning conditions such as ozone plasma, ozone, piranha, and NH_3/H_2O_2/water (caution: extremely hazardous [75]). (4) The most prominent advantage of gold surfaces is inertness to nearly any functional group except for thiols, disulfides, and thioethers [76]. The strong sulfur–gold interaction allows for convenient and stable formation of densely packed self-assembled monolayers (SAMs). The physical and chemical properties of such SAMs can easily be adjusted by varying the functional group on the outer surface of the SAM. In particular, protein repellence is observed when most alkane thiols

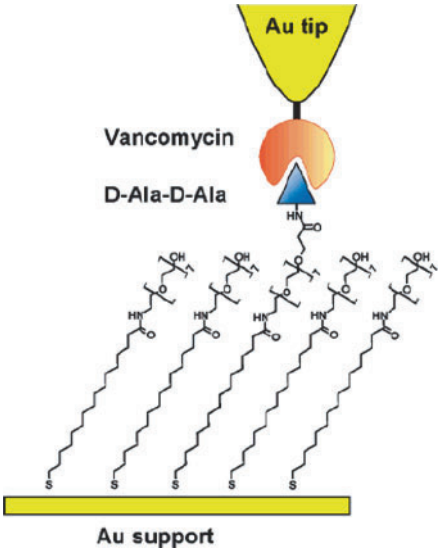


Figure 15.3. Single-molecule force spectroscopy and imaging of the vancomycin/D-Ala-D-Ala interaction using gold chemistry on the atomic force microscope tip as well as support. (From Gilbert et al. [172].)

carry an oligoethylene end group, and a wide range of coupling groups can be chosen at any desired molar fraction (usually between 2% and 20%), to which the biomolecules are then attached at the desired lateral density. Alternatively, the bare gold can directly be exposed to biomolecules that carry one or several free thiols/disulfides for binding to gold. Endogenous thiol or disulfide groups of different molecules can act as anchors for chemisorption. Fab fragments, antibodies [77], azurin [78], and fibronectin [79,80] were successfully bound via their endogenous thiol residues. Thiol-free ligands such as aggrecan or lectin were also bound by the same method after prederivatization with mercaptopropionyl groups [81] or an extended version of the same [82]. Chemisorption of thiolated ligands gives a dense layer of these molecules, which often inhibits recognition of a complementary molecule for steric reasons, as seen by lack of hybridization on a dense monolayer of thiolated single-strand DNA [83]. This problem has been solved by including other inert thiols that insert between the DNA molecules [83–85] or β-cyclodextrin [86–88] and reduce their lateral density in the mixed SAM to the optimal level. The more general approach in forming ligand-functionalized SAMs uses a modular multistep procedure. The first step is again the functionalization of gold with a mixed SAM consisting of inert components and of anchoring components with a suitable coupling function. The second step may be attachment of the ligand [89,90] or of a cross-linker to which the ligand is coupled in a third step [56].

In conclusion, common tip and support chemistry for molecular recognition force spectroscopy can either be based on silicon (or silicon nitride) tips or on gold tips, except for physisorptive methods, which can work in both cases but are only feasible in some instances. Silicon or silicon nitride tips are covered by a rather inert oxide layer, which can be activated by aminofunctionalization reagents such as APTES or ethanolamine. The resulting amino groups can be addressed either by ligand-containing linkers or by heterobifunctional cross-linkers. In the latter case the reactive ends can be reactive against a certain number of functional groups. The optimal choice depends on the chemistry, stability, and amount of ligand.

Although there is no "perfect" linker, coupling of nearly any ligand is possible by choosing a proper scheme. In contrast, gold-coated tips, if not used as physisorptive support, are functionalized by the use of SAMs. The addressing of such SAMs can be done either by single or by multiple step functionalizations. More extensive descriptions are given in a recent review by Ebner et al. [91] and in the cited literature.

15.2.2. The Force–Distance Cycle

Forces between biological proteins are measured by performing so called force–distance cycles using a ligand coupled to a tip that is mounted on a cantilever and a target surface with attached receptor molecules. Tethering the ligands to the tip using extended polymers such as PEG has several advantages over direct coupling: (1) the molecule can freely rotate, (2) it can diffuse within a volume determined by the tether length, and (3) the characteristic stretching behaviour allows discriminating between receptor–ligand interactions and nonspecific adhesion. At a fixed lateral position the cantilever is moved towards the surface and subsequently retracted. During this cycle the cantilever deflection is continuously monitored. A typical force–distance cycle is shown in Figure 15.4. At the beginning, the cantilever is far away from the surface and there is no interaction force. Consequently, the monitored cantilever deflection is zero. As the tip–sample distance is reduced, forces between the tip and the surface eventually develop. In case of attractive forces (van der Waals, electrostatic) the flexible cantilever is bent towards the surface, or, if repulsive forces are acting (electrostatic), the cantilever is pushed away. At each distance, the cantilever bends until its elastic

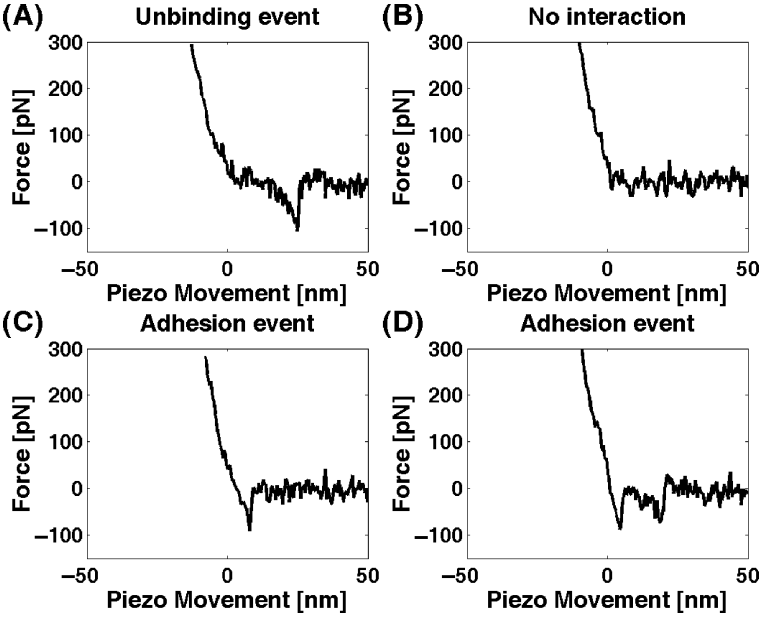


Figure 15.4. Results governed by performing force distance cycles with a ligand tethered to the atomic force microscope tip over a surface where a receptor is immobilized. **A.** A typical unbinding event. The nonlinear stretching of the poly(ethylene glycol) tether can be observed. **B.** No interaction of the tip and the sample. **C.** The tip adhered to the surface. In contrast to an unbinding event, no stretching of the tether is observed. **D.** Both unbinding and adhesion were observed. Due to the stretching of the tether, a separation in space took place.

force equals the tip–sample interaction force. The attractive forces can cause the tip to snap to the surface from a greater distance. If this jump to contact is caused by electrostatic attraction, it can be minimized by operating in electrolytes that screen these interactions [92]. Once in contact with the surface, the cantilever will experience an ever-increasing repulsive force, causing the cantilever to bend away from the surface as the electron orbitals of the atoms in the tip and sample overlap.

After a preset value of force is reached, the direction of motion is reversed and the cantilever is retracted from the surface. First, cantilever bending decreases until the repulsive force drops to zero. During contact of the tip with the surface, physical connections may manifest, such as adhesion forces or interaction forces between the receptor–ligand pair. In this case the cantilever is bent downwards on further retraction, and an increasing force is exerted on the connection until it breaks at a certain critical force, termed the rupture force. If no connection manifests, the cantilever stays in its equilibrium position during further retraction and the retrace looks the same as the approach. If the origin of the force is due to the receptor–ligand interaction, the observed force profile shows a nonlinear monotonic behaviour. This profile is determined by the elastic properties of the flexible PEG chain [53]. In case of nonspecific adhesion of the tip to the sample, the observed force profile is linear. It is mandatory that the specificity of the receptor–ligand interaction be demonstrated by blocking experiments. Either the receptors on the surface are inactivated by injecting free ligands, which in turn block the receptor-binding sites, or free receptors are injected to block the ligand tethered to the tip. As a consequence, almost all specific recognition signals completely disappear and adhesion events are only occasionally observed.

Most AFMs detect a sensor voltage that is proportional to the cantilever deflection. Therefore this voltage has to be converted into force values. The deflection of the cantilever is usually measured with the optical lever method. In this approach the cantilever deflection is obtained by measuring the position of the laser spot reflected from the cantilever on a split photodiode. First, for determining the proportionality factor between cantilever deflection and sensor voltage, force curves on a stiff surface must be taken. The deflection of the cantilever can be assumed to be equal to the travel distance of the z-piezo in the region of tip–surface contact. The slope of a linear fit of this region gives the so-called optical lever sensitivity. Dividing the output of the photodiode through the optical lever sensitivity yields the displacement z of the cantilever. Because the cantilever acts as a small spring, the deflection z can be accordingly converted into the force F acting on the cantilevers by using Hooke's law, $F = kz$, where k is the cantilever spring constant.

15.2.3. Spring Constant Determination

Several methods for determining the spring constant of a cantilever are available: using a calibrated reference [93], the added mass method [94], the thermal noise method [95], and the Sader method [96]. Each of these methods has advantages and drawbacks; a good overview can be found in Ref. 97. The thermal noise method and the Sader method are frequently used because they have good accuracy and are easy to use. The thermal noise is based on modelling the cantilever as a spring, making use of the equipartition theorem. The thermal motion is related to the thermal energy through

$$k = \frac{k_B T}{\langle z^2 \rangle} \tag{15.1}$$

where k_B is the Boltzmann constant (1.38×10^{-23} J/K), T is the temperature, and $\langle z^2 \rangle$ is the mean square displacement of the cantilever. The last-named quantity is determined by performing a power spectral density analysis of the cantilever oscillations, fitting a simple harmonic oscillator model

$$A = A_{\text{white}} + \frac{A_0 \omega_0^4}{(\omega^2 - \omega_0^2)^2 + (\omega\omega_0/Q)^2} \tag{15.2}$$

where A_{white} is a white noise floor, A_0 is the zero frequency amplitude, ω_0 is the radial resonance frequency, and Q is the quality factor), and integrating without the white noise contribution. Analysis in the frequency domain has the advantage that either external noise sources are usually broadband noise that can be baseline subtracted or they will occur at discrete frequencies that differ from the cantilever resonance and can thus be ignored. Butt and Jaschke [98] showed that two corrections are necessary. The first correction takes into account that cantilevers behave as multimode oscillators instead of a single-mode harmonic oscillator, and therefore energy is distributed over more than one mode. The second and more significant correction is that the optical lever detection scheme is proportional to angular changes in the cantilever position and not to the absolute deflection. These angular changes depend on the bending mode of the cantilever. By using beam theory, an analytic solution accounting for both effects was found for rectangular cantilevers, which results for the fundamental mode as follows:

$$k = 0.817 \frac{k_B T}{\langle (z_1^*)^2 \rangle} \tag{15.3}$$

where z_1^* is the deflection as measured by the optical lever method. Stark et al. –used finite-element analysis to examine one particular V-shaped cantilever, the Veeco Microlever E, which has a nominal spring constant of 0.1 N/m. The result of the numerical calculation was that the prefactor 0.817 has to be changed to 0.764.

The other technique to be discussed is the so-called Sader method. This technique requires the plane view dimensions, the resonance frequency, the quality factor, and the density and viscosity of the surrounding fluid (typically air) [96]. Although mathematically complex, it is experimentally very convenient. It models the contribution of the dissipation due to the fluid to the dynamic characteristics of the cantilever motion and yields

$$k = \rho L_0^3 \Omega(\text{Re}) \omega_0^2 Q \tag{15.4}$$

where ρ is the density of the surrounding fluid, L_0 is a length scale of the oscillator, ω_0 is the radial resonance frequency, Q is the quality factor in the fluid, and Ω is a dimensionless function, depending implicitly on the geometry and deflection function of the cantilever. Once determined for a body of given size, the function then holds universally for bodies of the same geometry but different size. Ω turns out to be a function of the Reynolds number Re*:*

$$\text{Re} = \frac{\rho \omega_0 L_0^2}{\eta} \tag{15.5}$$

where η is the viscosity of the fluid.

The resonant frequency ω_0 and the quality factor Q are determined by fitting the frequency response of the cantilever to Eq. (15.2). In principle, $\Omega(\mathrm{Re})$ can be calculated theoretically by solving the coupled fluid–structure interaction problem. However, this poses a formidable challenge. At least for rectangular cantilevers an expression was found when the length scale L_0 was chosen to be $b/2$, where b is the width of the cantilever:

$$\Omega(\mathrm{Re}) = 1.525\frac{L}{b}\Gamma_i(\omega_0) \tag{15.6}$$

and therefore

$$k = 0.1906\rho L b^2 Q\omega_0^2\Gamma_\mathrm{i} \tag{15.7}$$

where L is the length of the cantilever and Γ_i is the imaginary part of the so-called hydrodynamic function [99]. This function is quite complex; however, it is available on a web site [100] as a Java applet that performs the calculation and as downloadable Mathematica code.

For other geometries $\Omega(\mathrm{Re})$ can be evaluated as a function of Re by immersing a body with known spring constant in gas, such as air, adjusting the gas pressure, and making use of Eq. (15.4). This makes it possible to determine $\Omega(\mathrm{Re})$ over many orders of magnitude in Re because the gas density is proportional to pressure and the viscosity is independent of pressure. However, this procedure is only valid as long as the mean free path does not become comparable to or is larger than the length scale of the cantilever and the underlying continuum assumption in the analysis is valid. For Veeco Microlever cantilevers C and D an empirical formula for the spring constant was found using the aforementioned method [96] and taking L_0 to be equal to the width of the cantilever arms d:

$$k_\mathrm{C} = 3.57\rho d^2 L\mathrm{Re}^{-0.728+0.00915\ \ln\ \mathrm{Re}}\omega_0^2 Q k_\mathrm{D} = 2.97\rho d^2 L\mathrm{Re}^{-0.700+0.0215\ \ln\ \mathrm{Re}}\omega_0^2 Q \tag{15.8}$$

where

$$\mathrm{Re} = \frac{\rho d^2\omega_0}{\eta} \tag{15.9}$$

15.2.4. Force Distributions

Multiplying the deflection of the cantilever against the measured spring constant converts the data into forces. Force–distance cycles are typically repeated more than 1,000 times, usually resulting in more than 100 specific receptor–ligand rupture forces, finally yielding a distribution of rupture forces. Usually the probability density is estimated by histograms. For this the force axis is distributed into classes. The probability density estimate for a certain class $[a, b]$ is defined as the number of measured rupture forces within this class divided by the total number of measurements n and the width of the class $(b - a)$. Even though this procedure is simple and intuitive, it has several disadvantages: (1) the choice of classes is subjective, even though there exist some hints for choosing the optimal class width, and (2) the histogram is only piecewise continuous, with jump discontinuities at the class boundaries, and is therefore a poor estimate for typically continuously differentiable probability densities. These limitations can be overcome by using kernel estimation methods [101]. The probability density $p(x)$ is estimated by

$$p(x) = \frac{1}{nb} \sum_{i=1}^{n} K\left(\frac{x - x_i}{b}\right) \tag{15.10}$$

where x_i are measured data points and b is the bandwidth of the kernel K. The kernel K can be any function that is positive, symmetric, and normalized according to $\int_{-\infty}^{\infty} K(u)d(u) = 1$. Baumgartner et al. [102] suggested using the normal distribution as kernel and the measurement error of each force rupture event as bandwidth to estimate the probability density of rupture forces *pdf*:

$$pdf(F) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi s_i^2}} \exp\left(\frac{(F - F_i)^2}{2s_i^2}\right) \tag{15.11}$$

where F_i are the measured rupture force data and s_i the corresponding standard deviation. The advantage is that the values are weighted by their reliability and thus yield a better resolution. Furthermore, the problem of binning the data is avoided (Figure 15.5).

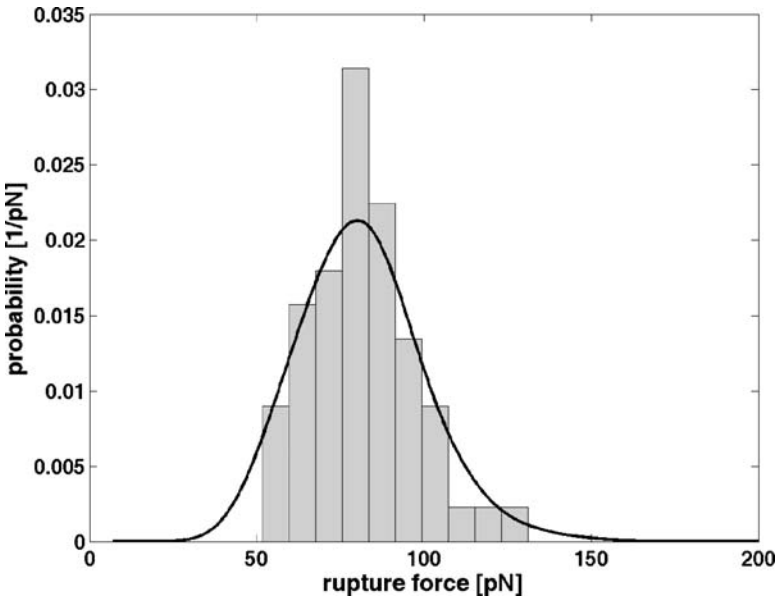


Figure 15.5. A histogram showing the distribution of measured rupture forces. It is overlaid with the experimentally determined probability density function, using Eq. (15.11). Both curves were normalized so that the area under the respective curve is 1.

15.2.5. Theory of Force Spectroscopy

In force spectroscopy experiments the applied force is weak, such that rupture times are longer than diffusional relaxation times. This implies that the ligand–receptor complex is always in a quasi-equilibrium state. Therefore, the survival probability $S(t)$ for the system

(i.e., the probability that rupture has not yet occurred at time t) under the influence of an external force $F(t)$ obeys a first-order kinetics with a time-dependent rate,

$$\frac{dS(t)}{dt} = -k(t)S(t) \tag{15.12}$$

and thus

$$S(t) = \exp\left[-\int_0^t k(t')dt'\right] \tag{15.13}$$

The probability distribution of forces F at rupture is related to the probability distribution of lifetimes by $p(F)dF = -\dot{S}(t)dt$, yielding

$$p(F) = \frac{k(F)}{\frac{dF}{dt}} \exp\left(-\int_0^F \left[\frac{k(F')}{\frac{dF}{dt}}\right] dF'\right) \tag{15.14}$$

During an irreversible molecular transition under the influence of a force, it is assumed that the molecule moves along a combined free energy surface $U(x) = U_0(x) - Fx$ in the pulling direction, x. The bare free energy $U_0(x)$ is assumed to have a single well at $x = 0$ and a barrier of height ΔG_β at $x = x_\beta$ (Figure 15.6). Due to the small size of the system, the surrounding heat bath causes notable energy fluctuation, resulting in a stochastic escape process. The applied force lowers the energy barrier and eases the dissociation caused by thermal energy fluctuations, resulting in a distribution $p(F)$ of measured rupture forces. Attempts to describe force spectroscopy experiments [11,12,103] involve Bell's expression to describe the increased rate of bond dissociation under external force [104]:

$$k(F) = k_{\text{off}} \exp\left(Fx_\beta / k_B T\right) \tag{15.15}$$

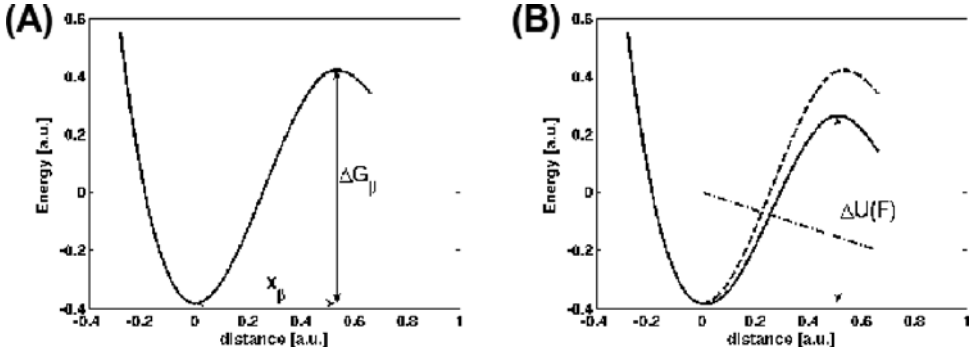


Figure 15.6. Schematic representation of a single barrier potential with and without applied force. **A.** Intrinsic free-energy surface $U_0(x)$ with minimum-to-barrier distance x_β and activation free energy ΔG_β. **B.** Combined free energy surface under the influence of an external force ramp, showing a new barrier height of $\Delta U(F)$.

where k_{off} is the kinetic off rate constant. The attractive feature of this phenomenological description is the apparent generality. No assumption is made concerning the nature of dynamics (it is subsumed into k_{off}), and the underlying free energy surface is characterized by a single parameter (x_β).

In typical force spectroscopy experiments the cantilever is retracted at a constant pulling speed, resulting in a linear force ramp, $F(t) = k_c vt$, where k_c is the spring constant of the cantilever and v is the retraction speed. If a linker is used, k_c has to be replaced by the resulting spring constant of the serial combination of the cantilever and tether, $(k_c^{-1} + k_{tether}^{-1})^{-1}$, where k_{tether} is the spring constant of the linker molecule. Usually this spring constant depends on the applied force, making it impossible to find an analytical solution of Eq. (15.14). A simple but powerful approximation is to use the resulting combined spring constant at rupture k_{eff} [105,106], again yielding a linear force ramp:

$$F(t) = k_{eff} vt \tag{15.16}$$

Substituting Eqs. (15.15) and (15.16) into Eq. (15.14) yields an analytical expression for the observed rupture force distribution:

$$p(F) = \frac{k_{eff}}{r} \exp\left\{\frac{Fx_\beta}{k_B T} - \frac{k_{off} k_B T}{r x_\beta}\left[\left(\exp\frac{Fx_\beta}{k_B T}\right) - 1\right]\right\} \tag{15.17}$$

Here k_{off} is the dissociation rate without an applied force, and the loading rate r is defined as $k_{eff} v$. A logarithmic dependence of the most probable rupture force F^* with respect to the loading rate r is found according to

$$F^* = \frac{k_B T}{x_\beta} \ln\left(\frac{x_\beta r}{k_B T k_{off}}\right) \tag{15.18}$$

Recently, calculations using model free energy surfaces yielded different expressions for $k(F)$. Hummer and Szabo [107] used a harmonic potential with a cusp-like feature at x_β: $U_0(F) = k_B T \Delta G_\beta (x/x_\beta)^2$ for $x < x_\beta$ and otherwise $-\infty$. Dudko and coworkers [108] used a linear-cubic surface: $U_0(F) = (3/2)\Delta G_\beta x/x_\beta - 2\Delta G_\beta (x/x_\beta)^3$. Applying Kramers' theory of escape from a potential well yields the respective expressions for $k(F)$. Substituting these expressions into Eq. (15.14) resulted in the distribution of rupture forces. The latter two models and the model described earlier can be unified within a single theoretical framework [109]:

$$k(F) = k_{off}\left(1 - \frac{\mu F x_\beta}{\Delta G_\beta}\right)^{1/\mu - 1} \exp\left\{\frac{\Delta G_\beta}{k_B T}\left[1 - \left(1 - \frac{\mu F x_\beta}{\Delta G_\beta}\right)^{1/\mu}\right]\right\} \tag{15.19}$$

$$p(F) = \frac{k(F)}{r} \exp\left(\frac{k_{off} k_B T}{r x_\beta}\right) \exp\left\{-\frac{k_B T k(F)}{r x_\beta}\left[1 - \left(1 - \frac{\mu F x_\beta}{\Delta G_\beta}\right)^{-1/\mu}\right]\right\} \tag{15.20}$$

where $\mu = 2/3$ and $1/2$ correspond to the linear-cubic and quadratic free-energy surfaces, respectively. For $\mu = 1$ and for $\Delta G_\beta \to \infty$ independent of μ, the expression reduces to the result of Evans and Ritchie [103]. When $\mu \neq 1$, permissible values of force F are limited from above by the value of the critical force $F_c = \Delta G_\beta/(\mu x_\beta)$ at which the barrier disappears, leading to wrong results for $k(F)$ because of the invalidity of Kramers' theory within this limit.

The most probable rupture force F^* and the variance of the force distributions σ_F were estimated to be, respectively,

$$F^* \approx \frac{\Delta G_\beta}{\mu}\left[1-\left(1-\frac{k_\mathrm{B}T}{\Delta G_\beta}\ln\frac{k_\mathrm{B}Tk_\mathrm{off}\exp\left(\Delta G_\beta/k_\mathrm{B}T\right)}{x_\beta r}\right)^{\mu}\right] \tag{15.21}$$

$$\sigma_F^2 \approx \frac{(k_\mathrm{B}T\pi)^2}{6x_\beta^2}\left(\frac{k_\mathrm{B}T}{\Delta G_\beta}\ln\frac{k_\mathrm{B}Tk_\mathrm{off}\exp\left(\Delta G_\beta/k_\mathrm{B}T+\tilde{\gamma}\right)}{x_\beta r}\right)^{2\mu-2} \tag{15.22}$$

Here $\tilde{\gamma}=\gamma^2-3/\pi^2\psi''(1)\approx 1.064$, where $\gamma=0.577$ is the Euler–Mascheroni constant and $\psi"(1)=-2.404$ is a particular value of the tetragamma function [110]. Within this model the most probable rupture force is proportional to $(\ln r)^{\mu}$.

A completely different model describing receptor–ligand unbinding was suggested by Katletz and Titulaer [111] and refined by Katletz in his thesis [112]. This work is based on the idea that dissociation of a molecule can be described by means of a thermally driven random walk with an absorbing barrier. It was assumed that:

- The dissociation can be described by a single coordinate, and the energy profile is harmonic, resulting in equally spaced energy levels $E_n = h\nu(n+1/2)$, $n = 0,\ldots, N-1$. The transition from level $N-1$ to N leads to dissociation.
- Only transitions from one level to a neighbouring one can occur. The transition probabilities per interaction is given by $P_{mn}=P_{nm}=\left[(m+1)\delta_{n-1,m}+m\delta_{n+1,m}\right]P_{10}$.

The transition probabilities per unit time were given by

$$\tilde{W}_{n,n+1} = ZcP_{n,n+1} \tag{15.23}$$

$$\tilde{W}_{n+1,n} = Zce^{-\beta\Delta E}P_{n+1,n} \tag{15.24}$$

where Z is the number of collisions between the heat bath molecules and the reacting molecule (with unit density), and c is the concentration of heat bath molecules. This factor can be eliminated by rescaling the time by $(ZcP_{10})^{-1}$. Here ΔE is the energy difference between two adjacent energy levels.

Protein structure and binding rely on noncovalent weak molecular interactions whose strength is on the order of the thermal energy. Consequently, the dissociation is described as a thermally driven random walk through states with different numbers of bonds. The Subscript of the transfer matrix is changed from energy level n used before to the number of bonds $i = N-n$. In the ground state E_0 all bonds are closed ($i=N$), whereas in the state E_N all bonds are open, $i=0$, and dissociation occurs. Therefore the state E_N is related to the activation energy E_a. Within this context the probability that a bond breaks is proportional to the number of intact bonds i, whereas the probability that an open bond closes again is proportional to the number of open bonds, $N-i$:

$$W_{i,i+1} = N-(i+1) \tag{15.25}$$

$$W_{i+1,i} = (i+1)\mathrm{e}^{-\beta\Delta E} \tag{15.26}$$

In force spectroscopy experiments the complex of the interacting biomolecules is stressed by an external force. Using the Bell ansatz [Eq. (15.15)] leads to a reduction of the activation

energy by $\beta F\lambda/i$, where λ determines the interaction range and F is the external force acting on the complex, which is equally distributed over i intact bonds:

$$W_{i,i+1} = N - (i\ + 1) \tag{15.27}$$

$$W_{i+1,i} = (i+1)e^{-\beta\left(\Delta E - \frac{F\lambda}{i+1}\right)} \tag{15.28}$$

In biology, individual bonds often act with respect to the state of other bonds ("cooperativity"). Thereby the behaviour is altered by the state of neighbouring bonds. Positive cooperativity means that the bond is more stable if bonds next to it are established, in contrast to negative cooperativity, where stability is reduced.

Cooperativity is introduced into the model by adapting the Ising model, which is used to describe the collective properties of magnetic spins. Here, it is not sufficient to deal just with the number of open bonds; instead, every possible configuration has to be taken into account. A certain configuration of m weak bonds is represented as a sequence of 0's and 1's with length m, where 0 means that the weak bond is open and 1 that the weak bond is established. This arrangement is identified with the number j whose binary representation is given by this sequence. All possible configurations are given by the numbers 0,. . ., $2^m - 1$. Here $2^m - 1$ is the ground state where all bonds are formed, and 0 is the highest-energy state and represents the unbound complex. Again transitions are only allowed between adjacent states that differ at a single position in the binary representation.

The elements of the cooperative transfer matrix $\boldsymbol{W}_{\mathbf{C}}$ correspond to the three elemental types of reactions:

$$\begin{aligned} \ldots 000 \ldots &\underset{k}{\overset{1}{\rightleftharpoons}} \ldots 010 \ldots \\ \ldots 100(001) \ldots &\underset{k/c}{\overset{1}{\rightleftharpoons}} \ldots 110(011) \ldots \\ \ldots 101 &\underset{k/c^2}{\overset{1}{\rightleftharpoons}} \ldots 111 \ldots \end{aligned}$$

It is assumed that the formation of a single bond has the same rate constant, set to 1 (because of time rescaling), whereas the rupture of a single bond depends on the state of its neighbouring sites. Every neighbouring bond contributes a factor $1/c$ to the rate constant. An external force is included by modifying the rate constant according to the Bell ansatz and taking into account that the force is distributed over the intact bonds.

For both models with and without cooperativity the transfer matrix $\boldsymbol{W}$ describes the population of states according to a system of linear rate equations:

$$\frac{dx_n}{dt} = \sum_m W_{nm}x_m - W_{mn}x_n = \sum_m A_{nm}x_m \tag{15.29}$$

with

$$A_{mn} = W_{nm} - \delta_{nm}\sum_l W_{\ln} \tag{15.30}$$

The transition state $n = 0$ acts as an absorbing barrier, meaning that molecules reaching this state are removed from the ensemble. In terms of mathematics this means

$$x_0 = 0 \tag{15.31}$$

$$A_{i0} = 0 \tag{15.32}$$

Equation (15.29) can be solved using different methods. First it can be numerically integrated. Because this system of differential equations is stiff, care has to be taken in choosing the right algorithm. Usually backward differentiation is a good choice, and a starting value for a Boltzmann distribution should be used. With the solution of $\vec{x}(t)$it is easy to compute the ratio of bound molecules as a function of time and then to determine the mean first passage time τ. A second method is to make use of the fact that the system is linear. The solution can be obtained by calculating the spectrum of $\boldsymbol{A}$. The magnitudes of the eigenvalues vary greatly, and a single decay is observed, as expected from the separation of time scales. Therefore the most interesting eigenvalue is the smallest one. The mean first passage time is simply the inverse of this eigenvalue τ. Because we are only interested in the smallest eigenvalue, a little "trick" developed by Farkas was applied [113]: We are interested in the rate by which the absorbing state gets populated. If all transitions into this state are fed back into the ground state, a flow through the system emerges. The inverse of this flow is the mean first passage time τ. Technically this flow is determined by calculating a new transition matrix $\boldsymbol{A}$ by adding the first line A_{i0} containing the transitions into the absorbing barrier to the last line of $\boldsymbol{A}$. The probability x_0 is not in equilibrium and therefore not a variable of the system. Accordingly, the first column and row are discarded, yielding the new submatrix $\boldsymbol{A}^*$. This new matrix is singular, corresponding to a stationary state:

$$\frac{d\vec{x}}{dt} = A^* \cdot \vec{x}^{stat} = 0 \tag{15.33}$$

In addition, the probabilities x_i must be normalized:

$$\sum_i x_i^{stat} = 1 \tag{15.34}$$

Solving this overdetermined linear equation system yields $\vec{x}^{stat}$, which allows calculation of the stationary rate given by

$$f = \sum_i A_{oi} x_i^{stat} \tag{15.35}$$

Combining both models allows a description of the interactions where N groups consisting of m weak bonds interact cooperatively. The total number of weak bonds is accordingly $N{\cdot}m$. The resulting transfer matrix is

$$\begin{aligned} W_{i,i+1} &= N - (i+1) \\ W_{i+1,i} &= (i+1)\left(\tau_m\left(k,c,\tfrac{F}{i+1}\right)\right)^{-1} \end{aligned} \tag{15.36}$$

where $\tau_m\left(k,c,\frac{F}{i+1}\right)$ is the lifetime of m weak bonds acting cooperatively with the parameters k and c and subject to an external force $F/(i + 1)$. The solution of the linear rate equation system (15.29) with transition matrix W defined by Eq. (15.36) yields the lifetime τ of the complex as a function of the external force F. The dissociation rate is given by

$$k(F) = k_{\text{off}}\frac{\tau(0)}{\tau(F)} \tag{15.37}$$

Inserting this equation into Eq. (15.14) allows us to calculate the distribution of rupture forces numerically.

15.2.6. The Effect of Hidden Barriers on Kinetic Parameters

The energy landscape of molecular bonds incorporates the interactions of many molecular groups, such as multiple hydrogen bonds and van der Waals interactions. Examples are the high-affinity biotin–avidin complex [103], the biotin–streptavidin complex [114], and other interactions described in Section 15.3. In many cases, the dissociation of the receptor–ligand bond proceeds via one ore more intermediate states. The simplest picture is the two-barrier model [103], in which the bound and the unbound states are separated along the pulling coordinate by two barriers, with an intermediate state in between. Figure 15.7 shows a conceptual, unperturbed energy landscape $E(x)$ along the pulling coordinate x for such a complex. In this case, the transition state is located at the outer (right) barrier. Applying a small force F_1 to the complex leads to a linear decrease of the potential by $F_1 \cdot x$, and the new combined potential is $E(x) - F_1 \cdot x$ with the transition state still located at the outer barrier. Under a higher force $F_2 > F_1$, the main barrier switches from the initial transition state to the inner transition state (inner barrier, left; combined potential $E(x) - F_2 \cdot x$). When the force is increased further, $F_3 >> F_1$, the energy barrier is lowered to the level of the initial minimum [combined potential $E(x) - F_3 \cdot x$], which finally would allow unhindered dissociation of the complex. The crossover in transition states leads to two distinct regimes with different slopes in the loading rate dependence of the most probable unbinding force and hence to two

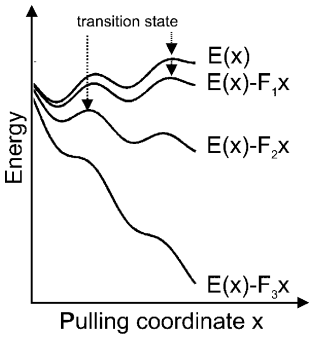


Figure 15.7. Conceptual free energy surface of the two-barrier model. For the unperturbed system, $E(x)$, and for small external forces, $E(x) - F_1x$, the transition state is located at the outer barrier. For higher forces, $E(x) - F_2x$, the transition state shifts inward from the outer barrier to the inner barrier. Upon further increase of the force, $E(x) - F_3x$, the energy barrier is lowered to the level of the initial energy minimum, ending in the unhindered dissociation of the bond.

different off-rates and potential widths in the Bell model. However, the assignment of different slopes, and therefore off-rates and potential to the position of the barrier along the pulling coordinate, is not unique [43]:

We assume a two-barrier model with the transition rates [from the bound to the intermediate state $k_{-1}(F)$, the backward rate $k_{+1}(F)$, and the rate from the intermediate to the unbound state $k_2(F)$] depending exponentially on the force, that is, $k_i(F) = k_i(0)^{eFxi/kT}$, with the corresponding potential widths x_{-1}, x_{+1} (negative for a transition opposite to the direction of the applied force), and x_{+2}, respectively. Following the formalism given in Ref. 43 for the generalized master equation, we obtain the mean dissociation time for a fixed force as

$$\tau(F) = \frac{k_{-1}(F) + k_{+1}(F) + k_{-2}(F)}{k_{-1}(F)k_{-2}(F)} \tag{15.38}$$

Given this formula, one can classify typical cases for the dependence of the effective dissociation rate $\tau^{-1}(F)$ on the force (Figure 15.8). Because the intermediate state has a higher energy than the ground state, $k_{+1}(0) >> k_{-1}(0)$ (assuming an energy difference of a few kT) holds. First we consider the cases in which the transition state from the intermediate to the unbound state has a higher energy than the transition to the ground state, so that $k_{+1}(0) >> k_{-2}(0)$ also holds (Figure 15.8A, B). For the limit of small forces in Eq. (15.38) we get the effective dissociation rate

$$\tau^{-1}(F) = \frac{k_{-1}(0)k_{-2}(0)}{k_{+1}(0)} e^{F(x_{-1}-x_{+1}+x_{-2})/kT} \tag{15.39}$$

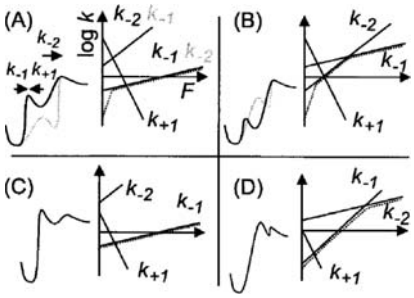


Figure 15.8. Different types of the dissociation via an intermediate state induced by an external force. **Left.** Conceptual energy landscapes. **Right.** Logarithm of the model transition rates $k_{21}(F)$, $k_{11}(F)$, and $k_{22}(F)$ as function of force. The effective dissociation rate is indicated by a dotted line. The gray potential scheme corresponds to an exchange of the functions $k_{21}(F)$ and $k_{22}(F)$, which does not change the effective dissociation rate. **A.** Two regimes in the dissociation process appear. At small forces all three transition rates determine the process. At high forces the transition from the ground to the intermediate state (from the intermediate to the unbound state, for the gray potential scheme) is rate determining. **B.** Similar situation as in panel A, but the rate-determining (forward) transitions cross over with increasing force. **C**, **D.** The intermediate state is only visited after the transition state with the highest energy has been passed. Because the dynamics of the complex is overdamped and fluctuation driven, the intermediate can still be dynamically relevant at an applied force **(D)**. (From Strunz et al. [43].)

From Eq. (15.39) we can see that the exponential increase of the thermal dissociation rate is governed by the distance of the bound state to the outermost barrier ($= x_{-1} - x_{+1} + x_{+2}$). When the force increases, the backward transition becomes negligible, $k_{-1}(0) >> k_{+1}(0)$, and the dissociation is dominated by either the transition rate to the intermediate state $k_{-1}(0)$ or

to the unbound state $k_{-2}(0)$. On further increase of the force it is also possible that a second transition occurs where the rate-dominating transition changes again (Figure 15.8B). For example, if initially the rate is dominated by the transition from the bound to the intermediate state, the increasing force can lead to a domination of the rate by the transition from the intermediate to the unbound state, leading to three distinct force intervals in the dissociation rate.

Equation (15.39) describes the dissociation rate in the first interval for small forces; in the second interval the dissociation rate is $\sim k_{-1}(F)$ [or k_{-2}(F)]; and in the third interval it is $\sim k_{-2}(F)$ [or $k_{-1}(F)$]. Only in this case all the parameters describing the two-state model (i.e., all three rates and potential widths) can be extracted directly by measuring the function $\tau(F)$. However, interchanging the parameters describing the functions $k_{-2}(F)$ and $k_{-1}(F)$ does not change $\tau(F)$, and therefore it is not possible to assign the measured parameters unambiguously to a transition. In addition, it is possible that even when the transition state from the bound to the intermediate state is the thermodynamically relevant transition state so that $k_{+1}(0) < k_{-2}(0)$ (Figure 15.8C, D), the intermediate state can be rate determining with an applied force (Figure 15.8D).

The model with one intermediate state is the simplest model to explain the experimentally observed behaviour of a receptor−ligand system like avidin–biotin [43]. The actual energy landscape of the bond may still be more complicated because only the rate-determining transitions are clearly detectable in the bond lifetime as a function of the force. Generally, the low-force regime is always associated with the thermodynamically relevant transition state, and the corresponding potential width is the distance to the bound state projected onto the direction of the applied force. The regimes at higher forces correspond to rate-determining transitions that can be located anywhere along the mechanical separation pathway.

15.2.7. Free Energy Surface Reconstruction from Nonequilibrium Single-Molecule Pulling Experiments

Single-molecule pulling experiments are typically carried out under nonequilibrium conditions because it is impossible to perform the pulling infinitely slowly. Is it still possible to extract equilibrium properties from such measurements that drive the system away from equilibrium? From the second law of thermodynamics, it is well known that the average mechanical work of pulling will be larger than the free energy difference. Only if the experiment is performed reversibly, that is, infinitely slowly, is the work is equal to the free energy difference. From this point of view it appears only possible to obtain free energy differences from such measurements by extrapolation to zero pulling velocity. However, this state of affairs changed in 1997 when Jarzynski [115,116] discovered a remarkable identity between thermodynamic free energy differences and the irreversible work required to switch a system between two states (in terms of single-molecule pulling experiments, the folded and unfolded states of a protein, or the bound and unbound states of a receptor and a ligand), thus extending the inequality of the second law of thermodynamics. This identity, although not directly applicable to AFM measurements, suggests that in principle one should be able to extract free energy surfaces from repeated pulling experiments. In 2001, Hummer and Szabo [117,118] adapted Jarzynski's identity for the analysis of single-molecule pulling experiments, which will be briefly reviewed in the following. Hummer and Szabo showed that Jarzynski's identity follows almost immediately from the Feynman–Kac theorem for path integrals.

Assume a system whose phase-space density evolves according to a Liouville-type equation

$$\frac{\partial f\,(x,t)}{\partial t} = L_t f\,(x,t) \tag{15.40}$$

Here, L_t is an explicitly time-dependent evolution operator that has the Boltzmann distribution as a stationary solution, $L_t e^{-\beta H(x,t)} = 0$, with $H(x, t)$ the time-dependent Hamiltonian of the system, x the phase-space coordinates, and $\beta^{-1} = k_B T$. The unnormalized Boltzmann distribution at time t reads

$$p\,(x,t) = \frac{e^{-\beta H(x,t)}}{\int e^{-\beta H(x',0)} dx'} \tag{15.41}$$

Because the distribution (15.41) is stationary ($L_t p = 0$) and $\partial p \big/ \partial t = -\beta\,(\partial H/\partial t)\,p$, it follows that $p(x, t)$ is a solution of the sink equation

$$\frac{\partial p}{\partial t} = L_t p - \beta \frac{\partial H}{\partial t} p \tag{15.42}$$

as can be easily verified by substitution. The solution of this sink equation (15.42) starting from an equilibrium distribution at $t = 0$ can also be expressed as a path integral using the Feynman–Kac theorem. By equating these two different solutions, one immediately gets

$$\frac{e^{-\beta H(x,t)}}{\int e^{-\beta H(x',0)} dx'} = \left\langle \delta(x - x_t) \exp\left[-\beta \int_0^t \frac{\partial H\left(x_{t'},t'\right)}{\partial t'} dt' \right] \right\rangle \tag{15.43}$$

where $\delta(x)$ is the Dirac δ-distribution, and the average $\langle ... \rangle$ is over an ensemble of trajectories starting from the equilibrium distribution at $t = 0$ and evolving according to Eq. (15.40). In this way, each trajectory is weighted with the Boltzmann factor of the external work w_t done on the system,

$$w_t = \int_0^t \frac{\partial H\left(x_{t'},t'\right)}{\partial t'} dt' \tag{15.44}$$

Integrating both sides of Eq. (15.43) with respect to x results in Jarzynski's identity [115,116], and one gets the following relation between the equilibrium free energy difference $\Delta G(t)$ and the Boltzmann-averaged work w_t between times t and 0:

$$e^{-\beta \Delta G(t)} \equiv \frac{\int e^{-\beta H(x,t)} dx}{\int e^{-\beta H(x,0)} dx} = \left\langle e^{-\beta w_t} \right\rangle \tag{15.45}$$

Equation (15.45) allows us to reconstruct the underlying free energy surface from repeated single-molecule AFM experiments, as sketched in Figure 15.9. On the assumption that the cantilever vibrates in a harmonic potential with spring constant k, the corresponding Hamiltonian reads $H(x, t) = H_0(x) + k(q - z(t))^2$, with $H_0(x)$ being the Hamiltonian of the resting, unperturbed system. For $t \geq 0$, $z(t) = z(0) + vt$ is the position of the anchor of the pulling spring (cantilever), and $q = q(x)$ is the point of attachment of the molecular system to the pulling spring projected on the pulling direction. By substituting this Hamiltonian into

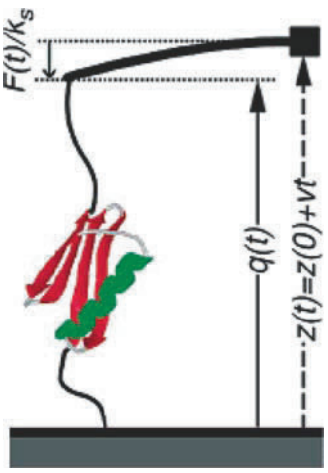


Figure 15.9. Schematic drawing of a single-molecule force spectroscopy experiment. The pulling apparatus (cantilever tip) is moved with a constant velocity v, where the controlled distance between the solid support and the anchor of the cantilever is $z(t) = z(0) + vt$. In contrast, the molecular extension indicated by $q(t)$, as well as the instantaneous force $F(t)$, underlies thermal fluctuations. (From Hummer and Szabo [117].)

Eq. (15.45), multiplying both sides by δ[q −q(x)], and integrating with respect to x and finally taking the logarithm, we obtain

$$G_0(q) \equiv -\beta^{-1} \ln \frac{\int \delta\left[q - q(x)\right] e^{-\beta H_0(x)} dx}{\int e^{-\beta H(x,0)} dx} = -\beta^{-1} \ln \left\langle \delta\left[q - q(x(t))\right] e^{-\beta \Delta w_t} \right\rangle \quad (15.46)$$

where $G_0(q)$ is the unperturbed free energy profile along the pulling coordinate q and Δw_t is the external work minus the instantaneous biasing potential, $\Delta w_t = kv(vt^2/2 - \int_0^t q_{t'} dt') - k(q_t - vt)^2/2$. It is required that the initial conditions of pulling trajectories $[x(t = 0)]$ are chosen from an equilibrium distribution corresponding to the Hamiltonian $H(x, 0) = H_0(x) + k(q - z(0))^2/2$. Therefore, the cantilever should be equilibrated at a fixed position $z(0)$ in all trajectories.

For an infinite number of pulling traces, the entire free energy surface could theoretically be reconstructed from observations at a single time t. In practice, at any time t the trajectories will be clustered near the location of the pulling spring, and therefore from observations collected at time t one can find reliable estimates of $G_0(q)$ only near $q \approx z(t)$. However, it is possible to combine multiple histograms obtained for different times t to improve the estimate of $G_0(q)$ (weighted histogram method [117]), which leads to the following expression for the free energy profile:

$$\exp\left[-\beta \Delta G_{WH}(q)\right] = \frac{\sum_t \frac{\left\langle \delta\left[q - q(t)\right] \exp(-\beta w_t) \right\rangle}{\langle \exp(-\beta w_t) \rangle}}{\sum_t \frac{\exp(-\beta V(q,t))}{\langle \exp(-\beta w_t) \rangle}} \quad (15.47)$$

where w_t is the accumulated work calculated by

$$w_t = \int_C F dq + V\left[q(t),t\right] - V\left[q(0),0\right] \tag{15.48}$$

Here, $F = k(z(t) - q(t))$ denotes the restoring force, where k is the cantilever spring constant and $V\left[q(t),t\right] = k\left(q_t - vt\right)^2/2$ is the harmonic biasing potential of the cantilever. The integral over q is along the position versus time contour connecting $q(0)$ and $q(t)$. If the cantilever is relatively stiff, then most trajectories will be clustered near $z(t)$, the position of the piezo actuator. In this case one can approximate the weighted distribution of molecular extensions by a Gaussian with mean

$$\bar{q}_t = \frac{\left\langle q(t)e^{-\beta w_t}\right\rangle}{\left\langle e^{-\beta w_t}\right\rangle} \tag{15.49}$$

and the corresponding variance $\sigma_t^2 = \overline{q_t^2} - \bar{q}_t^2$ [118]. Following this approximation (momentum-based approach), we can calculate the first derivative of the potential of mean force (i.e., the mean force) according to

$$G'_{MF}(\bar{q}_t) = \bar{F}_t = \frac{\left\langle F(t)e^{-\beta w_t}\right\rangle}{\left\langle e^{-\beta w_t}\right\rangle} \tag{15.50}$$

$G_{MF}(q)$ is then calculated by the cumulative integral of Eq. (15.50).

The equations were recently applied [119] to characterize the stability of the membrane proteins bacteriorhodopsin (BR), halorhodopsin (HR), and the Na^+/H^+ antiporter NhaA by calculating the unfolding free energies (i.e., the energy difference between the folded and unfolded states) of their transmembrane helices from single-molecule force measurements.

In contrast to its application to protein unfolding, Jarzynski's identity cannot be applied in a straightforward manner to unbinding because the elastic behaviour of the linker molecules that are used to tether the binding partners masks contributions of the latter. However the energy difference between the bound and unbound states of the receptor–ligand complex can be extracted by directly applying Jarzynski's identity, Eq. (15.45), to the deflection, $\mathrm{defl}(t) = z(t) - q(t)$, versus time curves obtained from repeated dissociation experiments to get the equilibrium binding energy of the corresponding complex.

15.3. Applications of Dynamic Force Spectroscopy to Protein Interactions

15.3.1. Load-Dependent Dynamics of Protein Interactions

Since their first application in 1992 [120] to measure the strength of individual hydrogen bonds, single-molecule force spectroscopy measurements have yielded valuable information on the mechanical properties of various biomolecules, the conformational transitions of polymers, and the strength of intermolecular bonds and the control of binding mechanisms. In this chapter, we limit ourselves to studies performed on the interactions of proteins with other proteins or with small ligands.

Ultimately, one wishes to use the data obtained from force spectroscopy to reconstruct the energy landscapes that underlie reactions of interest. Although a full reconstruction of energy landscapes of complex biological macromolecules is presently not feasible, the combination of sensitive measuring setups, protein engineering, and theoretical models and simulations are beginning to provide valuable glimpses into these landscapes.

Evans and his group set the guidelines for extracting barrier location from force spectroscopy measurements carried out over a broad range of loading rates [10]. Adhering to previous theoretical work, the authors used a biomembrane force probe (BFP) to probe bond rupture between biotin and its high-affinity receptors avidin and streptavidin with ramps of forces covering more than six orders of magnitude in rate [121]. As predicted by theory, and in contrast to early studies that reported fixed values of bond strength [122,123], a continuous spectrum of unbinding forces, ranging from 5 to 170 pN, was obtained as the loading rate increased from 0.05 to 60,000 pN/s. Concomitantly, interaction lifetime diminished from about 1 min to 0.001 sec, exposing the reciprocal relation between bond strength and lifetime anticipated for thermally activated processes subjected to a ramping force. Significantly, for both avidin and streptavidin, the force spectra revealed that unbinding kinetics is dominated by at least two prominent energy barriers located $\sim$0.1 and $\sim$0.5 nm, respectively, from the energy minimum along the direction of the force. The position of these barriers was found to match the location of transition states identified by force probe molecular dynamics simulations that reproduce the pulling of biotin from its binding pocket in avidin [11] or streptavidin [124] by means of external harmonic forces. The inner barrier was thus assumed to reflect the detachment of the biotin ureido ring from a network of hydrogen bonds and nonpolar interactions at the bottom of the binding pocket of (strept)avidin. The outer barrier was attributed to a subsequent disruption of interactions with the "3–4" loop of the two proteins, which closes behind biotin in the bound state. This latter assignment was validated directly by Yuan et al. [125], who showed that mutations of the W120 residue of this loop in streptavidin lead to a significant decrease in the height of the outer barrier but have no effect on the inner one.

Another interaction that was studied extensively by force spectroscopy is the one formed between selectins and their leukocyte-expressed ligands [126–129]. Selectins constitute one family of calcium-dependent lectin receptors present on the surface of endothelial cells. Together with their ligands, they mediate the emergence of leukocytes from the blood stream into sites of inflammation by modulating leukocyte rolling over the surface of the endothelium. This rolling motion occurs under hydrodynamic shear from the flow of the blood, rendering this an excellent system for investigation by a technique involving forces such as dynamic force spectroscopy (DFS). As discussed later, the ability to extract mechanical properties from the force extension profiles using theoretical frameworks such as the worm-like chain model provides important insights into the way in which mechanical properties relate to function.

Fritz et al. [127] used AFM to study the interaction between P-selectin and its cognate ligands—the leukocyte-expressed surface ligand P-selectin glycoprotein ligand-1 (PSGL-1). The complexes formed between the two proteins were found to sustain forces of up to $\sim$160 pN and exhibited chain-like elasticity characterized by a persistence length of 0.35 nm. The latter value is consistent with an extremely extended peptide conformation. This finding correlates well with the ability of selectins and their ligands to maintain cells tethered over very long distances, which necessitates adaptation of highly extensible structures. It was also found that the adhesion probability between PSGL-1 and P-selectin *increased* with increasing pulling velocities. This unusual dependence, which is typically reciprocal, is consistent with

the initial increase of tethering probability with shear stress observed in rolling. Only a single energy barrier was evident in the force spectra obtained in this study.

Using BFP, Evans et al. [128] studied of the interaction between PSGL-1 and another member of the selectin family, selectin L. As in the case of their previous study on strept(avidin)–biotin, the authors were able to probe bond strength over a very broad range of loading rates extending from 10 to 100,000 pN/s. In the presence of Ca^{2+}, the force spectra revealed two energy barriers along the unbinding path: an outer barrier located at $x_\beta \approx 4$ Å, which impeded dissociation under slow detachment, and an inner, Ca^{2+}-dependent barrier positioned at $x_\beta \approx 0.6$ Å, which dominated dissociation under rapid detachment. The differential response of the two barriers to the applied load was proposed to be important for the recruitment of multiple bonds during selectin-mediated function.

More recently, Marshall et al. [129] used force-clamp AFM to directly measure bond lifetimes of P-selectin–PSGL-1 complexes. Notably, the results indicated a biphasic relationship with bond lifetimes initially increasing with force ("catch") and, after reaching a maximum, decreasing ("slip"), revealing a force-induced switch from catch to slip bond behaviour. This switch, which was not observed in the initial DFS studies performed on this complex, was suggested to control leukocyte adhesion under varying shear stresses.

Another adhesion pair involved in leukocyte adhesion was studied by Wojcikiewicz, et al. [130], who used AFM to unbind complexes formed by the leukocyte function-associated antigen-1 (LFA-1) and its cognate ligands, intercellular adhesion molecules 1 and 2 (ICAM-1 and ICAM-2). The experimental setup consisted of LFA-1–expressing Jurkat T cells attached to the end of the AFM cantilever and surface immobilized ICAM-1 or ICAM-2. For both ICAM-1 and ICAM-2, the force spectra exhibited two linear regimes at fast and slow loading, amounting, respectively, to a sharp, inner energy barrier ($x_\beta \approx 0.56$ and 1.5 Å, for complexes formed with ICAM-1 and ICAM-2, respectively) and a shallow, outer barrier ($x_\beta \approx 3.6$ and 4.9 Å, respectively). Addition of Mg^{2+} led to an increase of the rupture forces measured in the slow-loading regime, indicating an increment of the outer barrier in the presence of the divalent cation. Comparison of the force spectra obtained for the complexes formed between LFA-1 and ICAM-1 or ICAM-2 indicated that, in the fast-loading regime, the rupture of LFA-1–ICAM-1 depends more steeply on the loading rate than does that of LFA-1–ICAM-2. The difference in dynamic strength between the two interactions was attributed to the presence of wider barriers in the LFA-1–ICAM-2 complex, which render the interaction more receptive to the applied load. The enhanced sensitivity of complexes with ICAM-2 to pulling forces was proposed to be important for the ability of ICAM-2 to carry out routine immune surveillance, which might be otherwise be impeded due to frequent adhesion events.

The last interaction we discuss in this section is the one between antibodies and their antigenic targets. Odorico et al .[131] used DFS to explore the energy landscape underlying the interaction between a chelated uranyl compound and a monoclonal antibody raised against the uranyl-dicarboxy-phenanthroline complex. To isolate contributions of the uranyl moiety to the binding interaction, measurements were performed with and without the ion in the chelating ligand. In the presence of uranyl, the force spectra contained two linear regimes, suggesting the presence of at least two major energy barriers along the unbinding pathway. To relate the experimental data to molecular events, the authors constructed a model with a variable fragment of the antibody and used computational graphics to dock the chelated uranyl ion into the binding pocket. The analysis suggested that the inner barrier ($x_\beta = 0.5$ Å) reflects the rupture of coordination bonds between the uranium atom and an Asp residue, whereas the outer barrier ($x_\beta = 3.9$ Å) amounts to the detachment of the entire ligand from the antibody-binding site.

In an earlier study, Schwesinger et al. [132] measured the unbinding forces between fluorescein and nine different monomeric Fv fragments derived from three unrelated antifluorescein antibodies. The on-rate, off-rate, equilibrium binding affinity, and activation energy for unbinding were determined in solution for all fragments, which included six point-mutated constructs that, together with the wild-type progenitors, provided a wide range of k_{off} and K_{D} values. In contrast to the study described earlier, the force spectra obtained for two of the constructs indicated a single energy barrier along the unbinding pathway. Extrapolation of the most probable unbinding forces to zero force gave k_{off} values very close to those measured in solution. Moreover, unbinding forces measured for all nine constructs at a fixed loading rate correlated well with the thermal off-rates of the constructs. This correlation held not only within homologs derived from a single antibody, but also across the three unrelated antibody species, suggesting a simple relationship between binding strengths and thermal off-rates that is rather insensitive to the details of the underlying energy surface.

A good correspondence between dissociation rates derived from mechanical unbinding experiments and from bulk assays was also reported by Neuert et al. [133]. In this case, the experimental system consisted of digoxigenin and its specific antibody. This pair is used as a noncovalent coupler in various applications, including forced unbinding experiments. The force spectra obtained for the complex suggested that the unbinding path is traversed by two activation energy barriers located at $x_\beta = 0.35$ and 1.15 nm, respectively. Linear fit of the low-force regime revealed a dissociation rate at zero force of 0.015 sec^{-1}, in close agreement with the 0.023 sec^{-1} obtained from bulk measurements made on antidigoxigenin Fv fragments.

The close agreement between dissociation rates derived from bulk measurements and from mechanical unbinding experiments observed in the two studies just described, is however, an exception rather than the rule. In most cases, the rates derived from extrapolation of the most probable unfolding forces to zero differs substantially from those measured in solution assays. Partly this is because outer barriers, which govern spontaneous dissociation kinetics, are not necessarily detected in force-probe experiments. This is particularly true if one uses AFM, which is quite limited in the low-force regimes where these barriers often reside. The second reason, which is more relevant to our discussion, is that the strict logarithmic dependence of the most probable force on the loading rate, which allows simple derivation of off-rates from the force spectra by linear extrapolation, is valid only if the barriers along the energy profile are sharp. However, if the curvature at the top of the barrier is small, the position of the transition state along the reaction coordinate is expected to *vary* with the force, which, in turn, changes with the loading rate. In such cases, derivation of spontaneous off-rates and, more relevant to this section, x_β from the force spectra can be associated with significant errors. Unfortunately, due to the limited range of loading rates available in AFM measurements and the usually limited number of data points used to construct the force spectra, distinction between a strict and nonstrict logarithmic behaviour is not a trivial task.

A way to overcome this problem is to fit the probability force distributions using Kramers' diffusion model. This approach was taken by Schlierf and Rief [134] to fit data that could not be reproduced by the Bell model (Figure 15.10). Notably, although the position of the transition state predicted by the Bell model was smaller than that predicted by the Kramers analysis by 6 Å, the most probable unfolding forces showed an almost perfect logarithmic dependence on the pulling velocity, indicating that great care should be taken before the linear theory of DFS is applied. A failure to fit force distributions using a Bell model was also reported in the digoxigenin-antidigoxigenin work described earlier, where poor matches

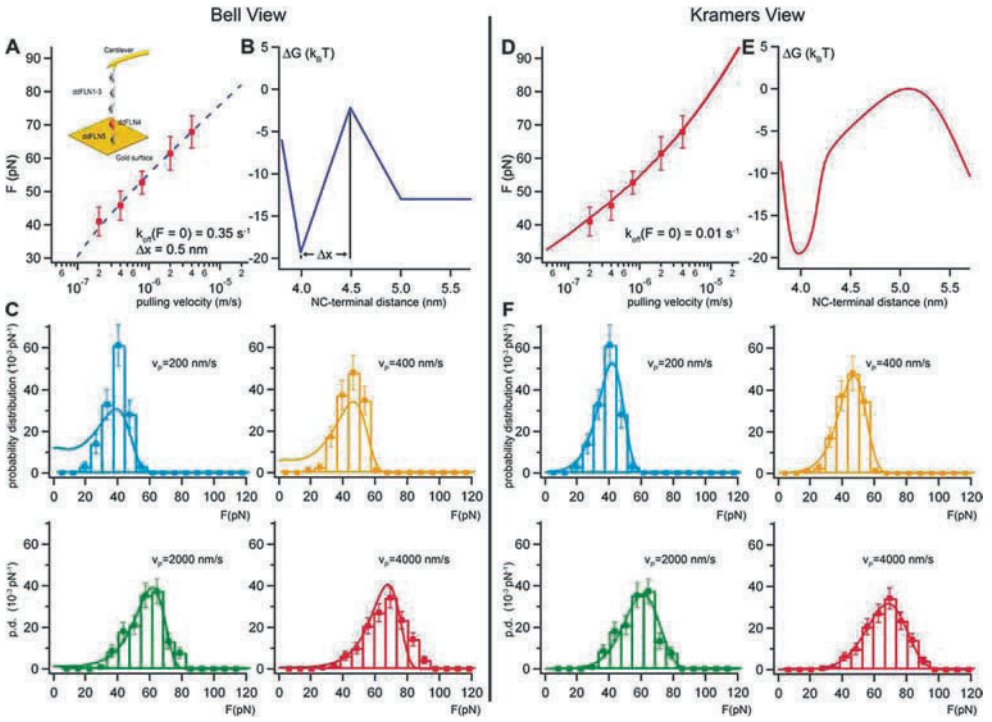


Figure 15.10. A comparison between Bell's and Kramers' views for the analysis of protein unfolding data. **A.** Force spectra of native ddFLN4 show a logarithmic behaviour within the experimental error. The inset shows a schematic illustration of the experimental setup. **B.** Reconstruction of the energy landscape using the Bell model. **C.** Measured force distributions at four different pulling velocities together with distributions calculated using the Bell model (*lines*). **D.** Force spectra modelled using Kramers' theory. **E.** The reconstructed energy landscape, which shows detailed curvature along the unfolding/folding pathway. **F.** Reproduction of force histograms using Kramers' model. (From Schlierf and M. Rief [134].)

were observed at high loading rates, as well as in the crossover region between the two linear regimes of the force spectrum. Some of these difficulties can be solved by using more detailed analysis such as the ones suggested by Klafter and coworkers [108] and Evstigeenv and Reimann [135], which are described in detail in Section 15.2.5.

DFS can be used to extract the number of bonds involved in an interaction by using the microscopic binding model of Katletz and Titulear [111,112,136] (see also Section 15.2.4). Such analyses can also yield information about cooperativity and activation barrier height. A first application of this model was the investigation of hen egg-white lysozyme (HEL) interacting with an antibody (HyHEL-10) [137]. It was found that the interaction consists of 6 cooperative groups, each made up of 3 bonds, for a total of 18 bonds. This latter value is in close agreement with results obtained from X-ray crystallography and molecular dynamics (MD) simulations, which indicated around 20 bonds. It is known that the antibody consists of six recognition sites. Whether all six sites contribute to binding is undetermined. Alanine screening indicated that at least four sites play a role in binding, whereas MD simulations found that all six recognition sites participate in the interaction, in agreement with the DFS data. The activation energy determined by DFS was lower than that derived from surface plasmon resonance measurements. A possible explanation for this discrepancy is that recognition

of lysozyme by its antibody is a two-step process. First an encounter state is established, which decays into the docking state. This decay takes place on a time scale comparable to that of the force pulling experiment. Therefore, most likely, the energetically higher encounter state was probed rather than the thermodynamically preferred docking state.

Another application of this theory was the interpretation of force spectroscopy data of human rhinoviruses (HRV) interacting with the very low density lipoprotein receptor (VLDLR) [138]. VLDLR exhibits a binding domain consisting of eight binding modules. To probe the contribution of these repeats, rupture forces of HRV2 interacting with artificial concatamers of VLDLR were measured. Force spectroscopy was carried out with different receptor constructs containing all eight imperfect binding repeats (V1 to V8), the first three binding repeats (V1, V2, and V3), and two (V33), three (V333), and five (V33333) copies of the third binding repeat fused to maltose binding protein and to a His_6 tag. The rupture force distributions of concatemers consisting of two (V33), three (V333), and five (V33333) copies of the third binding repeat were successfully fit to the statistical binding model. The numbers of cooperative blocks were equal to the number of copies of the third repeat. Every block consisted of six to seven bonds, which agreed well with X-ray crystallography data [139]. The rupture forces for V1 to V3 and V33 were similar, indicating that only two modules were bound, in agreement with neutralisation studies [140], which showed that the first binding repeat does not bind. Similar rupture forces for V1 to V8 and V33333 were found too, indicating that only five repeats of V1 to V8 bind to the virus.

15.3.2. Energy Landscape Roughness of Protein Binding Reactions

The "new view" of proteins [141] sees protein folding and binding as diffusion-like processes in which individual molecules reach their final state by gliding down the slopes of a funnel-shaped free energy landscape along multiple coexisting routes [8,142–147]. These energy landscapes are generally not smooth but rather are traversed by energy barriers of varying heights that render them highly corrugated or rugged (Figure 15.11A). This is because proteins have a complex, hierarchical structure whose energy minimization requires the simultaneous relief of multiple constraints, which are often imposed by competing processes occurring on different length and/or time scales [143,148–152]. Although proteins have evolved such that most of these constraints are minimized, some residual frustration—energetic or topological—always remains, leading to roughness [153,154]. Moreover, unlike innate systems in which the origin of frustration is purely physicochemical, frustration in protein energy landscapes has an additional source, which is imposed by functional constraints [155]. For example, the burial of polar or charged groups inside the hydrophobic core of globular proteins is energetically highly frustrating, yet such residues are often critical for binding. Another example is loops. The presence of loops in the structure acts to increase backbone flexibility, allowing breathing and correlated motions required for binding or catalysis. However, the introduction of loops is accompanied by an entropic penalty, causing a "localized" search problem due to the multiplicity of conformations that can be accommodated in these regions. This, in turn, decreases the rate of folding and may decrease thermodynamic stability.

The presence of multiple barriers in the free energy surface imposed by frustration should have strong effects on the folding and binding of proteins as well as on their equilibrium behaviour. Superimposed on the slopes of the funnel, the barriers are expected to attenuate transition rates. Present at the bottom of the funnel, they determine transition dynamics between conformational isomers and substates. Fluctuations between these isomers or

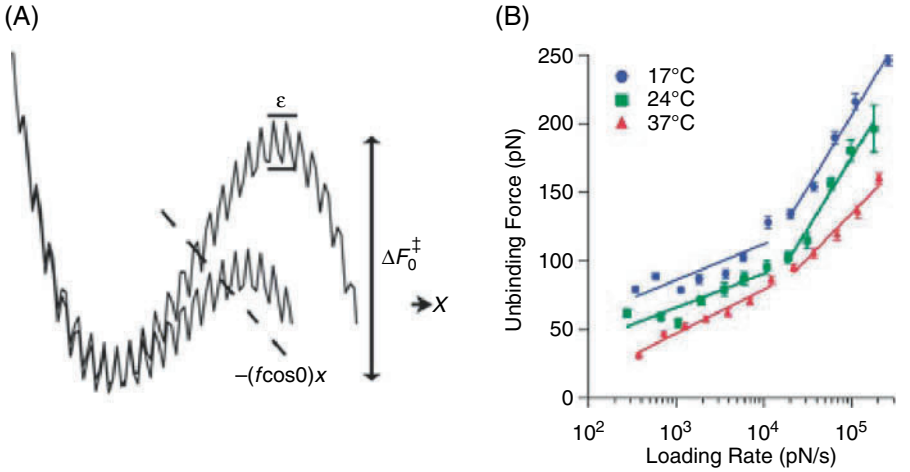


Figure 15.11. Measurement of energy landscape roughness. **A.** A rough energy landscape of an unbinding reaction driven by external mechanical force. Here ε and $\Delta F_0^{\ddagger}$ are the overall scale of roughness and the height of the activation barrier, respectively. (From Nevo et al. [160].) (**B**) Dynamic force spectra of the streptavidin–biotin interaction obtained at three different temperatures that were used to extract roughness using Eq. (15.54). The force spectra clearly show the presence of two barriers at the high (inner barrier) and low (outer barrier) loading rates. (From Rico and Moy [161].)

substates, in turn, define functional motions, such as those involved in catalysis or ligand exchange. They also allow proteins to accommodate different conformations and, hence, to have different properties that are important to their function.

From the foregoing description, it is clear that a full understanding of protein folding, binding, and function requires that the roughness of the underlying energy surfaces be known. However, experimental determination of roughness proved to be highly challenging. Thus, although the perception that protein energy landscapes are rough emerged in the mid 1980s, it was only in 2000 that the overall scale of roughness ε was estimated indirectly for the first time. This was achieved by following the rate of loop closure for flexible peptides, from which a value of $\varepsilon/k_B T \approx 2$ was inferred [156]. Three years later a theory suggesting how to measure ε directly from force spectroscopy measurements was put forward, and two additional years later this theory was tested experimentally.

The theory, published in 2003 by Hyeon and Thirumalai [157], is based on an earlier analysis made by Zwanzig [158], who applied mean-first-passage-time analysis to study diffusion in rough (1D) potentials. The results indicated that in such potentials the diffusion coefficient D is replaced by an effective diffusion coefficient D^*, which can be significantly smaller than D because of roughness. The effect was found to be particularly dramatic when the fluctuations in the potential have a Gaussian distribution, leading to the following relation:

$$D^* = D \exp(-\varepsilon^2 \beta^2) \tag{15.51}$$

where $\beta = 1/k_B T$. This expression approximates the dependence of D^* on roughness for several other distributions as well, provided that εβ is relatively small [159].

Extending Zwanzig's analysis, it was shown by Hyeon and Thirumalai that the rate of protein unfolding or unbinding induced by the application of a constant force is retarded

by roughness. Specifically, using Kramers' theory, they showed that (assuming a Gaussian distribution of ε) the unfolding/unbinding rate under a constant force $k_u(f)$, is given by

$$\log k_u(f) = \log\ k_0 - \beta(\Delta\, F^{\ddagger} - f\Delta x) - \varepsilon^2\beta^2 \tag{15.52}$$

where $\Delta F^{\ddagger}$ is the height of the activation energy barrier for the (two-state) transition, Δx is the characteristic length scale of roughness, taken to be close to x_β [160], and $k_0 = \omega_{ts}(f)\omega_0(f)/2\pi\gamma$, with ω_0 and ω_{ts} denoting the angular frequencies at the minimum and transition states, respectively, and $\boldsymbol{\gamma}$ the damping relaxation rate. Roughness thus manifests itself as a $1/T^2$ dependence of the unfolding rate, and therefore its amplitude can be extracted if log $k_u(f)$ is measured over a range of temperatures. Note that the term $\varepsilon^2\beta^2$ can lead to significant deviations from Arrhenius behaviour, particularly at low temperatures.

As discussed earlier in this chapter, most DFS experiments are not performed under a constant force but instead use a constant loading rate. The result obtained by Hyeon and Thirumalai indicates that, for this case, roughness manifests itself as an increase in the most probable force for unfolding/unbinding f^*. Provided that $\Delta F^{\ddagger}$, Δx, and k_0 are only weakly dependent on the applied force, f^*scales with ε as follows:

$$f^* = \frac{k_B T}{\Delta x\,(f^*)}\left\{\log\left(\frac{r_f \Delta x\,(f^*)}{\nu_D e^{-\beta\Delta F_0^{\ddagger}(f^*)} k_B T}\right) + \frac{\varepsilon^2}{\Delta x k_B T}\right\} \tag{15.53}$$

Following this approximation, the roughness scale can be obtained by comparing the loading rates that give rise to identical f^* values at two different temperatures, T_1 and T_2. Slightly modifying the result to account for temperature dependence of Δx (x_β), which is often observed for proteins (i.e., Hammond or anti-Hammond behaviour), Nevo et al. obtained the following expression [160]:

$$\begin{aligned}\varepsilon^2 \approx \frac{x_\beta\,(T)\,k_B T_1 x_\beta\,(T_2)\,k_B T_2}{x_\beta\,(T_2)\,k_B T_2 - x_\beta\,(T_1)\,k_B T_1}\Bigg[\Delta F_0^{+}\left(\frac{1}{x_\beta\,(T_1)} - \frac{1}{x_\beta\,(T_2)}\right)\\ + \frac{k_B T_1}{x_B\,(T_1)}\ln\frac{r_f\,(T_1)\,x_\beta\,(T_1)}{\nu_D\,(T_1)\,k_B T_1} - \frac{k_B T_2}{x_\beta\,(T_2)}\ln\frac{r_f\,(T_2)\,x_\beta\,(T_2)}{\nu_D\,(T_2)\,k_B T_2}\Bigg]\end{aligned} \tag{15.54}$$

Hyeon and Thirumalai's theory was first used to derive the roughness of the energy landscape underlying the interaction between two proteins belonging to the soluble machinery that mediates the transport of macromolecules between the nucleus and cytoplasm in interphase eukaryotic cells. One of the proteins was the prototypic nuclear import receptor importin β1 (impβ), and the other was the abundant and highly conserved GTPase Ran, which serves as a master regulator of nucleocytoplasmic exchange and is involved in various mitotic processes. Following Eq. (15.54), measurements were performed at different loading rates, spanning almost three orders of magnitude in scale, and at three temperatures: 7°C, 25°C, and 32°C. Justifying the modification made by Nevo et al., analysis of the slopes of the force spectra indicated a movement of the transition state towards the bound state as temperature increases. Values of ε were derived from nine f^*-T pair combinations and were found to be nearly the same, yielding a value of $\sim 6k_B T$, which is about one fifth of the height of the major barrier for unfolding, $\Delta F^{\ddagger}$. The rather bumpy energy surface was attributed to the unusually high flexibility of impβ, which is required to allow this protein to interact avidly with multiple ligands that have grossly different structures and binding sites over it.

More recently, the foregoing analysis was applied to derive ε for the well-studied streptavidin–biotin interaction [161]. The temperatures used in the study were 17°C, 24°C, and 37°C and the pulling speed ranged from ∼100 to 38,000 nm/sec (Figure 15.11B). Analysis of the Bell parameters revealed a considerable widening of the inner barrier for the transition with temperature, reflecting, perhaps, softening of the dominant hydrogen bond network that stabilizes the ground state of the complex. In contrast, the position of the outer barrier did not change significantly upon increase of the temperature. Estimations of ε were made at four different forces—75, 90, 135, and 156 pN—with the first two forces belonging to the first linear loading regime of the force spectrum (outer barrier) and the last two to the second (inner barrier). The values obtained were consistent *within each of the two regimes*, averaging at $7.5k_BT$ and $\sim 5.5k_BT$ along the outer and inner barriers of the transition, respectively. The difference was attributed to contributions from the intermediate state of the reaction, which is suppressed (along with outer barrier) at high loading rates. The origin of roughness was attributed to a competition of solvent water molecules with some of the hydrogen bonds that stabilize the complex and to the aforementioned "3–4" loop of streptavidin, which is highly flexible and therefore may induce the formation of multiple conformational substates in the complex. It was also proposed by the authors that the large roughness detected in the energy landscape of streptavidin–biotin is a significant contributor to the unusually slow dissociation kinetics of the complex and may account for the discrepancies in the unbinding forces measured for this pair.

In addition to the foregoing work, the theory of Hyeon and Thirumalai has also been used to derive the roughness of energy landscapes underlying *unfolding* reactions. Two such reactions were analyzed. The first was the forced unfolding of the fourth domain of the actin cross-linking protein filamin of *Dictyostelium discoideum* (ddFLN4) [162]. For this domain, which exhibited an anti-Hammond behaviour, application of Eq. (15.54) to the data yielded $\varepsilon/k_BT \approx 4$. The authors, however, preferred to interpret the data in terms of a temperature-dependant switch in the nature of the interactions that control the response of the domain to the mechanical force rather than by effects produced by roughness. The other study used the experimental scheme proposed by Hyeon and Thirumalai to derive the energy landscape roughness of single transmembrane helices of the archaeal light-driven proton pump bacteriorhodopsin. The analysis, performed on five of the seven helices of the protein, yielded roughness values ranging from $\sim 4k_BT$ to $\sim 6k_BT$ [163]. The somewhat unexpectedly high magnitude of roughness obtained for these secondary structures was suggested to reflect topological and curvature-associated frustration imposed by the surrounding lipid bilayers.

15.3.3. Discrimination between Modes of Protein Activation

We conclude this part of the chapter by describing two experiments in which single-molecule mechanical unbinding was applied to distinguish between alternative modes of protein activation. The activation of proteins by other proteins or by small ligands is a key process in biology: Signalling pathways, enzyme activity, gating of ion channels, and gene activation and inactivation often depend on the switching of proteins between different functional states. Two general mechanisms have been proposed to underlie this switching. The induced-fit model assigns changes in protein activity to conformational changes triggered by effector binding [164,165]. The population-shift model, on the other hand, ascribes these changes to a redistribution of *preexisting* conformational isomers [166]. According to this model, known also as the preequilibrium or conformational selection model, switching occurs through a

shift in equilibrium between the existing isomers induced by the binding of the ligand to the conformation most complementary to it [166].

Distinguishing between the two models of activation by conventional bulk methods is typically hard because it requires that the distribution of conformational isomers in the ensemble be known *a priori*. This information in turn is hard to derive from data obtained from bulk assays because of ensemble averaging. If the populations are skewed, an accurate probing of the ensemble becomes even harder. Consequently, there is relatively little experimental evidence that distinguishes between the two activation modes (for notable exceptions, see Refs. 167 and 168).

In an attempt to overcome the limitations associated with ensemble analyses, Nevo and coworkers used DFS to study the activation of the complex discussed in the previous section between impβ and the Ran [169,170]. The effector was the small Ran-binding protein RanBP1, which functions in the dissociation of the complex in the cytoplasm. In the absence of the effector, the complex between impβ and Ran loaded with a guanosine triphosphate (GTP) analog populates two distinct states characterized by relatively low and high adhesion strength, respectively. Notably, addition of free RanBP1 to the liquid cell diminished the higher-strength population. Concomitantly, the lower-strength conformation became more populated. The *means* of the distributions, however, did not change, indicating that the effector had not altered the strength of the interaction in the two states. These results fit a dynamic population-shift mechanism in which RanBP1 binds selectively to the lower-strength conformation of RanGTP–impβ, changing the properties and function of the complex by shifting the equilibrium between its two states.

Continuing the analysis, the authors repeated the foregoing measurements, but this time with GDP-loaded Ran. Loaded as such, Ran associates very weakly with impβ to form a single bound state characterized by unimodal distributions of unbinding forces. Addition of RanBP1 to this complex resulted in a marked shift of the distributions to higher unbinding forces. It also led to a significant increase in the number of unbinding events recorded for the complex, reflecting the ability of the effector to facilitate association between the two binding partners. These results are consistent with an induced-fit mechanism in which the binding of RanBP1 induces a conformational change in the complex, which, in turn, promotes and strengthens the interaction between impβ and RanGDP.

In a more recent study, Ritco-Vonsovici et al. [171] used DFS and MD simulations to elucidate the molecular basis underlying different activities of β-catenin, a multifunctional protein that plays essential roles in two related cellular processes. The first is cell adhesion and cytoskeletal organization, and the second is cell differentiation and development through the Wnt signalling pathway, in which it acts as a transcriptional coactivator. It was hypothesized by the authors that the two cellular functions are mediated by distinct molecular forms of β-catenin. To test this hypothesis, they studied the interaction of β-catenin with its Wnt-dependent target, transcription factor Tcf4. As in the case of impβ–Ran(GTP), unbinding of Tcf4 from β-catenin gave rise to bimodal force distributions, indicating the presence of two populations of molecular complexes. Each population exhibited a distinct dependence on the loading rate, yielding unconventional force spectra containing two independent curves (Figure 15.12). Upon addition of inhibitor of β-catenin and T cell factor (ICAT), a selective inhibitor of the interaction between β-catenin and Tcf4, the fractional occupancy of the low-strength state was reduced by 30%–40% at low loading rates and vanished altogether at high pulling speeds, strongly suggesting that binding of the inhibitor follows conformational selection mechanism. Comparative structural analysis and molecular dynamics simulations suggested that the two complexes of β-catenin–Tcf4 originate from two conformational

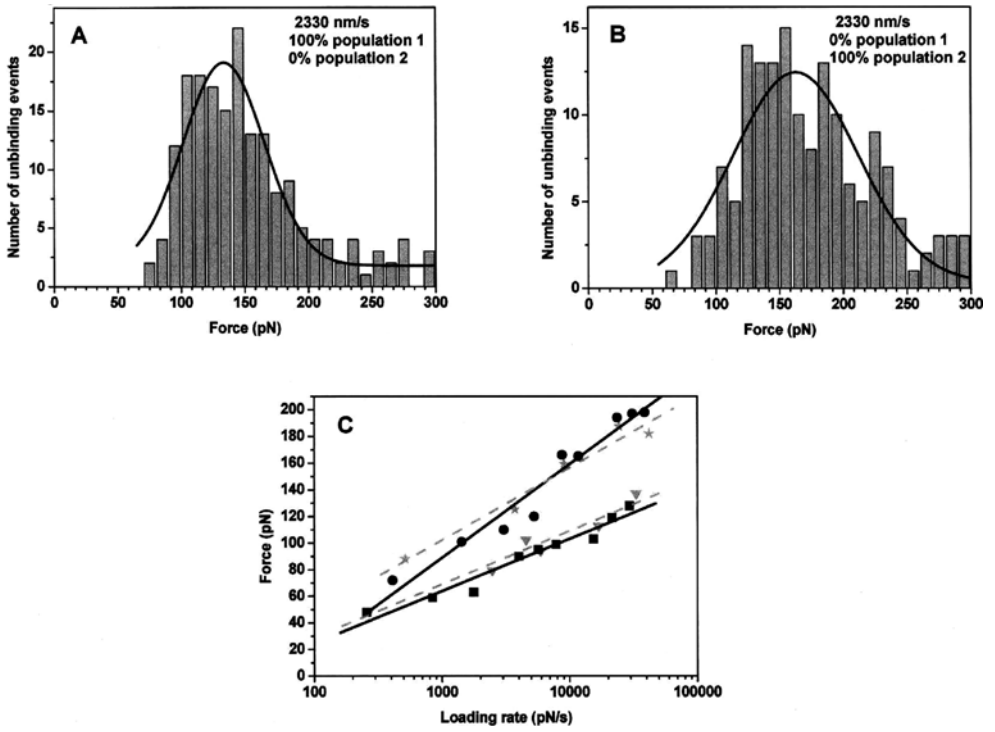


Figure 15.12. Molecular plasticity of β-catenin: New insights from single-molecule measurements. A population shift is induced by the presence of inhibitor of β-catenin and T cell factor (ICAT), as seen in the statistical distribution and force spectra of unbinding forces for the armadillo β-catenin/Tcf4 interaction. The presence of one population of molecules in each histogram is evident by the fact that they have a single maximum. **A.** The low-strength complex is populated in the absence of ICAT. **B**. However, in the presence of ICAT the high-strength population is favoured. **C.** Force spectra corresponding to the two populations of bound armadillo β-catenin/Tcf4 complexes in the presence (*dashed lines*) or absence of ICAT (*solid lines*). (From Ritco-Vonsovici et al. [171].)

isomers of β-catenin differing mainly in flexible regions of its binding site. *In vivo*, isomers of β-catenin, like those detected in this work or others, may similarly interconvert when encountering different ligands, giving rise to different functional outcomes.

References

1. Hunjan, J., A. Tovchigrechko, Y. Gao, and I. A. Vakser. 2008. The size of the intermolecular energy funnel in protein–protein interactions. Proteins 72:344–352.
2. Levy, Y., S. S. Cho, J. N. Onuchic, and P. G. Wolynes. 2005. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J Mol Biol 346:1121–1145.
3. Miller, D. W., and K. A. Dill. 1997. Ligand binding to proteins: the binding landscape model. Protein Sci 6:2166–2179.
4. O'Toole, N., and I. A. Vakser. 2008. Large-scale characteristics of the energy landscape in protein–protein interactions. Proteins 71:144–152.
5. Thielges, M. C., J. Zimmermann, W. Yu, M. Oda, and F. E. Romesberg. 2008. Exploring the energy landscape of antibody–antigen complexes: protein dynamics, flexibility, and molecular recognition. Biochemistry 47:7237–7247.

6. Tovchigrechko, A., and I. A. Vakser. 2001. How common is the funnel-like energy landscape in protein–protein interactions? Protein Sci 10:1572–1583.
7. Wang, J., L. Xu, and E. Wang. 2007. Optimal specificity and function for flexible biomolecular recognition. Biophys J 92:L109–111.
8. Wolynes, P. G. 2005. Recent successes of the energy landscape theory of protein folding and function. Q Rev Biophys 38:405–410.
9. Bustamante, C., Y. R. Chemla, N. R. Forde, and D. Izhaky. 2004. Mechanical processes in biochemistry. Annu Rev Biochem 73:705–748.
10. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. Biophys J 72:1541–1555.
11. Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. 1997. Molecular dynamics study of unbinding of the avidin–biotin complex. Biophys J 72:1568–1581.
12. Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. Science 276:1109–1112.
13. Benoit, M., D. Gabriel, G. Gerisch, and H. E. Gaub. 2000. Discrete interactions in cell adhesion measured by single-molecule force spectroscopy. Nat Cell Biol 2:313–317.
14. Li, F., S. D. Redick, H. P. Erickson, and V. T. Moy. 2003. Force measurements of the $\alpha_5\beta_1$ integrin-fibronectin interaction. Biophys J 84:1252–1262.
15. Baumgartner, W., N. Golenhofen, N. Grundhofer, J. Wiegand, and D. Drenckhahn. 2003. Ca2+ Dependency of N-cadherin function probed by laser tweezer and atomic force microscopy. J Neurosci 23:11008–11014.
16. Fritz, J., A. G. Katopodis, F. Kolbinger, and D. Anselmetti. 1998. Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy. Proc Natl Acad Sci USA 95:12283–12288.
17. Chen, A., and V. T. Moy. 2000. Cross-linking of cell surface receptors enhances cooperativity of molecular adhesion. Biophys J 78:2814–2820.
18. Wong, J., A. Chilkoti, and V. T. Moy. 1999. Direct force measurements of the streptavidin–biotin interaction. Biomol Eng 16:45–55.
19. Riener, C. K., C. M. Stroh, A. Ebner, C. Klampfl, A. A. Gall, C. Romanin, Y. L. Lyubchenko, P. Hinterdorfer, and H. J. Gruber. 2003. Simple test system for single molecule recognition force microscopy. Anal Chim Acta 479:59–75.
20. Ebner, A., P. Hinterdorfer, H. J. Gruber. 2007. Comparison of different aminofunctionalization strategies for attachment of single antibodies to AFM cantilevers. Ultramicroscopy 107:922–927.
21. Ebner, A., F. Kienberger, C. Huber, A. S. Kamruzzahan, V. P. Pastushenko, J. Tang, G. Kada, H. J. Gruber, U. B. Sleytr, M. Sara, and P. Hinterdorfer. 2006. Atomic-force-microscopy imaging and molecular-recognition-force microscopy of recrystallized heterotetramers comprising an S-layer-streptavidin fusion protein. Chembiochem 7:588–591.
22. Lee, G. U., D. A. Kidwell, and R. J. Colton. 1994. Sensing discrete streptavidin–biotin interactions with atomic force microscopy. Langmuir 10:354–357.
23. De Paris, R., T. Strunz, K. Oroszlan, H.-J. Güntherrodt, and M. Hegner. 2000. Force spectroscopy and dynamics of the biotin–avidin bond studied by scanning force microscopy. Single Molecules 4:285–290.
24. Andrade, J. D., V. Hlady, A.-P. Wei, C.-H. Ho, A. S. Lea, S. I. Jeon, Y. S. Lin, and E. Stroup. 1992. Proteins at interfaces: principles, multivariate aspects, protein resistant surfaces, and direct imaging and manipulation of adsorbed proteins. Clin Mater 11:67–84.
25. Green, N. M. 1990. Avidin and streptavidin. Methods Enzymol 184:51–67.
26. Sen, S., S. Subramanian, and D. E. Discher. 2005. Indentation and adhesive probing of a cell membrane with AFM: theoretical model and experiments. Biophys J 89:3203–3213.
27. Almqvist, N., R. Bhatia, G. Primbs, N. Desai, S. Banerjee, and R. Lal. 2004. Elasticity and adhesion force mapping reveals real-time clustering of growth factor receptors and associated changes in local cellular rheological properties. Biophys J 86:1753–1762.
28. Scheuring, S., D. J. Muller, P. Ringler, J. B. Heymann, and A. Engel. 1999. Imaging streptavidin 2D crystals on biotinylated lipid monolayers at high resolution with the atomic force microscope. J Microsc 193:28–35.
29. Bustamante, C., J. Vesenka, C. L. Tang, W. Rees, M. Guthold, and R. Keller. 1992. Circular DNA molecules imaged in air by scanning force microscopy. Biochemistry 31:22–26.
30. Feng, X. Z., R. Bash, P. Balagurumoorthy, D. Lohr, R. E. Harrington, and S. M. Lindsay. 2000. Conformational transition in DNA on a cold surface. Nucl Acids Res 28:593–596.
31. Stroh, C., H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S. M. Lindsay, and P. Hinterdorfer. 2004. Single-molecule recognition imaging-microscopy. Proc Natl Acad Sci USA 101:12503–12507.
32. Baumgartner, W., P. Hinterdorfer, W. Ness, A. Raab, D. Vestweber, H. Schindler, and D. Drenckhahn. 2000. Cadherin interaction probed by atomic force microscopy. Proc Natl Acad Sci USA 97:4005–4010.

33. Ebner, A., F. Kienberger, G. Kada, C. M. Stroh, M. Geretschlager, A. S. M. Kamruzzahan, L. Wildling, W. T. Johnson, B. Ashcroft, J. Nelson, S. M. Lindsay, H. J. Gruber, and P. Hinterdorfer. 2005. Localization of single avidin-biotin interactions using simultaneous topography and molecular recognition imaging. ChemPhysChem 6:897–900.
34. Stroh, C. M., A. Ebner, M. Geretschlager, G. Freudenthaler, F. Kienberger, A. S. M. Kamruzzahan, S. J. Smith-Gill, H. J. Gruber, and P. Hinterdorfer. 2004. Simultaneous topography and recognition imaging using force microscopy. Biophys J 87:1981–1990.
35. Willemsen, O., M. Snel, K. O. van der Werf, B. G. de Grooth, J. Greve, P. Hinterdorfer, H. Gruber, H. Schindler, Y. van Kooyk, and C. Figdor. 1998. Simultaneous height and adhesion imaging of antibody–antigen interactions by atomic force microscopy. Biophys J 75:2220–2228.
36. Chen, X., M. C. Davies, C. J. Roberts, S. J. B. Tendler, P. M. Williams, J. Davies, A. C. Dawkes, and J. C. Edwards. 1997. Recognition of protein adsorption onto polymer surfaces by scanning force microscopy and probe-surface adhesion measurements with protein-coated probes. Langmuir 13:4106–4111.
37. Cross, B., F. Ronzon, B. Roux, and J. P. Rieu. 2005. Measurement of the anchorage force between GPI-anchored alkaline phosphatase and supported membranes by AFM force spectroscopy. Langmuir 21: 5149–5153.
38. Grange, W., T. Strunz, I. Schumakovitch, H.-J. Güntherodt, and M. Hegner. 2001. Molecular recognition and adhesion of individual DNA strands studied by dynamic force microscopy. Single Molecules 2:75–78.
39. Jiang, Y., C. Zhu, L. Ling, L. Wan, X. Fang, and C. Bai. 2003. Specific aptamer–protein interaction studied by atomic force microscopy. Anal Chem 75:2112–2116.
40. Ros, R., F. Schwesinger, D. Anselmetti, M. Kubon, R. Schafer, A. Pluckthun, and L. Tiefenauer. 1998. Antigen binding forces of individually addressed single-chain Fv antibody molecules. Proc Natl Acad Sci USA 95:7402–7405.
41. Schwesinger, F., R. Ros, T. Strunz, D. Anselmetti, H. Guntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, and A. Pluckthun. 2000. Unbinding forces of single antibody–antigen complexes correlate with their thermal dissociation rates. Proc Natl Acad Sci USA 97:9972–9977.
42. Strunz, T., K. Oroszlan, R. Schafer, and H. J. Guntherodt. 1999. Dynamic force spectroscopy of single DNA molecules. Proc Natl Acad Sci USA 96:11277–11282.
43. Strunz, T., K. Oroszlan, I. Schumakovitch, H. Guntherodt, and M. Hegner. 2000. Model energy landscapes and the force-induced dissociation of ligand–receptor bonds. Biophys J 79:1206–1212.
44. Kamruzzahan, A. S. M., A. Ebner, L. Wildling, F. Kienberger, C. K. Riener, C. D. Hahn, P. D. Pollheimer, P. Winklehner, M. Holzl, B. Lackner, D. M. Schorkl, P. Hinterdorfer, and H. J. Gruber. 2006. Antibody linking to atomic force microscope tips via disulfide bond formation. Bioconjugate Chem 17:1473–1481.
45. Ebner, A., L. Wildling, A. S. M. Kamruzzahan, C. Rankl, J. Wruss, C. D. Hahn, M. Hölzl, F. Kienberger, D. Blaas, P. Hinterdorfer, and H. J. Gruber. 2007. A new, simple method for linking of antibodies to atomic force microscopy tips. Bioconjugate Chem 18:1176–1184.
46. Puntheeranurak, T., L. Wildling, H. J. Gruber, R. K. Kinne, and P. Hinterdorfer. 2006. Ligands on the string: single-molecule AFM studies on the interaction of antibodies and substrates with the Na+-glucose co-transporter SGLT1 in living cells. J Cell Sci 119:2960–2967.
47. Hinterdorfer, P., W. Baumgartner, H. J. Gruber, K. Schilcher, and H. Schindler. 1996. Detection and localization of individual antibody–antigen recognition events by atomic force microscopy. Proc Natl Acad Sci USA 93:3477–3481.
48. Raab, A., W. Han, D. Badt, S. J. Smith-Gill, S. M. Lindsay, H. Schindler, and P. Hinterdorfer. 1999. Antibody recognition imaging by force microscopy. Nat Biotechnol 17:902–905.
49. Avci, R., M. Schweitzer, R. D. Boyd, J. Wittmeyer, A. Steele, J. Toporski, W. Beech, F. T. Arce, B. Spangler, K. M. Cole, and D. S. McKay. 2004. Comparison of antibody–antigen interactions on collagen measured by conventional immunological techniques and atomic force microscopy. Langmuir 20:11053–11063.
50. Hinterdorfer, P., K. Schilcher, W. Baumgartner, H. J. Gruber, and H. Schindler. 1998. A mechanistic study of the dissociation of individual antibody–antigen pairs by atomic force microscopy. Nanobiology 4:177–188.
51. Kada, G., L. Blayney, L. H. Jeyakumar, F. Kienberger, V. P. Pastushenko, S. Fleischer, H. Schindler, F. A. Lai, and P. Hinterdorfer. 2001. Recognition force microscopy/spectroscopy of ion channels: applications to the skeletal muscle Ca2+ release channel (RYR1). Ultramicroscopy 86:129–137.
52. Kienberger, F., Kada, G., Gruber, H. J., Pastushenko, P., Riener, C., Trieb, M., Knaus, H.-G., Schindler, H., and Hinterdorfer, P. 2000. Recognition force spectroscopy studies of the NTA-His6 bond. Single Molecules 1:25–31.
53. Kienberger, F., V. P. Pastushenko, G. Kada, H. J. Gruber, C. Riener, H. Schindler, and P. Hinterdorfer. 2000. Static and dynamical properties of single poly(ethylene glycol) molecules investigated by force spectroscopy. Single Molecules 1:123–128.

54. Puntheeranurak, T., B. Wimmer, F. Castaneda, H. J. Gruber, P. Hinterdorfer, and R. K. H. Kinne. 2007. Substrate specificity of sugar transport by rabbit SGLT1: single-molecule atomic force microscopy versus transport studies. Biochemistry 46:2797–2804.
55. Wang, H., R. Bash, S. M. Lindsay, and D. Lohr. 2005. Solution AFM studies of human Swi-Snf and its interactions with MMTV DNA and chromatin. Biophys J 89:3386–3398.
56. Wieland, J. A., A. A. Gewirth, and D. E. Leckband. 2005. Single molecule adhesion measurements reveal two homophilic neural cell adhesion molecule bonds with mechanically distinct properties. J Biol Chem 280:41037–41046.
57. Johnsson, B., S. Lofas, and G. Lindquist. 1991. Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. Anal Biochem 198: 268–277.
58. Hahn, C. D., C. Leitner, T. Weinbrenner, R. Schlapak, A. Tinazli, R. Tampe, B. Lackner, C. Steindl, P. Hinterdorfer, H. J. Gruber, and M. Holzl. 2007. Self-assembled monolayers with latent aldehydes for protein immobilization. Bioconjugate Chem 18:247–253.
59. Zara, J. J., R. D. Wooda, P. Boonb, C.-H. Kimc, N. Pomatoc, R. Bredehorst, and C.-W. Vogel. 1991. A carbohydrate-directed heterobifunctional cross-linking reagent for the synthesis of immunoconjugates. Anal Biochem 194:156–162.
60. Wielert-Badt, S., P. Hinterdorfer, H. Gruber, J. Lin, D. Badt, B. Wimmer, H. Schindler, and R. Kinne. 2002. Single molecule recognition of protein binding epitopes in brush border membranes by force microscopy. Biophys J 82:2767–2774.
61. Haselgrübler, T., A. Amerstorfer, H. Schindler, and H. Gruber. 1995. Synthesis and applications of a new poly(ethylene glycol) derivative for the crosslinking of amines with thiols. Bioconjugate Chem 6: 242–248.
62. Riener, C. K., F. Kienberger, C. D. Hahn, G. M. Buchinger, I. O. C. Egwim, T. Haselgrubler, A. Ebner, C. Romanin, C. Klampfl, B. Lackner, and H. J. Gruber. 2003. Heterobifunctional crosslinkers for tethering single ligand molecules to scanning probes. Anal Chim Acta 497:101–114.
63. Nevo, R., Stroh, C., Kienberger, F., Kaftan, D., Brumfeld, V., Elbaum, M., Reich, Z., and Hinterdorfer, P. 2003. A molecular switch between alternative conformational states in the complex of Ran and importin β1. Nat Struct Biol 10:553–557.
64. Hölzl, M., A. Tinazli, C. Leitner, C. D. Hahn, B. Lackner, R. Tampe, and H. J. Gruber. 2007. Protein-resistant self-assembled monolayers on gold with latent aldehyde functions. Langmuir 23:5571–5577.
65. Peelen, D., and L. M. Smith. 2005. Immobilization of amine-modified oligonucleotides on aldehyde-terminated alkanethiol monolayers on gold. Langmuir 21:266–271.
66. Ebner, A., L. Wildling, A. S. Kamruzzahan, C. Rankl, J. Wruss, C. D. Hahn, M. Holzl, R. Zhu, F. Kienberger, D. Blaas, P. Hinterdorfer, and H. J. Gruber. 2007. A new, simple method for linking of antibodies to atomic force microscopy tips. Bioconjugate chemistry 18:1176–1184.
67. Langry, K. C., T. V. Ratto, R. E. Rudd, and M. W. McElfresh. 2005. The AFM measured force required to rupture the dithiolate linkage of thioctic acid to gold is less than the rupture force of a simple gold–alkyl thiolate bond. Langmuir 21:12064–12067.
68. Tournier, E. J. M., J. Wallach, and P. Blond. 1998. Sulfosuccinimidyl 4-(N-maleimidomethyl)-1-cyclohexane carboxylate as a bifunctional immobilization agent. Optimization of the coupling conditions. Anal Chim Acta 361:33–44.
69. Walsh, M. K., X. Wang, and B. C. Weimer. 2001. Optimizing the immobilization of single-stranded DNA onto glass beads. J Biochem Biophys Methods 47:221–231.
70. Karrasch, S., M. Dolder, F. Schabert, J. Ramsden, and A. Engel. 1993. Covalent binding of biological samples to solid supports for scanning probe microscopy in buffer solution. Biophys J 65:2437–2446.
71. Horn, R., S. Ding, and H. Gruber. 2000. Immobilizing the moving parts of voltage-gated ion channels. J Gen Physiol 116:461–476.
72. Hegner, M., P. Wagner, and G. Semenza. 1993. Ultralarge atomically flat template-stripped Au surfaces for scanning probe microscopy. Surf Sci 291:39–46.
73. Tinazli, A., J. Tang, R. Valiokas, S. Picuric, S. Lata, J. Piehler, B. Liedberg, and R. Tampe. 2005. High-affinity chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. Chem Eur J 11:5249–5259.
74. Wagner, P., M. Hegner, H.-J. Güntherodt, and G. Semenza. 1995. Formation and in situ modification of monolayers chemisorbed on ultraflat template-stripped gold surfaces. Langmuir 11:3867–3875.
75. Hahn, C. D., A. Tinazli, M. Hölzl, C. Leitner, F. Frederix, B. Lackner, N. Müller, C. Klampfl, R. Tampé, and H. J. Gruber. 2007. Pragmatic studies on protein-resistant self-assembled monolayers. Chem Monthly 138: 245–252.

76. Love, J. C., L. A. Estroff, J. K. Kriebel, R. G. Nuzzo, and G. M. Whitesides. 2005. Self-assembled monolayers of thiolates on metals as a form of nanotechnology. Chem Rev 105:1103–1170.
77. Harada, Y., M. Kuroda, and A. Ishida. 2000. Specific and quantized antigen-antibody interaction measured by atomic force microscopy. Langmuir 16:708–715.
78. Bonanni, B., A. S. M. Kamruzzahan, A. R. Bizzarri, C. Rankl, H. J. Gruber, P. Hinterdorfer, and S. Cannistraro. 2005. Single molecule recognition between cytochrome c 551 and gold-immobilized azurin by force spectroscopy. Biophys J 89:2783–2791.
79. Bustanji, Y., C. R. Arciola, M. Conti, E. Mandello, L. Montanaro, and B. Samori. 2003. Dynamics of the interaction between a fibronectin molecule and a living bacterium under mechanical force. Proc Natl Acad Sci USA 100:13292–13297.
80. Conti, M., G. Donati, G. Cianciolo, S. Stefoni, and B. Samorì. 2002. Force spectroscopy study of the adhesion of plasma proteins to the surface of a dialysis membrane: role of the nanoscale surface hydrophobicity and topography. J Biomed Mater Res 61:370–379.
81. Dean, D., L. Han, A. J. Grodzinsky, and C. Ortiz. 2006. Compressive nanomechanics of opposing aggrecan macromolecules. J Biomech 39:2555–2565.
82. Gad, M., A. Itoh, and A. Ikai. 1997. Mapping cell wall polysaccharides of living microbial cells using atomic force microscopy. Cell Biol Int 21:697–706.
83. Boozer, C., S. Chen, and S. Jiang. 2006. Controlling DNA orientation on mixed ssDNA/OEG SAMs. Langmuir 22:4694–4698.
84. Green, N. H., P. M. Williams, O. Wahab, M. C. Davies, C. J. Roberts, S. J. B. Tendler, and S. Allen. 2004. Single-molecule investigations of RNA dissociation. Biophys J 86:3811–3821.
85. Noy, A., D. V. Vezenov, J. F. Kayyem, T. J. Meade, and C. M. Lieber. 1997. Stretching and breaking duplex DNA by chemical force microscopy. Chem Biol 4:519–527.
86. Auletta, T., M. R. de Jong, A. Mulder, F. C. van Veggel, J. Huskens, D. N. Reinhoudt, S. Zou, S. Zapotoczny, H. Schonherr, G. J. Vancso, and L. Kuipers. 2004. β-Cyclodextrin host–guest complexes probed under thermodynamic equilibrium: thermodynamics and AFM force spectroscopy. J Am Chem Soc 126: 1577–1584.
87. Schönherr, H., M. W. J. Beulen, J. Bugler, J. Huskens, F. van Veggel, D. N. Reinhoudt, and G. J. Vancso. 2000. Individual supramolecular host–guest interactions studied by dynamic single molecule force spectroscopy. J Am Chem Soc 122:4963–4967.
88. Zapotoczny, S., T. Auletta, M. R. de Jong, H. Schonherr, J. Huskens, F. C. J. M. van Veggel, D. N. Reinhoudt, and G. J. Vancso. 2002. Chain length and concentration dependence of β-cyclodextrin–ferrocene host–guest complex rupture forces probed by dynamic force spectroscopy. Langmuir 18:6988–6994.
89. Touhami, A., M. H. Jericho, and T. J. Beveridge. 2007. Molecular recognition forces between immunoglobulin G and a surface protein adhesin on living *Staphylococcus aureus*. Langmuir 23:2755–2760.
90. Touhami, A., B. Hoffmann, A. Vasella, F. A. Denis, and Y. F. Dufrene. 2003. Probing specific lectin-carbohydrate interactions using atomic force microscopy imaging and force measurements. Langmuir 19: 1745–1751.
91. Ebner, A., L. Wildling, R. Zhu, C. Rankl, T. Haselgruuebler, P. Hinterdorfer, and H. J. Gruber. 2008. Functionalization of probe tips and supports for single-molecule recognition force spectroscopy. Top Curr Chem 285:29–76.
92. Muller, D. J., and A. Engel. 1997. The height of biomolecules measured with the atomic force microscope depends on electrostatic interactions. Biophys J 73:1633–1644.
93. Gibson, C. T., G. S. Watson, and S. Myhra. 1996. Determination of the spring constants of probes for force microscopy/spectroscopy. Nanotechnology 7:259–262.
94. Cleveland, J. P., S. Manne, D. Bocek, and P. K. Hansma. 1993. A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. Rev Sci Instrum 64:403–405.
95. Hutter, J. L., and J. Bechhoefer. 1993. Calibration of atomic-force microscope tips. Rev Sci Instrum 64:1868–1873.
96. Sader, J. E., J. Pacifico, C. P. Green, and P. Mulvaney. 2005. General scaling law for stiffness measurement of small bodies with applications to the atomic force microscope. J Appl Phys 97:124903–124907.
97. Ohler, B. 2007. *Practical Advice on the Determination of Cantilever Spring Constants*. Veeco Instruments, Plainview, NY.
98. Butt, H. J., and M. Jaschke. 1995. Calculation of thermal noise in atomic force microscopy. Nanotechnology 6:1–7.
99. Sader, J. E. 1998. Frequency response of cantilever beams immersed in viscous fluids with applications to the atomic force microscope. J Appl Phys 84:64–76.

100. Sader, J. E. 2008. Atomic force microscope cantilevers (Calibration method of Sader). Available at: http://www.ampc.ms.unimelb.edu.au/afm/. Accessed February 23, 2009.
101. Rosenblatt, M. 1956. Remarks on some nonparametric estimates of a density function. Ann Math Statist 27:832–837.
102. Baumgartner, W., P. Hinterdorfer, and H. Schindler. 2000. Data analysis of interaction forces measured with the atomic force microscope. Ultramicroscopy 82:85–95.
103. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. Biophys J 72:1541–1555.
104. Bell, G. I. 1978. Models for the specific adhesion of cells to cells. Science 200:618–627.
105. Evans, E., and K. Ritchie. 1999. Strength of a weak bond connecting flexible polymer chains. Biophys J 76:2439–2447.
106. Friedsam, C., A. K. Wehle, F. Kühner, and H. E. Gaub. 2003. Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length. J Phys: Condens Matter 15:S1709–S1723.
107. Hummer, G., and A. Szabo. 2003. Kinetics from nonequilibrium single-molecule pulling experiments Biophys J 85:5–15.
108. Dudko, O. K., A. E. Filippov, J. Klafter, and M. Urbakh. 2003. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. Proc Natl Acad Sci USA 100:11378–11381.
109. Dudko, O. K., G. Hummer, and A. Szabo. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. Phys Rev Lett 96:108101–108104.
110. Abramowitz, M., and I. A. Stegun. 1964. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.
111. Katletz, S., and U. M. Titulaer. 1999. A statistical model for antibody–antigeb binding. Condens Matter Phys 2:361–268.
112. Katletz, S. 2002. A statistical treatment of the dissociation of cooperative systems: application to the antibody–antigen system. Thesis, University of Linz, Linz, Austria.
113. Farkas, L. 1927. Keimbildungsgeschwindigkeit inubersattigten Dampfen. Z Physik Chem 125:236–242.
114. Yuan, C., A. Chen, P. Kolb, and V. T. Moy. 2000. Energy landscape of streptavidin-biotin complexes measured by atomic force microscopy. Biochemistry 39:10219–10223.
115. Jarzynski, C. 1997. Nonequillibrium equality for free energy differences. Phys Rev Lett 78:2690–2693.
116. Jarzynski, C. 1997. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. Phys Rev E 56:5018–5035.
117. Hummer, G., and A. Szabo. 2001. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. Proc Natl Acad Sci USA 98:3658–3661.
118. Hummer, G., and A. Szabo. 2005. Free energy surfaces from single-molecule force spectroscopy. Acc Chem Res 38:504–513.
119. Preiner, J., H. Janovjak, C. Rankl, H. Knaus, D. A. Cisneros, A. Kedrov, F. Kienberger, D. J. Muller, and P. Hinterdorfer. 2007. Free energy of membrane protein unfolding derived from single-molecule force measurements. Biophys J 93:930.
120. Hoh, J. H., J. P. Cleveland, C. B. Prater, J. P. Revel, and P. K. Hansma. 1992. Quantized adhesion detected with the atomic force microscope. J Am Chem Soc 114:4917–4918.
121. Merkel, R., P. Nassoy, A. Leung, K. Ritchie, and E. Evans. 1999. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. Nature 397:50–53.
122. Moy, V. T., E. L. Florin, and H. E. Gaub. 1994. Intermolecular forces and energies between ligands and receptors. Science 266:257–259.
123. Chilkoti, A., T. Boland, B. D. Ratner, and P. S. Stayton. 1995. The relationship between ligand-binding thermodynamics and protein–ligand interaction forces measured by atomic force microscopy. Biophys J 69:2125–2130.
124. Grubmuller, H., B. Heymann, and P. Tavan. 1996. Ligand binding: molecular mechanics calculation of the streptavidin–biotin rupture force. Science 271:997–999.
125. Yuan, C., A. Chen, P. Kolb, and V. T. Moy. 2000. Energy landscape of streptavidin–biotin complexes measured by atomic force microscopy. Biochemistry 39:10219–10223.
126. Hanley, W. D., D. Wirtz, and K. Konstantopoulos. 2004. Distinct kinetic and mechanical properties govern selectin–leukocyte interactions. J Cell Sci 117:2503–2511.
127. Fritz, J., A. G. Katopodis, F. Kolbinger, and D. Anselmetti. 1998. Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy. Proc Natl Acad Sci USA 95:12283–12288.
128. Evans, E., A. Leung, D. Hammer, and S. Simon. 2001. Chemically distinct transition states govern rapid dissociation of single L-selectin bonds under force. Proc Natl Acad Sci USA 98:3784–3789.
129. Marshall, B. T., M. Long, J. W. Piper, T. Yago, R. P. McEver, and C. Zhu. 2003. Direct observation of catch bonds involving cell-adhesion molecules. Nature 423:190–193.

130. Wojcikiewicz, E. P., M. H. Abdulreda, X. Zhang, and V. T. Moy. 2006. Force spectroscopy of LFA-1 and its ligands, ICAM-1 and ICAM-2. Biomacromolecules 7:3188–3195.
131. Odorico, M., J. M. Teulon, T. Bessou, C. Vidaud, L. Bellanger, S. W. W. Chen, E. Quemeneur, P. Parot, and J. L. Pellequer. 2007. Energy landscape of chelated uranyl: Antibody interactions by dynamic force spectroscopy. Biophysical Journal 93: 645–654.
132. Schwesinger, F., R. Ros, T. Strunz, D. Anselmetti, H. J. Guntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, and A. Pluckthun. 2000. Unbinding forces of single antibody–antigen complexes correlate with their thermal dissociation rates. Proc Natl Acad Sci USA 97:9972–9977.
133. Neuert, G., C. Albrecht, E. Pamir, and H. E. Gaub. 2006. Dynamic force spectroscopy of the digoxigenin-antibody complex. FEBS Lett 580:505–509.
134. Schlierf, M., and M. Rief. 2006. Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. Biophys J 90:L33–L35.
135. Evstigneev, M., and P. Reimann. 2003. Dynamic force spectroscopy: optimized data analysis. Phys Rev E 68, 045103 1–4.
136. Katletz, S., C. Borken, H. Schindler, P. Hinterdorfer, and U. M. Titulaer. 2000. A statistical model for antibody–antigen unbinding and its application to lysozyme-anti lysozyme. Single Molecules 1:173.
137. Katletz, S., C. M. Stroh, P. Hinterdorfer, and U. Titulaer. In preparation.
138. Rankl, C., F. Kienberger, Z. Rong, H. J. Gruber, J. Wruss, D. Blaas, and P. Hinterdorfer. Concatemers of ligand binding repeat 3 of human VLDL-receptor attach to HRV2 via multiple modules: determination of the unbinding force on the single molecule level by force spectrometry. In preparation.
139. Verdaguer, N., I. Fita, M. Reithmayer, R. Moser, and D. Blaas. 2004. X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat Struct Mol Biol 11:429–434.
140. Moser, R., L. Snyers, J. Wruss, J. Angulo, H. Peters, T. Peters, and D. Blaas. 2005. Neutralization of a common cold virus by concatemers of the third ligand binding module of the VLDL-receptor strongly depends on the number of modules. Virology 338:259–269.
141. Baldwin, R. L. 1994. Protein-folding—matching speed and stability. Nature 369:183–184.
142. Bryngelson, J. D., J. N. Onuchic, N. D. Socci, and P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein-folding—a synthesis. Proteins 21:167–195.
143. Thirumalai, D., and S. A. Woodson. 1996. Kinetics of folding of proteins and RNA. Acc Chem Res 29: 433–439.
144. Dill, K. A. 1999. Polymer principles and protein folding. Protein Sci 8:1166–1180.
145. Dobson, C. M., and M. Karplus. 1999. The fundamentals of protein folding: bringing together theory and experiment. Curr Opin Struct Biol 9:92–101.
146. Kumar, S., B. Ma, C. J. Tsai, N. Sinha, and R. Nussinov. 2000. Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci 9:10–19.
147. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. Curr Opin Struct Biol 14:70–75.
148. Ansari, A., J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young. 1985. Protein states and proteinquakes. Proc Natl Acad Sci USA 82:5000–5004.
149. Bryngelson, J. D., and P. G. Wolynes. 1987. Spin glasses and the statistical mechanics of protein folding. Proc Natl Acad Sci USA 84:7524–7528.
150. Bryngelson, J. D., and P. G. Wolynes. 1989. Intermediates and barrier crossing in a random energy-model (with applications to protein folding). J Phys Chem 93:6902–6915.
151. Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. J Mol Biol 298:937–953.
152. Brockwell, D. J., and S. E. Radford. 2007. Intermediates: ubiquitous species on folding energy landscapes? Curr Opin Struct Biol 17:30–37.
153. Guo, Z. Y., and D. Thirumalai. 1995. Kinetics of protein-folding—nucleation mechanism, time scales, and pathways. Biopolymers 36:83–102.
154. Thirumalai, D., and S. A. Woodson. 2000. Maximizing RNA folding rates: A balancing act. RNA 6: 790–794.
155. Gruebele, M. 2005. Downhill protein folding: evolution meets physics. C R Biologies 328:701–712.
156. Lapidus, L. J., W. A. Eaton, and J. Hofrichter. 2000. Measuring the rate of intramolecular contact formation in polypeptides. Proc Natl Acad Sci USA 97:7220–7225.
157. Hyeon, C. B., and D. Thirumalai. 2003. Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? Proc Natl Acad Sci USA 100:10249–10253.
158. Zwanzig, R. 1988. Diffusion in a rough potential. Proc Natl Acad Sci USA 85:2029–2030.

159. Hyeon, C., and D. Thirumalai. 2007. Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. J Phys Condens Mat 19:(11)1–27.
160. Nevo, R., V. Brumfeld, R. Kapon, P. Hinterdorfer, and Z. Reich. 2005. Direct measurement of protein energy landscape roughness. EMBO Rep 6:482–486.
161. Rico, F., and V. T. Moy. 2007. Energy landscape roughness of the streptavidin–biotin interaction. J Mol Recognit 20:495–501.
162. Schlierf, M., and M. Rief. 2005. Temperature softening of a protein in single-molecule experiments. J Mol Biol 354:497–503.
163. Janovjak, H., H. Knaus, and D. J. Muller. 2007. Transmembrane helices have rough energy surfaces. J Am Chem Soc 129:246–247.
164. Koshland, D. E. 1958. Application of a theory of enzyme specificity to protein synthesis. Proc Natl Acad Sci USA 44:98–104.
165. Koshland, D. E., Jr., G. Nemethy, and D. Filmer. 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5:365–385.
166. Monod, J., J. Wyman, and J. P. Changeux. 1965. On the nature of allosteric transitions: a plausible model. J Mol Biol 12:88–118.
167. James, L. C., P. Roversi, and D. S. Tawfik. 2003. Antibody multispecificity mediated by conformational diversity. Science 299:1362–1367.
168. Volkman, B. F., D. Lipson, D. E. Wemmer, and D. Kern. 2001. Two-state allosteric behavior in a single-domain signaling protein. Science 291:2429–2433.
169. Nevo, R., C. Stroh, F. Kienberger, D. Kaftan, V. Brumfeld, M. Elbaum, Z. Reich, and P. Hinterdorfer. 2003. A molecular switch between alternative conformational states in the complex of Ran and importin beta 1. Nat Struct Biol 10:553–557.
170. Nevo, R., V. Brumfeld, M. Elbaum, P. Hinterdorfer, and Z. Reich. 2004. Direct discrimination between models of protein activation by single-molecule force measurements. Biophys J 87:2630–2634.
171. Ritco-Vonsovici, M., A. Ababou, and M. Horton. 2007. Molecular plasticity of beta-catenin: new insights from single-molecule measurements and MD simulation. Protein Sci 16:1984–1998.
172. Gilbert, Y., M. Deghorain, L. Wang, B. Xu, P. D. Pollheimer, H. J. Gruber, J. Errington, B. Hallet, X. Haulot, C. Verbelen, P. Hols, and Y. F. Dufrene. 2007. Single-molecule force spectroscopy and imaging of the vancomycin/D-Ala-D-Ala interaction. Nano Lett 7:796–801.

16

Probing Single Membrane Proteins by Atomic Force Microscopy

S. Scheuring, K. Tanuj Sapra, and Daniel J. Müller

Abstract In this book chapter, we describe the working principle of the atomic force microscope (AFM), followed by the applications of AFM in high-resolution imaging and single-molecule force spectroscopy of membrane proteins. In the imaging mode, AFM allows observing the assembly of membrane proteins directly in native membranes approaching a resolution of ~0.5 nm with an outstanding signal-to-noise ratio. Conformational deviations of individual membrane proteins can be observed and their functional states directly imaged. Time-lapse AFM can image membrane proteins at work. In conjunction with high-resolution imaging, the use of the AFM as a single-molecule force spectroscope (SMFS) has gained tremendous importance in recent years. This combination allows to locate the inter- and intramolecular interactions of single membrane proteins. SMFS allows characterization of interactions that guide the folding of proteins and describe the parameters that lead to their destabilization, malfunction and misfolding. Moreover, it enables to measure the interactions established by ligand- and inhibitor-binding and in membrane protein assemblies. Because of its practical use in characterizing various parameters of membrane proteins in their native environment, AFM can be aptly described as a 'lab on a tip' device.

16.1. Introduction

16.1.1. A Short Synopsis on Membrane Proteins

One of the crucial events leading to the formation of the first cell must have been the development of the membrane (Alberts et al., 2002). From the pool of various saturated and unsaturated lipids with different head groups, nature has very beautifully constructed the thin, self-sealing, insulating boundaries that are used by cells to create various specialized compartments. The Singer and Nicolson fluid mosaic model (Singer and Nicolson, 1972) considered the cell membrane as a dynamic, fluid structure where most molecules are able to diffuse freely about in the plane of the membrane, and not as a static structure as thought previously. Since the presentation of this model, our understanding of the cell membrane has advanced tremendously, to the point where we now know that the cell membrane is more mosaic than fluid (Dupuy and Engelman, 2008; Engelman, 2005).

S. Scheuring, K. T. Sapara, and D. J. Müller • Institut Curie, UMR168-CNRS, 26 Rue d'Ulm, 75248 Paris, France, and Biotechnology Center, Technische Universität Dresden, 01307 Dresden, Germany

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_16,

Refinements in the cell barrier during the course of evolution must have led to the insertion of proteins spanning the whole thickness of the hydrophobic membrane core for energy-storage purposes and to the creation of ways to control information transfer across this boundary. We can never hope to understand how cells work if we ignore the machineries and structure of the cellular membranes, because these connect the cells to their environment. Moreover, improving our understanding of membrane proteins would have a significant impact on medicine. The pharmaceutical importance of membrane proteins stems from the fact that they include structural proteins, channels, and receptors that are accessible through the exterior of cells and thus are formidable drug targets. Mutations in genes encoding membrane proteins are the causative agents of various diseases, including cystic fibrosis, retinitis pigmentosa, congenital nephrogenic diabetes insipidus, arrhythmias, hearing loss, and amyloid diseases (Agre and Kozono, 2003; Sanders and Myers, 2004). A better understanding of these defects is important for designing new therapies. For example, guanine nucleotide protein–coupled receptors (GPCRs) form the largest known family of cell-surface receptors, responding to diverse stimuli such as hormones, neurotransmitters, odorant molecules, and light, thus implicating them in virtually all physiological processes and making them targets for $\sim$60% of the new drug molecules. Ion channels, a diverse family of membrane-spanning proteins, regulate the ionic flux in virtually all cells by permitting the passive flow of ions down their electrochemical gradients. This process establishes the basis for membrane excitability and neurotransmission essential for the proper functioning of neurons and cardiac and muscle cells. The diverse physiological importance of ion channels makes them appropriate molecular drug targets as well (Overington et al., 2006). Thus, it is necessary to obtain structural, biochemical and biophysical information to improve our understanding of these fascinating machineries of the membrane and increase the efficiency of drug discovery.

The study of membrane proteins remains an important challenge for the structural biologist. The molecular-level understanding of membrane proteins lags far behind that of water-soluble proteins owing to experimental difficulties in obtaining high-resolution structural information due to their strict confinements in the lipid bilayer. Structure determination of membrane proteins is an arduous task, as emphasized by the fact that <1% of the structures in the Protein Data Bank (Berman et al., 2000) are membrane proteins. The situation has improved markedly in recent years, and we now know more than 168 unique structures from about 30 families (White, 2008).

Membrane proteins are hard to handle, and consequently the purification of functional protein in milligram quantities such as needed for structural determination using X-ray crystallography is a major problem. A common problem that precludes the characterization of many membrane proteins is their metastable nature once outside the membrane and their tendency to aggregate in solution, which leads to rapid inactivation (Booth, 2003; Bowie, 2001). The complex nature of eukaryotic membrane proteins makes their expression, solubilization, purification, and crystallization even more difficult. Therefore, it is conceivable that many more years will elapse before high-resolution structures of eukaryotic transmembrane proteins emerge.

However, understanding the function of a membrane protein requires much more than structural knowledge. Over the years we have learned that membrane proteins dynamically assemble into supramolecular structures to fulfill the functional tasks of the cell (Engelman, 2005; Rajendran and Simons, 2005). Within such assemblies membrane proteins and lipids interact with each other, which in turn modulates the functional state of the membrane protein. We are beginning to learn that the functional state of, for example, GPCRs is modulated by homo- and heterodimerization and that the dimerization and functional tuning are controlled

by the cell (Fotiadis et al., 2006; Park and Palczewski, 2005). Many answers in science beget curiosity and further questions. What do membrane protein assemblies look like? How does a cell control these assemblies? What are the mechanisms that guide the assemblies and finally fine-tune the functional state of such important cellular machineries? Unraveling the answers to these and related questions does not necessarily require an immediate structural knowledge. Integrated approaches that combine biochemical, biophysical, and computational analyses with low-resolution structures are likely to have increasingly important roles in providing frameworks for the mechanistic understanding of membrane protein structure and function.

16.1.2. Atomic Force Microscope as a Multifunctional Tool for Characterizing Membrane Protein Structure and Function

Soon after its introduction in 1986, the atomic force microscope (AFM) (Binnig et al., 1986) was applied to study membrane proteins (Butt et al., 1990). Whereas these first experiments were conducted in a dehydrated environment, the sample preparation and AFM imaging procedures have been continuously developed toward imaging membrane proteins in conditions simulating the *in vivo* environment. The AFM provides novel ways to characterize structure–function relationships of live cells and cellular compartments down to the resolution of single molecules (Drake et al., 1989; Engel and Müller, 2000; Muller and Dufrene, 2008; Muller et al., 2006; Scheuring, 2006; Scheuring et al., 2005). The great advantage of AFM is that the biological sample can be investigated in a buffer solution at ambient temperature and pressure. Another key advantage is that AFM imaging bypasses the labeling procedures required for studying biological specimens in fluorescence microscopy. Membrane proteins exemplify the versatility of AFM in providing structural and functional insights into the different oligomeric and conformational states in their native environment that are difficult to obtain with conventional approaches. The AFM allows imaging nonordered membrane proteins directly in native membranes, a unique feature compared to the other techniques applied in structural biology (Buzhynskyy et al., 2007b; Scheuring and Sturgis, 2005). Although the AFM images only surfaces, having a lateral resolution of $\sim$0.5 nm and a vertical resolution of $\sim$0.1 nm with an outstanding signal-to-noise ratio enables the observation of single membrane proteins in great details. Conformational deviations of individual membrane proteins can be observed and their functional states directly imaged. Moreover, recently time lapse AFM has been used to monitor the spatial and temporal dynamics of bacterial surface-layer proteins (Müller et al., 1996), nuclear pore complexes (Stoffler et al., 1999), outer membrane porins (Müller and Engel, 1999), gap junction hemichannels (Müller et al., 2002; Yu et al., 2007), cytolysin pores (Czajkowsky et al., 2004), and potassium channel KirBac3.1 (Jaroslawski et al., 2007a) thus revealing functionally related conformational changes.

In recent years the AFM has emerged as an indispensable nanotechnological tool that makes it possible to image biological specimens and at the same time to specifically probe biological, chemical, and physical information (Gerber and Lang, 2006; Muller and Dufrene, 2008). The AFM is readily used to sense and locate biomolecular inter- and intramolecular interactions of tissues, living cells, and single molecules. In conjunction with high-resolution imaging, the use of AFM as a single-molecule force spectroscope (SMFS) has gained tremendous importance in recent years (Borgia et al., 2008; Engel and Gaub, 2008; Kedrov et al., 2007). This combination allows the characterization of interactions that stabilize functional proteins, as well as those that destabilize proteins, leading to malfunction and misfolding (Janovjak et al., 2006). Because of its practical use in characterizing various parameters of

membrane proteins in their native environment, AFM can be aptly described as a "lab on a tip" device (Müller et al., 2006).

In the following sections of this review we describe the working principle of the AFM, followed by the applications of AFM in high-resolution imaging and single-molecule mechanical unfolding and folding of membrane proteins. We end by describing the future prospects of AFM-based methodologies.

16.2. The Atomic Force Microscope

The early 1980s saw the advent of scanning probe microscopy (SPM) when Binnig and Rohrer invented the scanning tunneling microscope (STM) (Binnig et al., 1982). Albeit developed with the sole interest of studying the electrical properties of conducting surfaces, it was soon discovered that the use of STM is not restricted to imaging but can also be extended for manipulating single atoms. In 1986, the year Binnig and Rohrer won the Nobel Prize in Physics, Binnig, Quate, and Gerber demonstrated the use of short-range interactions between a scanning probe and an object to contour nonconducting and conducting surfaces at atomic (or close to atomic) resolution (Binnig et al., 1986). This so-called scanning force microscope (SFM) or atomic force microscope was the second member of the SPM family. The principle of scanning an atomically sharp probe over an object and detecting specific interactions initiated the development of new SPM techniques that could be applied to characterize multiple physical, chemical, and biochemical parameters of the object. AFM and all the following SPM methodologies opened the floodgates to the nanoworld (Gerber and Lang, 2006).

16.2.1. Principle and Setup

The AFM works on a very simple principle, making it probably one of the easiest forms of microscopy to understand. Figure 16.1 illustrates the main parts of an AFM setup. A central component of the equipment is a micromachined (microfabricated or nanofabricated) probe mounted at the end of a cantilever. A voltage-driven piezoelectric transducer/actuator enables spatial positioning of the cantilever probe in the *x, y*, and *z* directions at subnanometer precision. The fluid cell with the cantilever is mounted inside the so-called head of the AFM. The AFM head houses the optical deflection system, consisting of a laser diode and a quadrant photodetector. AFM images are created by scanning a sharp AFM probe in the *x*-*y* plane over a sample surface. The interaction force between the probe and the sample is monitored by measuring the deflection (vertical bending) of the cantilever, which is usually detected by the laser beam focused on the upper part of the free end of the cantilever and reflected onto a quadrant photodiode. Recording of the force at every pixel scanned in the *x*-*y* plane is then used to reconstruct the surface topography. Of the many different AFM imaging modes, the constant-force mode and the oscillation mode represent the two most commonly used modes for biological applications. In the constant-force mode, the force detected between the probe and the object is held constant by a feedback loop that moves the piezo element in the *z* direction to keep the cantilever deflection, that is, the force, constant. The resulting AFM image detects the height change for each pixel scanned and thus reflects the sample topography. Oscillation mode AFM oscillates the AFM probe sinusoidally. A feedback loop ensures that the sample surface is touched only at the very end of every downward movement of the AFM probe. This interaction changes the amplitude of the cantilever, and the feedback loop keeps this change constant by controlling the *z* direction of the piezo. The resulting AFM

image corresponds to a topograph. However, many factors can influence the amplitude of the oscillating probe, including surface charge (roughness), elasticity (material contrast), and hydrophobicity. Thus, the factors causing the apparent topographic contrast are sometimes difficult to interpret.

Whereas during AFM imaging the piezoelectric transducer or the cantilever is raster scanned or oscillated along the object (Figure 16.1), for SMFS measurements the transducer is moved only in the z (up–down) direction (Figure 16.2), thus enabling the measurement of interaction forces over the probe-sample distance. The pattern of these force–distance (F-D) curves is characteristic of the specific interactions occurring between the probe and the sample. Biochemical, chemical, or physical modification of the probe and/or support allows the specific detection of inter- and intramolecular interactions of interest (Gerber and Lang, 2006; Muller and Dufrene, 2008).

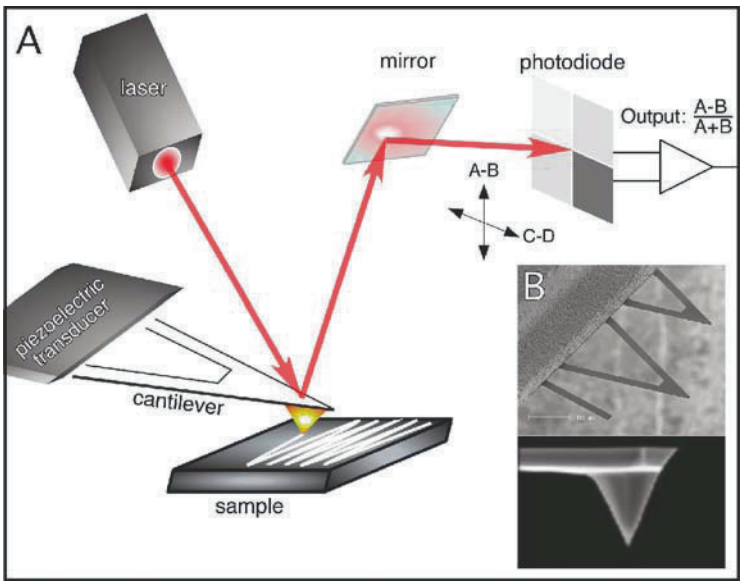


Figure 16.1. Setup of an atomic force microscope (AFM) instrument. **A.** The scheme shows the different parts and working of an AFM. Piezoelectric transducer, laser, mirror, and photodiode are housed in the "head" of the microscope. The cantilever **(B)** is fixed in a fluid cell (not shown) and placed inside the head, which is then placed over the sample. A laser beam (emitting in the red or infrared range) is adjusted on the edge of the cantilever so as to obtain a maximum sum signal on the photodiode. During raster scanning of the probe over the sample the beam is deflected with the up–down movement of the cantilever. The voltage signal is converted to force (piconewtons) or deflection (nanometers) with the help of a software. **B.** V- and bar-shaped cantilevers with pyramidal tips are used, depending on the experimental task and the sample studied. (Cantilever images from http://www.veecoprobes.com.)

16.3. High-Resolution Imaging of Single Native Proteins

The AFM makes it possible to image water-soluble and membrane proteins at a lateral resolution of 1 nm or better. The vertical resolution of high-resolution AFM topographs can approach the Angström range. Much effort has been made to demonstrate that the high-resolution AFM images represent the true structural features of the native unperturbed

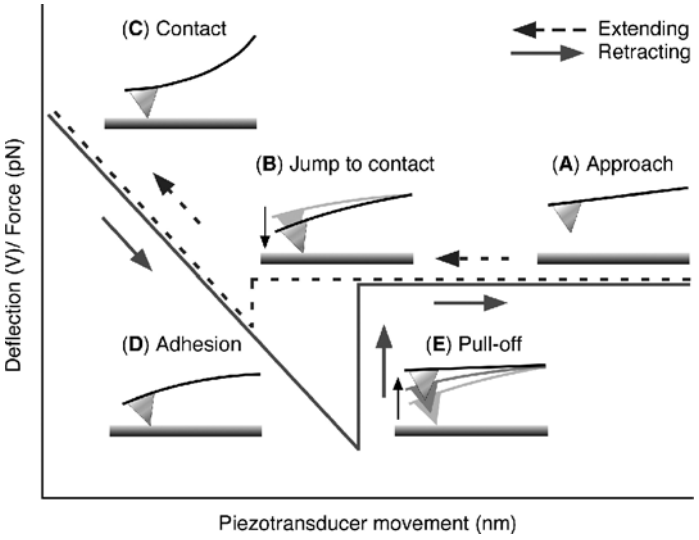


Figure 16.2. A typical force–distance (F-D) curve. Schematic representation of the F-D curve with the corresponding stages in probe–sample interaction during the up–down movement of the piezoelectric transducer. The slope of the region between B and D is used to determine the deflection sensitivity of the cantilever (volts/nanometer), which in conjunction with the spring constant value of the cantilever is required to convert a voltage–displacement curve into a force–displacement curve. It should be noted, however, that the F-D curve (*solid line*) shown here does not represent the unfolding of a polypeptide chain but instead denotes an adhesion peak between the cantilever probe and the surface.

protein. However, sample preparation and imaging conditions must be precisely followed to prevent possible structural and functional distortion of the fragile biological sample (Müller et al., 1997; Muller and Engel, 2007; Müller et al., 1999). Whereas oscillation mode imaging is in most cases appropriate to obtain high-resolution AFM topographs of water-soluble proteins and DNA (Hansma and Laney, 1996; Hansma et al., 1994), contact mode imaging is used to reveal AFM topographs of membrane proteins down to the subnanometer resolution (Möller et al., 1999; Scheuring et al., 1999). For several reasons, such as increased sampling time for each pixel probed, contact mode topographs of membrane proteins can approach a higher resolution than oscillation mode topographs. Most important, however, is that membrane proteins are particularly well-suited for high resolution imaging because they are embedded in a lipid bilayer that orients and stabilizes them. Figures 16.3 and 16.4 show two examples of high-resolution topographs of water-soluble proteins and membrane proteins in buffer solution, respectively. The resolution achieved clearly allows one to observe characteristic substructures of the proteins. The figures show only a selection of a plethora of proteins and nucleic acids that have been imaged by AFM. For an overview see reviews by Müller et al. (2006) and Scheuring et al. (2001, 2007c).

16.3.1. High-Resolution Imaging of Protein Assemblies

AFM-based high-resolution single-molecule imaging shows promising potential for imaging individual membrane proteins for structural studies and obtaining information on the dynamic processes of complex formation (Scheuring, 2006) and supramolecular

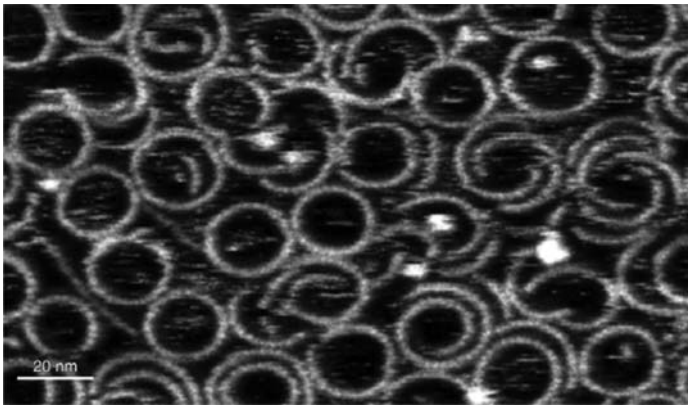


Figure 16.3. High-resolution atomic force microscope (AFM) topographs of tubulin protofilaments. Shortly after preparation (<0.5 hr), microtubules depolymerized into protofilaments that consist of α- and β-tubulins. The topograph was recorded using oscillation-mode AFM in buffer solution (Elie-Caille et al., 2007).

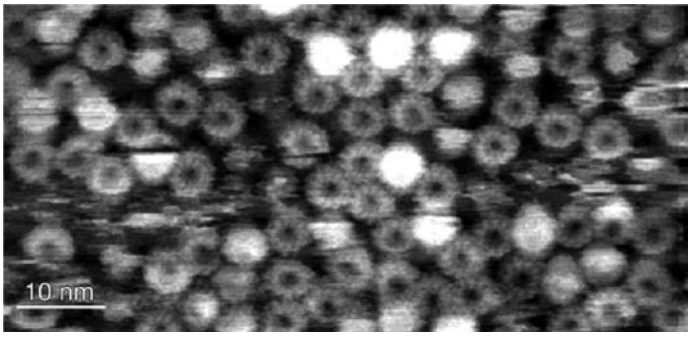


Figure 16.4. High-resolution atomic force microscope (AFM) topographs of sodium-driven rotors from the F_0F_1 ATP synthase from *Ilyobacter tartaricus*. The topograph reveals the densely packed reconstitution of membrane-embedded rotors. The exceptional signal-to-noise ratio allows one to structurally analyze individual rotors. Intact rotors are composed of 11 subunits. The constant-force mode AFM topograph was recorded in buffer solution (Stahlberg et al., 2001).

assembly (Fotiadis et al., 2003; Scheuring et al., 2004b). Most important, the assemblies of membrane proteins can be characterized in their native membranes. Thus, it becomes possible to monitor the assembly of different membrane proteins into functional units. The AFM is the only technique that can provide such data on nonordered membrane systems. This is due to its striking signal-to-noise ratio, which allows the structural identification of individual proteins without the need of labeling. This approach has provided clear images of membrane protein assemblies in native chromatophore preparations from *Rhodospirillum photometricum* (Scheuring et al., 2004a, b; Scheuring and Sturgis, 2005). These studies yielded striking and novel findings concerning antenna heterogeneity, antenna domain formation, and complex assembly. The majority of light-harvesting complexes 2 (LH2) assembled as nonameric rings. However, on closer examination LH2 complexes of various sizes were found within the native membranes. Diameter distribution and image processing analysis showed heterogeneity of the

LH2 complex stoichiometry around the general nonameric assembly (∼70%) with smaller octamers (∼15%) and larger decamers (∼15%). This finding was qualitatively corroborated by examination of individual complexes in raw data images (Scheuring et al., 2004a). It seems probable that a heterogeneous stoichiometry is an inherent feature of LH2, as it has also been observed in *Phaeospirillum molischianum* (Gonçalves et al., 2005) and *Rhodopseudomonas palustris* (Scheuring et al., 2006). In contrast to the heterogeneity found for LH2 complexes, the reaction center–light-harvesting complex 1 (RC-LH1) core complexes appeared uniform in size; each monomeric reaction center (RC) was surrounded by a closed elliptical assembly of 16 LH1 subunits, with long and short axes of 95 and 85 Å, respectively, following the long RC axis. Analysis of the distribution of the photosynthetic complexes of *R. photometricum* showed significant clustering of both antenna complexes and core complexes. Membranes were found that contained domains densely packed with photosynthetic proteins, in addition to other regions composed of protein-free lipid bilayers. Clustering of complexes is a functional necessity because each light-harvesting component must pass its harvested energy to a neighboring complex and eventually to a RC (Scheuring et al., 2004a). There is no fixed structural assembly of LH2 and core complexes. Core complexes completely surrounded by LH2 (Figure 16.5A; 2) and core complexes making multiple core–core contacts (Figure 16.5A; 1) were found (Scheuring et al., 2004a; Scheuring and Sturgis, 2005). These high-resolution topographs (Figure 16.5A) were used to compute reliable atomic models of the supramolecular structures (Figure 16.5B) (Scheuring et al., 2007a). Detailed pair correlation function analysis of the complex assembly showed that the most frequent assembly was two core complexes separated by an intercalated LH2 (Scheuring and Sturgis, 2005).

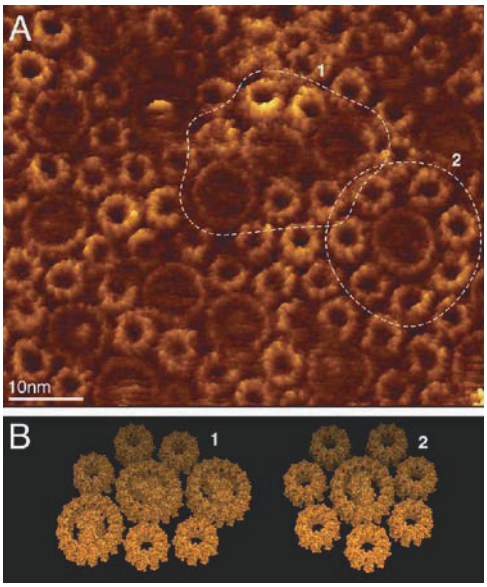


Figure 16.5. High-resolution atomic force microscope (AFM) topograph of a high-light-adapted membrane of *Rhodospirillum photometricum.* **A.** The mixture of light-harvesting complexes 2 (LH2), small nonameric rings with ∼50 Å diameter, and core complexes—elliptical reaction center–light-harvesting complex 1 (RC-LH1) assemblies ∼10 nm in diameter housing a central RC—are visualized in a non-ordered dense packing. Some core complexes are completely surrounded by LH2 rings (2), whereas others make core–core contacts (1). [Adapted from Scheuring and Sturgis (2005).] **B.** The high-resolution AFM topographs can be docked by the atomic structures, providing structural models of protein ensemble machines. [Adapted from Scheuring et al. (2007a).]

Adaptation of membrane architecture in response to changing environmental conditions is a general process in living cells. In photosynthetic organisms membranes must adapt to changing photon intensity. The adaptation of the supramolecular assembly of the photosynthetic complexes in *R. photometricum* membranes can be observed through a comparative study of membranes from cells grown under different light intensities. In membranes from low-light adapted cells increased quantities of peripheral LH2 for better capture of the less abundant photons were found. Additional LH2 complexes were not randomly inserted into the membrane but formed paracrystalline, hexagonally packed antenna domains (Figure 16.6A). Core complexes remained in domains in which they were locally more highly concentrated (LH2 rings/core complex • 3.5), when compared with the average density under low-light cell growth (LH2 rings/core complex • 7). Indeed, these domains in the low-light-adapted membranes resembled the high-light-adapted membranes in terms of protein composition and complex distribution (Scheuring and Sturgis, 2005). This indicated that complex assembly followed a eutectic phase behavior with an ideal LH2 rings/core complex ratio of ~3.5 independent of the growth conditions, and additional LH2 being synthesized under low-light conditions was integrated in specialized antenna domains (Figure 16.6A). The LH2 packing (Figure 16.6B) in antenna domains was observed to be extremely dense (Scheuring et al., 2007a) and rigid (Scheuring and Sturgis, 2006), probably excluding quinone/quinol diffusion. RCs that are grouped together, independent of the growth conditions, and formation of antenna domains under low-light conditions prevent photodamage under high-light conditions and ensure efficient photon capture under low-light conditions (Scheuring and Sturgis, 2005).

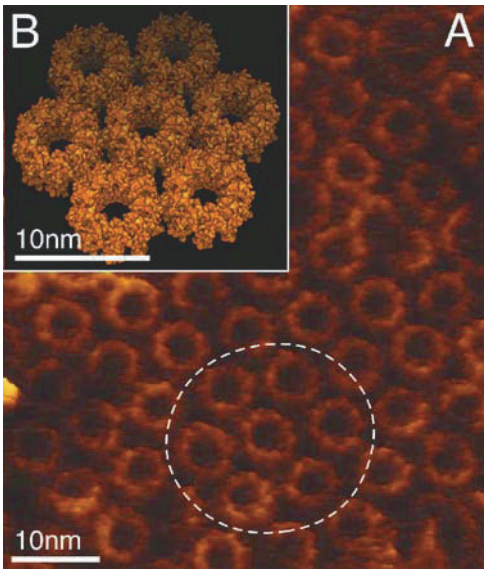


Figure 16.6 High-resolution atomic force microscope (AFM) topograph of antenna domains in a low-light-adapted membrane of *Rhodospirillum photometricum.* **A.** The raw data AFM topograph shows light-harvesting complexes 2 arranged in a paracrystalline hexagonal packing. This packing is the densest possible assembly of these complexes ensuring efficient excitation energy transfer. [Adapted from Scheuring and Sturgis (2005).] **B.** The structural model of the supramolecular antenna assembly used to compute the intercomplex pigment distances. [Adapted from Scheuring et al. (2007a).]

Moving on from bacterial systems, a challenging task is the study of supramolecular membrane protein assemblies in mammalian membranes. Only a few examples of single-component specialized membranes have been reported. A breakthrough study documented the assembly of bovine rhodopsin in native rod outer segment disk membranes (Fotiadis et al., 2003). Recently, the assembly of adenosine triphosphate (ATP) synthases in inner mitochondrial membrane (Buzhynskyy et al., 2007c) and of the voltage-gated anion channel (VDAC) in outer (Gonçalves et al., 2007) mitochondrial membrane was studied. Only one high-resolution AFM study of a multicomponent mammalian membrane has been reported—the specialized lens core cell membranes (Buzhynskyy et al., 2007b). To ensure its transparency and to change its shape for focusing at different distances (accommodation) the eyes lens has developed some remarkable adaptations (Donaldson et al., 2001). The avascular nature of the lens minimizes light scattering. Furthermore, the lens cells have degraded organelles during differentiation and contain a high concentration of crystalline proteins (Andley, 2007). The cells are tightly packed to make intercellular distances smaller than the wavelength of visible light. Lacking organelles, notably mitochondria, the lens cells are very restricted in their possibility of performing metabolic biochemical processes. To explain how such cells located deeply inside the lens are nourished, an internal microcirculation model was proposed. In this microcirculation system, solutes flow in the extracellular space, enter deeper-lying lens cells, and flow through cell-to-cell channels back toward the lens periphery (Donaldson et al., 2001) thus maintaining a flow of water, ions and metabolites.

Two types of membrane proteins—aquaporin-0 (AQP0) (Alcala et al., 1975) and connexins (Cx) (Fleschner and Cenedella, 1991)—form cell-to-cell junctions ensuring metabolite transport, waste evacuation, water homeostasis, and adhesion between fiber cells. Together, AQP0 and connexons form junctional microdomains—functional platforms in the cell membranes that warrant water, ion, and metabolite transport and facilitate cell adhesion (Figure 16.7A). In these junctional microdoamins AQP0 forms extended square arrays that are separated and confined by connexons (Figure 16.7B). Very rarely, individual connexon molecules integrated in AQP0 arrays are found. The edge regions can either be terminated by single rows of connexon channels or by connexons that are densely packed (Figure 16.7C). In the densely packed regions, individual AQP0 tetramers were identified to mix with connexons (Buzhynskyy et al., 2007b). Statistical analysis showed that ∼70% of the AQP0 array circumference is lined by connexons that defined and confined the junctional microdomain shape and size (Mangenot et al., 2008).

16.3.2. High-Resolution Imaging of Membrane Protein Diffusion

Time-lapse high-resolution AFM allows taking sequential snapshots of biological systems. Applied to membrane proteins, it was possible to follow how single membrane-spanning rotors of the F_0F_1 ATP synthase diffused in the plane of the lipid bilayer (Müller et al., 2003). Similarly, the movements of AQP0 in the lens membrane could be analyzed. Individual water channel tetramers showed a clear tendency to adjust their position and rotational alignment with respect to the optimal assembly of AQP0 into microdomain arrays (Buzhynskyy et al., 2007b; Scheuring et al., 2007b) (Figure 16.8). The analysis of the single trajectories of diffusing membrane proteins made it possible to determine their diffusion behavior in a supported lipid bilayer. These examples demonstrate that high-resolution AFM imaging is promising not only for imaging membrane proteins for structural studies, but also for obtaining information on the dynamic processes of their assembly and complex formation.

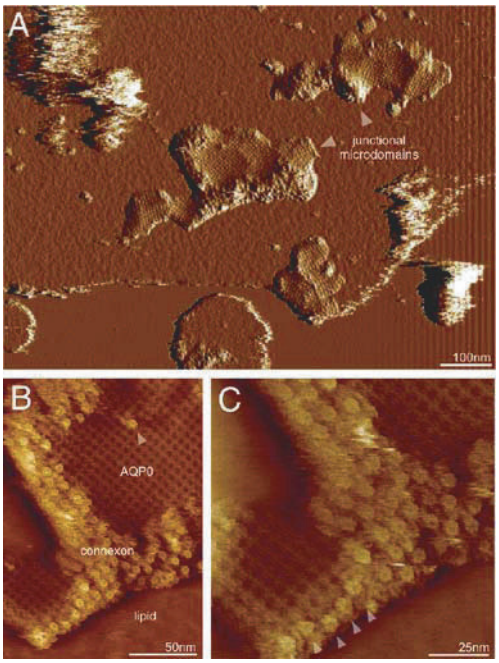


Figure 16.7. High-resolution atomic force microscope (AFM) analysis of junctional microdomains in native lens core membranes. **A.** Junctional microdomains represent ensemble functional platforms in the cell membranes comprising of lens-specific water channel AQP0 and gap-junction connexons. **B.** The junctional microdomains consist of extended AQP0 arrays edged by connexons. **C.** Connexons confine the junctional microdomains, which are either densely packed or exist as single rows (*arrowheads*).

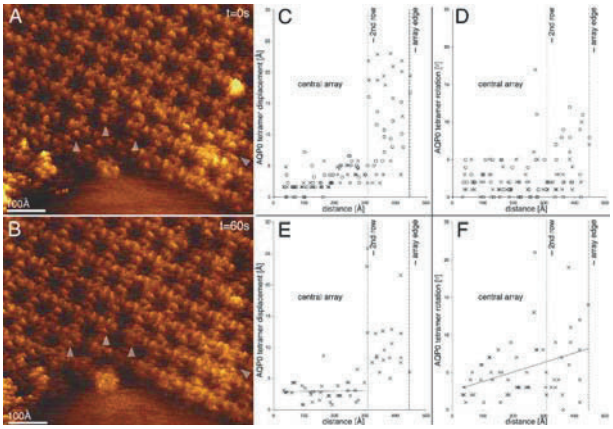


Figure 16.8. Time-lapse high-resolution atomic force microscope (AFM) topographs reveal the movements of AQP0 in native lens membrane junctional microdomains. **A, B.** Consecutive high-resolution topographs. Arrowheads indicate the AQP0 tetramers that undergo the most significant movements. Displacement **(C)** and rotational misalignment **(D)** of the AQP0 tetramers shown in panel A (*crosses*) and panel B (*circles*) with respect to a perfect AQP0 lattice. Movement **(E)** and rotation **(F)** of AQP0 tetramers when comparing images shown in panels A and B. Whereas tetramers in the center of the AQP0 array are essentially immobile, tetramers in the last two rows show significant freedom to move. Rotational movements of the tetramers also increase with proximity to the edge of the array.

16.3.3. High-Resolution Imaging of Proteins at Work

Time-lapse high-resolution AFM imaging of biological macromolecules in their native environment has the potential to allow their direct observation at work. Studies include bacterial surface layers (Müller et al., 1996), bacterial outer membrane pores (Müller and Engel, 1999), inwardly rectifying potassium channels (Jaroslawski et al., 2007b), gap junction communication channels (Müller et al., 2002; Yu et al., 2007), and RNA polymerases (Guthold et al., 1999). Commercial AFMs require $\geq$60 sec to record a high-resolution topograph. In contrast, many biological processes take place on a much faster time scale. To circumvent this limitation, high-speed AFMs have been developed that might soon be commercially available (Ando et al., 2003; Humphris et al., 2005; Viani et al., 1999). Depending on the approach, these high-speed AFMs can record up to 200 topographs in 1 sec. Such temporal resolution combined with the high-resolution imaging capabilities of the AFM will open new avenues for characterizing dynamic processes of biological systems. Breakthrough time-lapse imaging studies using high-speed AFM setups reported protein interactions between molecular chaperones GroEL/GroES (Viani et al., 2000) and the movement of motor proteins in real time (Ando et al., 2007). Another molecular phenomenon observed using high-speed AFM is protein diffusion, which was impressively shown by imaging individual loosely distributed streptavidin molecules diffusing on a supported bilayer (Figure 16.9A) (Ando et al., 2008) and by following the displacements of single molecule defects in two-dimensional lattices of streptavidin (Figure 16.9B) (Yamamoto et al., 2008).

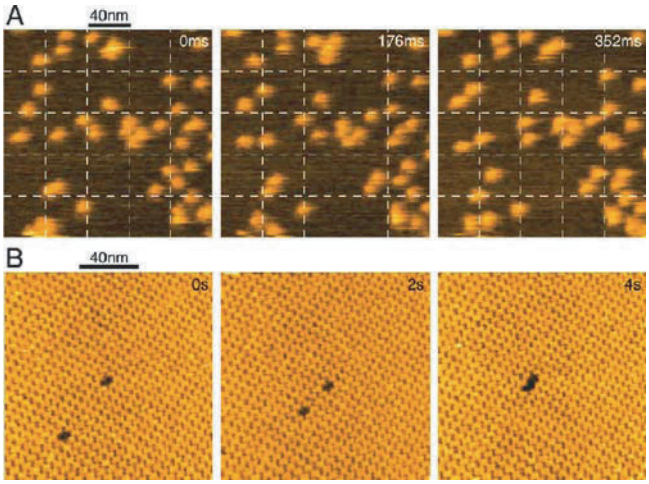


Figure 16.9. Protein diffusion visualized by high-speed atomic force microscope (AFM). **A.** Three consecutive images (time range 352 ms) show diffusion of streptavidin molecules on biotin-containing lipid bilayers (DPPC-DOPE-biotin). [Adapted from Ando et al. (2008).] **B.** Three consecutive images (time range 3 sec), showing the fusion of two monovacancy point defects in a streptavidin two-dimensional crystal. [Adapted from Yamamoto et al. (2008).]

16.4. Single-Molecule Force Spectroscopy of Membrane Proteins

The confines of the AFM are not limited to imaging. It is possible to detect interactions at the single-molecule level by using a technique called single-molecule force spectroscopy (SMFS). As the name implies, force is applied to a molecular system of interest attached

between a cantilever probe and a well-chosen surface (Figure 16.10). Other single-molecule tools such as optical or magnetic tweezers, where a dielectric or a magnetic bead is the probe, can also be used to conduct SMFS experiments (Neuman and Nagy, 2008). In the following we focus on SMFS with AFM. Unlike the tweezers techniques, AFM-based SMFS is the only approach that can be successfully applied for simultaneously imaging protein motions and detecting the interactions shaping the dynamic energy landscapes of membrane proteins. The SMFS approach relies on a simple working principle: A biomolecule tethered to the AFM probe is withdrawn from the surface, which might specifically or nonspecifically anchor the biomolecule as well. As soon as the mechanical pulling force supersedes the stability of the interactions that stabilize the biomolecule, the interactions are broken. The ensuing F-D curve reveals the mechanical response of the biomolecule under an external force (Figures 16.2 and 16.11B). In the case of a receptor–ligand pair this approach allows the measurement of binding strength, in the case of a DNA molecule the F-D curve reveals the stretching behavior of the nucleic acid as the force is increased, and in the case of an oligomeric protein construct the F-D records the forces required to unfold individual proteins. Owing to the nature of membrane protein unfolding the F-D signals can show a very unique response to the applied force, as explained in the next sections.

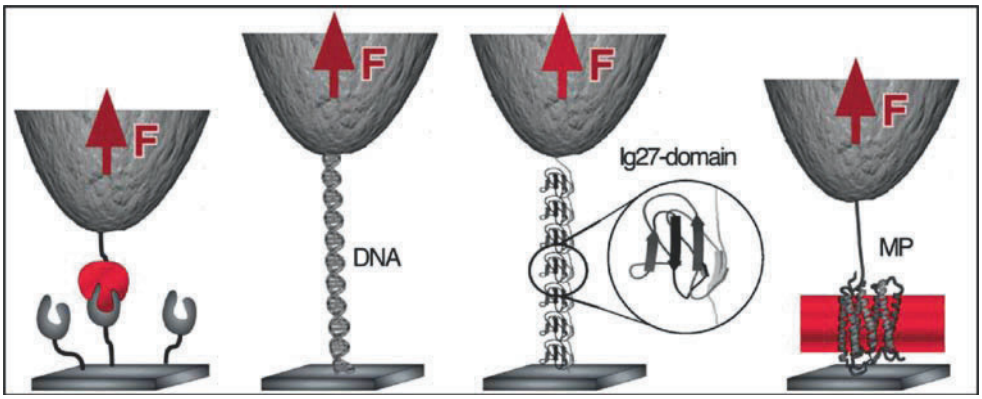


Figure 16.10. Detecting biomolecular interactions using the molecular probe of an atomic force microscope (AFM). Chemical modification of the AFM probe allows detection of various specific and nonspecific biological interactions. **Left to right**, examples of detecting receptor–ligand interactions, mechanical properties of DNA, interactions stabilizing Ig27 domains, and interactions of membrane proteins.

16.4.1. Detecting Unfolding Intermediates and Pathways

SMFS experiments on membrane proteins are performed by mechanically pulling on a terminal end of the protein tethered to the AFM probe while separating the probe from the membrane surface at a certain speed (Figure 16.11). In contrast to water-soluble proteins, which merely unfold in one spontaneous event, it is a general observation that membrane proteins unfold in a stepwise manner, each step representing an unfolding intermediate (Figure 16.11) (Kedrov et al., 2007). The sequence of unfolding intermediates constitutes the unfolding pathway of the membrane protein (Figure 16.12). With the possibility to investigate membrane proteins in the native membrane and at physiological conditions, the SMFS approach creates a range of possibilities to characterize the factors that influence the mechanical unfolding pathways of membrane proteins. It can be shown that several unfolding pathways can coexist and that the probability of a membrane protein to choose certain unfolding

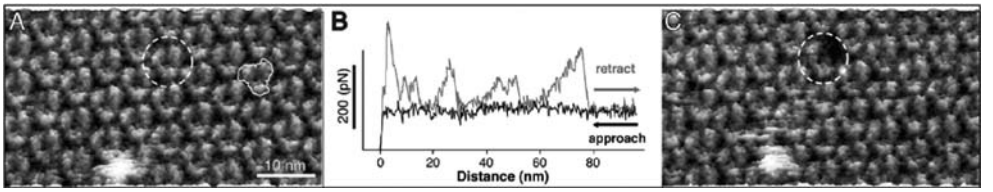


Figure 16.11. Controlled unfolding of one bacteriorhodopsin from native purple membrane. **A.** Atomic force microscope (AFM) topograph of the cytoplasmic purple membrane surface. Bacteriorhodopsins forming trimers (*outlined*) are clearly visible. To select an individual protein we zoomed in (*circled area*) by reducing the pixel number and frame size. After positioning, the AFM probe was kept in contact with the protein surface for ~1 sec. Most of the extension curves recorded on the cytoplasmic surface showed no adhesion peaks. However, 5% of the force–distance curves showed discontinuous adhesion forces similar to that shown in panel B. The length of the force spectrum extended up to 75 nm, corresponding to the length of an entirely unfolded bacteriorhodopsin (248 amino acids). **C.** Topograph recorded after the force spectrum showed one bacteriorhodopsin missing (*circle*). Experiments were performed in buffer solution at room temperature.

pathways depends on, for example, buffer solution, electrolyte, pH, temperature, oligomeric assembly, mutation, drug binding, and specific protein interactions. In the following we provide a short overview of why it is important to characterize membrane protein folding and unfolding and the advantages of doing so at single-molecule resolution.

16.4.2. Importance of Studying Membrane Protein (Un)Folding

Besides their pharmacological relevance, studying membrane protein folding *in vitro* (Booth and Curran, 1999) is of considerable interest at the basic level. Like the folding of water-soluble proteins, the folding of membrane proteins probably proceeds down a funnel-shaped energy landscape to an energy minimum (Dill and Chan, 1997). A major difference between soluble protein folding and membrane protein folding, however, is that the starting point in membrane proteins is much more constrained because the secondary structure and topology are set by the strict confines of the lipid membrane. Thus, the unfolded protein is much farther down the folding funnel and closer to the folded state than in soluble proteins. It seems clear that the folding-energy landscape is defined by a complex interplay among various forces, including polypeptide partitioning in the bilayer, interactions between the lipid chains and the protein, and interactions within the protein itself.

Protein folding studies are excellent starting points to help us understand the effects of point mutations in membrane protein misassembly and misfolding in the cell, which can be extremely complex processes involving insertion in the endoplasmic reticulum (ER), trafficking to the Golgi, and quality control of the ER (ER-associated degradation or ERAD), followed by degradation by proteasomes (Brodsky and McCracken, 1999; Tsai et al., 2002). The exact mechanisms by which single–amino acid changes lead to disease phenotypes are largely unknown. Many disease-linked mutations also occur in integral membrane proteins causing the protein to misfold or misassemble (Sanders and Myers, 2004; Sanders and Nagy, 2000) and leading to diseases such as cystic fibrosis, retinitis pigmentosa, Charcot-Marie-Tooth disease, and hereditary hearing loss (Aridor and Hannan, 2000, 2002; Cotton and Horaitis, 2002). Knowledge of membrane protein folding mechanisms *in vitro* will take us to the root of these problems and help us work toward strategies for disaggregating and folding proteins correctly. Their implication in varied functions and involvement in every cellular process and many diseased states has galvanized research on membrane proteins.

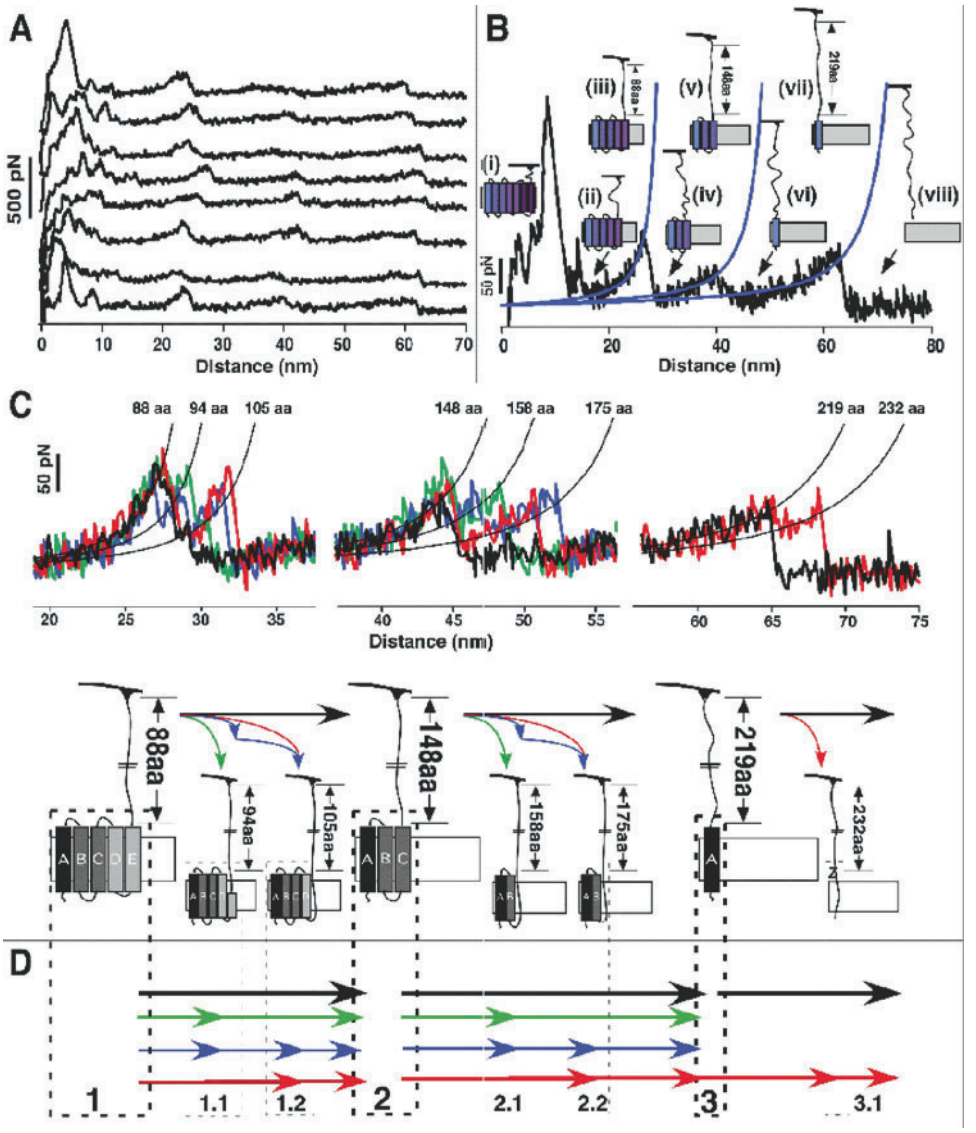


Figure 16.12. Assigning unfolding barriers, unfolding pathways, and unfolding probabilities of a membrane protein. **A.** Force–distance (F-D) curves, each one recorded on unfolding a single bacteriorhodopsin from the native purple membrane. **B.** The unfolding pathway described by each F-D spectrum. **C.** Whereas some unfolding barriers unfold individually, others have a certain probability to unfold together with other barriers. **D.** Individual secondary structure elements of the membrane protein can unfold in different pathways. In the simplest pathway, two helices and their connecting loop unfold in one single step (*thick line*, *black arrow*). In other pathways the helices can unfold in single steps (*colored arrows*). These unfolding steps reflect unfolding intermediates, and the sequence of the unfolding intermediates represents the unfolding pathway. The probability with which a bacteriorhodopsin selects an individual unfolding pathway depends on the environmental conditions, such as temperature, electrolyte, oligomeric assembly, or structural modifications (Janovjak et al., 2003; Müller et al., 2002c; Sapra et al., 2006a).

16.4.3. Why Study Membrane Protein (Un)Folding Under a Mechanical Load?

The use of sodium dodecyl sulfate (SDS) or lipid vesicles for studying membrane protein folding and unfolding (Bowie, 2001; Seddon et al., 2004) poses the problem that these do not mimic the *in vivo* folding environment (Booth, 2003; McKibbin et al., 2007). In addition, these bulk assays do not take into account the protein–protein interactions that crucially determine the folding and functional kinetics of a membrane protein (Engelman, 2005). Single-molecule AFM assays partially circumvent these problems. Importantly SMFS the experiments are always performed with the membrane protein being inserted in its native membrane, thus approaching physiological conditions. It was shown that AAA proteases mechanically extract membrane proteins from mitochondrial and ER membranes *in vivo* (Leonhard et al., 2000; Ye et al., 2001). Similar mechanisms have been previously proposed for *Escherichia coli* protease FtsH (Kihara et al., 1999). Although the final stages of helix insertion in the lipid bilayer involve lateral release of the helix (Sadlish et al., 2005; Xie et al., 2007), the vectorial nature of mechanical unfolding is a close approximation to the vectorial insertion of a helix facilitated by the translocons (Rapoport, 2007), Sec61 in the rough ER (Van den Berg et al., 2004), or the bacterial SecYEG complex machinery (Breyton et al., 2002). This highlights the fact that studying membrane protein unfolding under an applied force is not an artifactual process and may offer advantages beyond those of conventional methods.

16.4.4. Unfolding Forces Reflect Interactions That Stabilize Structural Regions

As discussed in the previous section, the force peaks of a F-D spectrum denote the unfolding intermediates of a membrane protein. These intermediates constitute stable mechanical barriers that a membrane protein establishes against an external force. As soon as the stability of a barrier is overcome by an externally applied pulling force, the structural segment unfold spontaneously and the unfolding process is resolved within the detection limit of the SMFS setup. Thus, the applied force required to initiate unfolding of a structural segment reflects the strength of stabilizing interactions in the segment. Figure 16.13 portrays a map of stable structural segments detected in different membrane proteins. These stable structural segments can unfold in single events or together with adjacent structural segments. The probability that structural segments unfold in individual or in grouped events depends on the experimental conditions, such as temperature or loading rate (Janovjak et al., 2003; Janovjak et al., 2004). In addition, it was shown that the origin of unfolding barriers established by structural segments depends on the pulling direction, that is if the membrane protein is being pulled from the N- or the C-terminal (Kedrov et al., 2004; Kessler and Gaub, 2006). Whereas some common unfolding barriers were detected when unfolding the membrane protein by pulling its C-terminal and N-terminal end, other barriers only showed up when pulling from either one of the two termini. Thus, mechanically pulling different termini forces a membrane protein along different unfolding pathways of its energy landscape.

16.4.5. Origin of Unfolding Forces

A change in the population of individual unfolding pathways of a membrane protein with changing experimental conditions indicates altered dynamics of the protein's inter- and intramolecular interactions. Accordingly, the unfolding forces, which are a direct measure

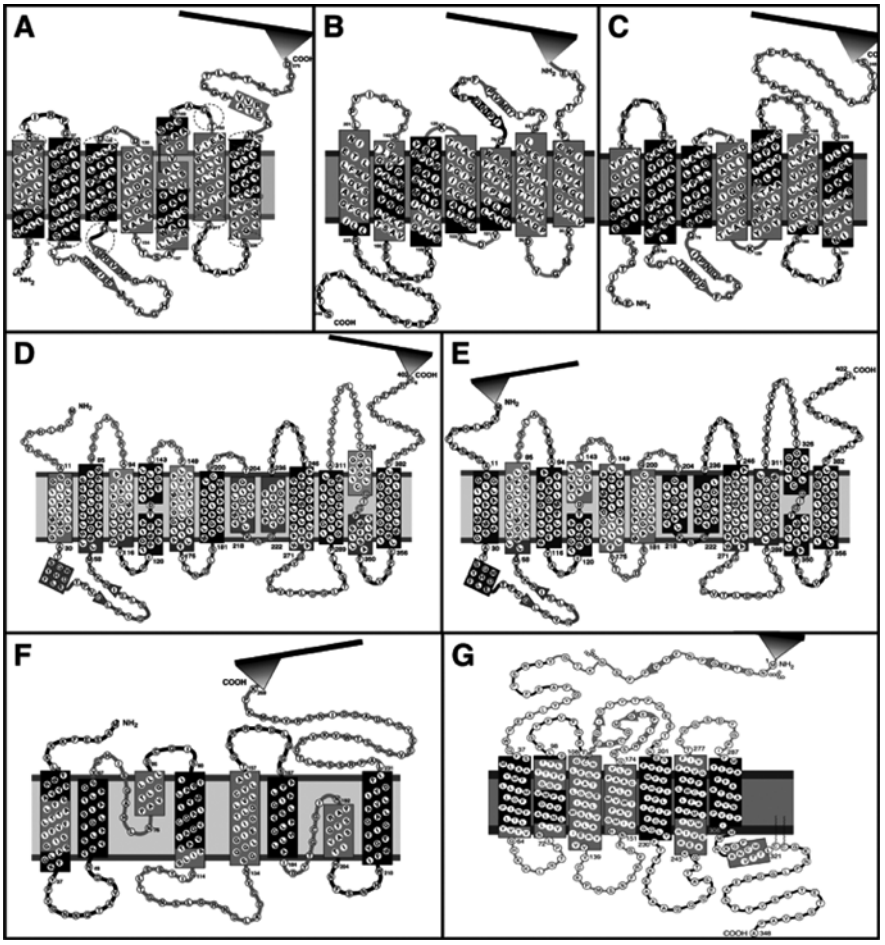


Figure 16.13. Single-molecule force spectroscopy (SMFS) locates molecular interactions stabilizing the structural segments of membrane proteins. **A.** Mapping the secondary structure of halorhodopsin with structural segments that establish unfolding barriers (Cisneros et al., 2005). Single halorhodopsin molecules were mechanically unfolded by pulling their C-terminal end. Mapping the stable structural segments of bacteriorhodopsin detected on unfolding the protein from the N-terminus **(B)** and the C-terminus **(C)** (Kessler and Gaub, 2006). Mapping NhaA with structural segments detected on unfolding the protein from the C-terminus **(D)** and the N-terminus **(E)** (Kedrov et al., 2004). **F**. Mapping stable structural segments of human aquaporin-1 from erythrocytes (Möller et al., 2003). **G.** Stable structural segments of bovine rhodopsin (Sapra et al., 2006b).

of interactions, depend on the environment of the biological system. However, what exactly does the mechanical rupture force signify? Intuitively, one would think the force magnitude to be the strength of a bond. However, this is not the case. A good way to think about forced unfolding is in the context of the lifetime of a bond. Weak bonds and structures with limited lifetimes, t_{off}, will dissociate under almost any level of force if pulled for modest periods of time. If pulled apart faster than t_{off}, a bond will resist detachment, leading to an increase in the measured force. The detachment or unfolding force thus depends on the loading rate. This arises from the fact that when loaded at higher rates a bond lifetime is decreased, and at low loading rates it experiences longer lifetimes close to it's natural lifetime (Evans, 1998, 2001).

A bond under force can be described by using a classical two-state model. An energy profile diagram of a two-state reaction involves a low-energy folded state and a high-energy unfolded state separated by a barrier that needs to be crossed for conversion from one state to the other (Figure 16.14).

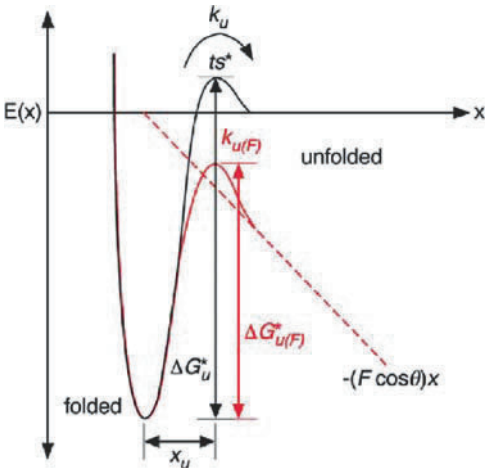


Figure 16.14. Energy barrier under no force and an externally applied force. Shown is a simple two-state energy profile that separates a folded state from an unfolded state. A bond (or folded state of a protein) is confined in a potential well by an energetic barrier $\Delta G_u{}^*$ and can be broken (or unfolded in the case of a protein) with an unfolding rate of k_u via the transition state ts^* at a distance x_u from the bond (or folded) state. An externally applied force F pulls the bond (or protein) at an angle θ to the molecular reaction coordinate x and adds a mechanical potential, $-(F \cos\theta)x$, that tilts the energy barrier and lowers it to $\Delta G^*{}_{u(F)}$. The new unfolding rate under an externally applied force F is $k_{u(F)}$. However, it is assumed that the distance to the transition state, x_u, remains unchanged.

Within the framework of a two-state model it can be conceptualized that the rate of a reaction under an externally applied force is increased due to the lowering of the energy barrier by the external load (Figure 16.14). An external pulling force, F, lowers the energy barrier, ΔG^*_u, at the transition state ($x = x_{ts}$),

$$\Delta G^*_{u(F)} = \Delta G^*_u - Fx_u \tag{16.1}$$

$\Delta G^*_{u(F)}$ and ΔG^*_u represent the free energy difference between the native folded state and the transition state under applied force and zero force, respectively, and $x_u = \langle x_{ts} \cos\theta \rangle$ is the thermally averaged projection (Evans, 2001).

The unfolding rate constant, k_u, of a reaction under no force is given by

$$k_u = \left(\frac{1}{t_D}\right) \exp\left(-\frac{\Delta G^*_u}{k_B T}\right), \quad \Delta G^*_u = G^* - G_n \tag{16.2}$$

where G^* is the free energy of the transition state, G_n is the free energy of the native state, t_D is the relaxation time, k_B is the Boltzmann constant, and T is the absolute temperature. In his

seminal work, Bell (1978) postulated the exponential increase of the rate of escape, k_{off}, with an externally applied force, F,

$$k_{\mathrm{off}} \approx \left(\frac{1}{t_{\mathrm{off}}}\right) \exp\left(\frac{F}{F_{\mathrm{u}}}\right) \tag{16.3}$$

where $F_{\mathrm{u}} = \dfrac{k_{\mathrm{B}}\mathrm{T}}{x_{\mathrm{u}}}$. For a reaction under force, a similar relation is obtained from Eqs. (16.1) and (16.2),

$$k_{\mathrm{u}}(F) \cong \left(\frac{1}{t_{\mathrm{D}}}\right) \exp\left(-\frac{\Delta G_{\mathrm{u}}^{*} - Fx_{\mathrm{u}}}{k_{\mathrm{B}}\mathrm{T}}\right) = k_{\mathrm{u}} \exp\left(\frac{Fx_{\mathrm{u}}}{k_{\mathrm{B}}\mathrm{T}}\right) \tag{16.4}$$

16.4.6. Elucidating the Unfolding Routes of Membrane Proteins

Bacteriorhodopsin is a paradigm for studying membrane protein folding (Curnow and Booth, 2007; Joh et al., 2008). This made it the protein of choice for the first proof-of-principle experiment demonstrating the pulling of a single membrane protein by SMFS (Oesterhelt et al., 2000). It was found that bacteriorhodopsin unfolded via different pathways under mechanical force. Further detailed SMFS experiments revealed a number of interesting facts about the mechanical unfolding pathways of this membrane protein (Müller et al., 2002c). The varied pathways observed during the unfolding of single membrane proteins support the "new view" of protein folding in which a protein slides along the folding funnel to reach its final native folded state (Frauenfelder et al., 1991). Bulk unfolding experiments employing the use of chemical or thermal denaturants do not allow the determination of different populations of unfolding pathways in one experiment. In this section we give a comprehensive overview of what has been learned by unfolding membrane proteins by SMFS, where we stand now, and what the open vistas are for the future.

16.4.7. Molecular Nature of Unfolding Intermediates

Mechanical unfolding experiments of membrane proteins show that certain structural regions unfold stepwise, establishing unfolding barriers against unfolding. The rest of the folded protein forms an unfolding intermediate along the unfolding reaction. However, what stabilizes these unfolding intermediates? The experimental results strongly argue that the amino acids of individual structural segments act collectively to form unfolding barriers and unfold cooperatively over a certain threshold of applied force. There is no general consensus on the number of amino acids required to form a stable structural segment. The results from unfolding membrane proteins indicate that the smallest segment could contain ~6 amino acids and the largest segment up to 40 amino acids (Müller et al., 2002c). Stable structural segments do not necessarily correlate with the secondary structure elements of the protein because a transmembrane helix together with a polypeptide loop can establish a mechanical unfolding barrier.

A key question is to determine the contribution of a single amino acid to the comprehensive interactions stabilizing the intermediate. The first experiments investigating the influence of single point mutations and multiple mutations on the stability of structural segments indicated that there is no straightforward relationship between an amino acid and its contribution towards the stability of a structural segment (Sapra et al., 2008a, b). For example, whereas some mutations change the stability of the structural segments hosting them, other mutations

influence the stability of distant structural segments. More experiments in conjunction with molecular dynamic simulations are needed to shed light on these and related questions.

16.4.8. Membrane Proteins of Similar Structures Show Similar Unfolding Patterns

Despite only ∼30% sequence identity between the topologically similar bacteriorhodopsin and halorhodopsin, the two structures unfold via very similar pathways under an applied load (Cisneros et al., 2005). However, a new unfolding intermediate was detected in helix E of halorhodopsin. This subtle difference introduced by a pi-bulk interaction located between Ala178 and Trp183 in helix E points to an interesting fact. The comprehensive interactions stabilizing the structural segments are disturbed by a local change in the protein structure, thus inducing an interruption in the stabilization of the structural segment. This emphasizes the significance of the evolutionary design of proteins for optimum folding and function. Recently, Bowie and coworkers proposed a hypothesis that during the course of evolution certain key residues in a protein may be replaced by other amino acids leaving the structural constraints of the protein intact (Yohannan et al., 2004). The gist of this idea is that the cumulative interactions of the surrounding amino acids maintain the structural framework of a protein. This may, however, lead to the formation of proteins with very different functions. For example, substituting three key residues in bacteriorhodopsin—a light-driven proton pump—makes it function as sensory rhodopsin II (SRII)—a phototactic receptor (Sudo and Spudich, 2006). One would assume that more extensive substitutions in bacteriorhodopsin are needed to convert it to functional SRII, given that bacteriorhodopsin and SRII share only 26% sequence identity. The similar pathways detected during unfolding of bacteriorhodopsin, halorhodopsin (Cisneros et al., 2005), SRII (Cisneros et al., 2008), and proteorhodopsin (Klyszejko et al., 2008) in SMFS experiments suggest that a minimum number of key residues are involved in the function of these proteins, and that despite their low sequence identity these proteins use similar mechanisms to stabilize their folded functional states. Using the same structural scaffold for different protein functions seems to be an efficient way that nature has adopted to maximize the functional pool of proteins (Vendruscolo and Dobson, 2005).

16.4.9. Detecting Intermediates and Pathways during Refolding of Membrane Proteins

Conventional bulk assays study the folding of membrane proteins in detergent or lipid micelles (Seddon et al., 2004). The denatured states in these assays can represent a broad selection of undefined conformational states with residual helical content estimated to be up to 50% (Hunt et al., 1997a; Riley et al., 1997). In many such folding and unfolding studies the denatured state is merely described by the nonfunctional state of the protein determined by a change in one of its physical properties, for example, a change in the fluorescence intensity. In the absence of a well-defined denatured state, determining the thermodynamic and kinetic stability may be ambiguous. A change in the physical property of the membrane protein may not be a realistic indicator of its final natively folded functional state. To remove such ambiguous parameters and to set the results on a comparable basis requires characterization of the unfolding and refolding of membrane proteins between two well-defined folded and unfolded states. Furthermore, it is crucial to study membrane protein folding and unfolding

with the necessary boundary conditions *in vitro* that mimic the constraints experienced by a membrane protein *in vivo*.

SMFS provides a unique opportunity to follow the folding of individually unfolded polypeptides in the context of the lipid bilayer. Using the AFM cantilever as an anchor, it is possible to refold single membrane proteins inside the lipid bilayer and determine the refolding kinetics of individual structural segments. The refolding protocol consists in the simple steps of mechanical unfolding of the membrane protein followed by relaxation of the partially unfolded polypeptide, and then waiting for a certain refolding time. The final step is mechanically to unfold the refolded polypeptide and to monitor which structural segments have been folded during the waiting time (Figure 16.15). Because the starting unfolded conformation of the protein is a well-defined stretched state, every experiment characterizing the refolding of a single polypeptide starts from an identical condition. Besides membrane proteins this methodology has been successfully used to determine the refolding kinetics of water-soluble proteins (Li et al., 2000; Rief et al., 1997).

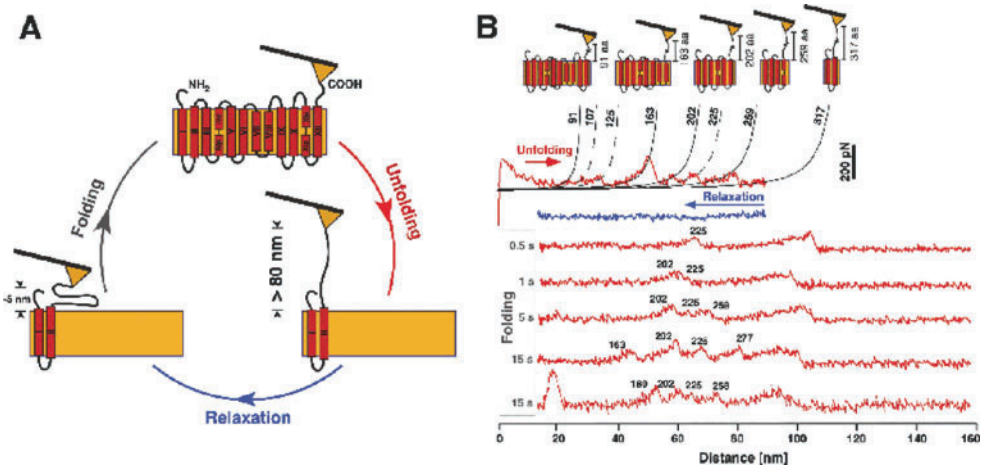


Figure 16.15. Refolding a single antiporter NhaA in the membrane bilayer. **A.** Experimental scheme. Single NhaA molecules were partly unfolded by mechanically pulling their C-terminal end. The unfolded polypeptide was then relaxed close (∼5 nm) to the membrane surface to allow spontaneous refolding (Kedrov et al., 2006a). **B.** Experimental data. Initial unfolding of a single NhaA resulted in a characteristic force–distance (F-D) curve (*upper red curve*). After this, the unfolded polypeptide was relaxed; the relaxation F-D curve is shown in blue. Repeated pulling of the polypeptide (*red curve*) after a certain refolding time provided information on which structural domains have been refolded (Kedrov et al., 2006a). Individual force peaks were fitted with the worm-like chain (WLC) model (*black lines and numbers above*) to locate reestablished structural domains.

The refolding kinetics of membrane proteins was first determined using this strategy for NhaA (Figure 16.15) (Kedrov et al., 2006a) and later for bacteriorhodopsin (Kessler et al., 2006). Characteristic folding rates of structural segments from NhaA ranged from 0.3–2.2 sec^{-1} and were comparable to the folding rates obtained from ensemble folding studies (0.002–13 sec^{-1}) (Booth, 2000). However, transmembrane helix V had the fastest folding rate of ∼50 sec^{-1}, and it could refold even under an external load of ∼30 pN. This structural segment formed the folding core, which was followed by the folding of other structural segments. Helix V hosts two highly conserved residues, Asp163 and Asp164 (Hunte et al., 2005), which carry a negative charge *in vivo* and form an important part of the ligand-binding pocket of NhaA. Under the acidic conditions of the experiment (pH 4.0) these residues are

protonated in the unfolded conformation of the polypeptide, which promotes the rapid folding of helix V. Similar effects were observed while investigating the folding kinetics of bacteriorhodopsin transmembrane α-helix C (Hunt et al., 1997b). The polypeptide's hydrophobicity was tuned via pH-driven protonation of Asp residues to promote proper folding and membrane insertion of helix C. In the case of NhaA it was suggested that, the primary sequence of helix V (Hunte et al., 2005) may have been fine-tuned during evolution to ensure folding priority and stability of the functional NhaA domain.

In agreement with the two-stage folding model (Popot and Engelman, 1990), structural segments of NhaA folded and inserted stepwise in the membrane bilayer. Individual pairs of transmembrane helices that have their connecting polypeptide loop at the opposite membrane surface predominantly folded and partitioned into the lipid bilayer as a single structural unit. Folding of individual helices was observed at a much lower probability. This important trend was previously observed in ensemble folding studies (Engelman et al., 2003; Popot and Engelman, 1990) and molecular dynamic simulations (Khutorsky, 2003). A possible explanation for this phenomenon may be that the pairwise insertion of transmembrane helices is energetically favorable for the transfer of the connecting hydrophilic loop across the hydrophobic core of the bilayer.

The refolding of bacteriorhodopsin was studied under an external load of several piconewtons (Kessler et al., 2006). Unlike the NhaA kinetic measurements, refolding of bacteriorhodopsin was carried out to obtain mechanistic insights into the refolding process of individual transmembrane helices. It was found that helices E and D exerted a force of ∼50 pN on the AFM cantilever during insertion into the lipid bilayer. The area between the snap-in curve and the worm-like chain (WLC) fit provided an estimate of the work done ($8k_BT$ to 43 k_BT) during the folding of the helical hairpin constituted of helices E and D. The estimated free energy gain of $\sim 0.84k_BT$ per residue for the helical hairpin E and D (Kessler et al., 2006) was in good agreement with the values estimated at a membrane–water interface ($0.2k_BT$ to 0.69 k_BT per residue) (Ladokhin and White, 1999; Wieprecht et al., 2002), from helix insertion experiments of M13 coat protein ($0.58k_BT$ per residue) (Soekarjo et al., 1996), and an artificially designed transmembrane peptide ($0.17k_BT$ per residue) (Ladokhin and White, 2004).

16.4.10. Dynamic Energy Landscape

The energy landscape model put forth the "new view" of protein folding (Bryngelson et al., 1995; Oliveberg and Wolynes, 2006; Wolynes et al., 1995). The energy landscape model is based on the premise that unfolded polypeptides from a high-energy state travel in an entropy funnel along different trajectories until they finally reach the energy minimum of the natively folded state. This multidimensional energy landscape is rough and made of bumpy paths with high barriers and deep kinetic traps that a folding polypeptide has to overcome to reach its low-energy destination at the bottom of the funnel. The roughness of the energy landscape has important implications in determining the functional and structural dynamics of a protein. For example, the ruggedness of the landscape is indicative of the coexisting conformational substates of a protein and thus the conformational flexibility of the structure. Consequently, competing interactions stabilizing these substates cannot be simultaneously minimized into a single or a few conformations resulting in an energetically frustrated landscape. Whereas a flexible structure is defined by a rough energy landscape exhibiting numerous energy minima separated by low-energy barriers, a smooth energy surface with one or a few minima is indicative of a rigid protein with all of its energetic interactions satisfied

simultaneously (Oliveberg and Wolynes, 2006). Such an energy landscape is supposed to be minimally frustrated.

A relevant experimental approach for mapping the one-dimensional energy landscape is dynamic force spectroscopy (DFS). DFS induces the unfolding a protein toward the mechanical pulling trajectory at increased pulling speeds (Evans, 1999). Plotting the most probable unfolding force F_p versus $\log_e$(loading rate) maps the most prominent energy barriers in the energy landscape along the force-driven unfolding pathway (Evans, 1998; Evans and Ritchie, 1997). F_p can be described by

$$F_p = \frac{k_B T}{x_u} \ln\left(\frac{x_u \cdot r_f}{k_B T \cdot k_u}\right) \tag{16.5}$$

with r_f the loading rate, x_u the distance between the folded and the transition states, and k_u the unfolding rate in the absence of an externally applied force load. DFS can be used to determine the distance between the transition state and the folded state, x_u, and the unfolding rate k_u of the structural segments of membrane proteins (Janovjak et al., 2004; Sapra et al., 2008a, c). These parameters denote useful values for characterizing individual unfolding barriers taken along an unfolding pathway.

The unfolding energy landscape of bacteriorhodopsin determined by DFS revealed some detailed insights into the stability and folding of the membrane protein (Janovjak et al., 2004). It was found that the unfolding of all structural segments, such as single helices, paired helices with connecting loops, and individual loops, can be described by a two-state process with unfolding rates under zero applied force in the range of $10^{-6} - 10^{-2}$ sec^{-1}. Single transmembrane helices of bacteriorhodopsin showed comparable stability to that observed for the mechanically induced unfolding of small globular proteins like barnase (2.3×10^{-5} sec^{-1}) (Best et al., 2001) and titin immunoglobulin 27 (1.2×10^{-4} sec^{-1}) (Carrion-Vazquez et al., 1999; Williams et al., 2003). The unfolding probability of transmembrane helices was highly dependent on the loading rate. Increasing the loading rates favored the helices unfolding individually, whereas on decreasing the loading rates the helices unfolded in a pairwise fashion (Janovjak et al., 2004). This pairwise unfolding that is favored near equilibrium conditions is in confirmation with the helical hairpin hypothesis of Engelman and Steitz, which proposes that pairwise association of helices plays an important role in stability and folding (Engelman and Steitz, 1981). The distance to the transition state provides a direct measure of the distance that an α-helix has to be stretched for spontaneous unfolding to be initiated. This pulling distance was on the order of a few amino acids (0.3−0.9 nm) and was independent of the length of the structural segment pulled.

Single-molecule DFS can also be applied to determine the energy landscape roughness. Hyeon and Thirumalai have proposed that the energy landscape roughness can be estimated from DFS experiments at different temperatures (Hyeon and Thirumalai, 2003). Recently, the energy landscape roughness of individual transmembrane helices of bacteriorhodopsin was estimated by DFS to be $\sim 5 k_B T$ (Janovjak et al., 2007).

16.4.11. Following the Unfolding Contours of Mutant Proteins in an Energy Landscape

Protein folding has been studied in part to decipher the mechanisms of how and why proteins misfold in certain cases and not in others. A misfolded protein might be thermodynamically more, less, or equally stable as the native folded protein. However, the low

resolution and ensemble nature of bulk unfolding measurements can sometimes be misleading in determining the mechanistic details of such a complicated process as protein folding. Because of the difference in unfolding energy landscapes of proteins *in vivo* and *in vitro* (bulk assays or under an external load), a comparison of single-molecule experiments with bulk experiments is a futile attempt in many cases (Brockwell et al., 2002; Carrion-Vazquez et al., 1999). Here, we present a case study illustrating this point using bacteriorhodopsin as a model system.

The influence of alanine substitutions on the structure and thermodynamic stability of bacteriorhodopsin was recently tested (Faham et al., 2004). Unfolding of bacteriorhodopsin in micelles was induced by increasing the detergent (SDS) concentration. Point mutations (Figure 16.16) were found that increased, decreased, or had no effect on the thermodynamic stability of bacteriorhodopsin. The most destabilizing mutant, Y57A in α-helix B, altered the unfolding free energy, $\Delta\Delta G_u$, by -3.7 ± 0.5 kcal/mol (Faham et al., 2004). Surprisingly, an adjacent mutation M56A increased the stability of the protein ($\Delta\Delta G_u = 1.4 \pm 0.1$ kcal/mol). However, these energetic changes in bacteriorhodopsin were not accompanied by major structural alterations. A number of proline residues at transmembrane α-helix kinks were also replaced by alanine (Yohannan et al., 2004). P50A in α-helix B had no effect on stability, whereas both P91A (α-helix C) and P186A (α-helix F) were found to be somewhat destabilizing, with associated changes in $\Delta\Delta G_u$ of -1.3 ± 0.3 and -0.9 ± 0.1 kcal/mol, respectively. Similar to Y57A and M56A, none of the proline substitutions altered the structure of the proteins beyond local adjustments near the kinks (Faham et al., 2004; Yohannan et al., 2004).

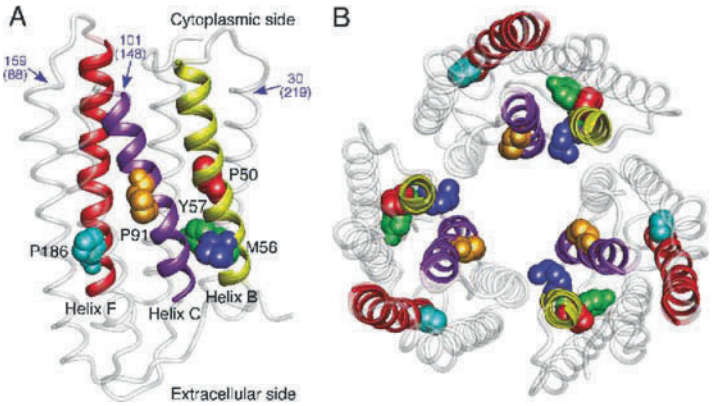


Figure 16.16. Point mutations inserted to characterize their influence on bacteriorhodopsin stability, unfolding pathways, and energy landscape. **A.** Bacteriorhodopsin monomer showing the mutations P50A (*red*), M56A (*blue*), and Y57A (*green*) in α-helix B (*yellow*), P91A (*orange*) in α-helix C (*purple*), and P186A (*cyan*) in α-helix F (*red*). **B.** Top view of the bacteriorhodopsin trimer from the extracellular surface. None of the mutations were at the monomeric interface within the bacteriorhodopsin trimer (Protein Data Bank code 1BRR) (Mitsuoka et al., 1999), and thus they did not affect the integrity of the bacteriorhodopsin assembly (Sapra et al., 2006a). Blue values denote the amino acid positions at which major force peaks (occurring with 100% frequency) were detected in the force–distance curves during single-molecule force spectroscopy (SMFS) unfolding. Values in brackets give amino acid positions as counted from the C-terminal, that is, the direction of pulling by SMFS.

The unfolding response of all of these mutations was studied using DFS (Sapra et al., 2008a). The aim was to characterize the changes in molecular interactions between and within structural segments of bacteriorhodopsin introduced by the above mentioned mutations and to investigate the effect of those changes on the unfolding pathways and the energy landscape

of bacteriorhodopsin. Statistical analyses of all unfolding pathways of the bacteriorhodopsin mutants showed the existence of a major pathway and an exponentially decreasing distribution of all other pathways. The bacteriorhodopsin mutants did not unfold via all of the pathways occurring during wild-type bacteriorhodopsin unfolding. Whereas several unfolding pathways were missing for the mutants, they could also unfold via new alternative pathways. The probabilities of unfolding pathways of wild-type and mutated bacteriorhodopsin were clearly different.

The distance of the transition states from the energy minima of the folded states of the structural segments and the associated changes in the heights of energy barriers in a two-state energy profile were estimated. Interestingly, it was observed that the free energies of the transition states of the bacteriorhodopsin mutants decreased as the distance of those transition states to the folded intermediate states decreased. These results imply that irrespective of the point mutant ΔG_u^* changed proportionally with a change in x_u. Figure 16.17 gives an example of such an energy profile of α-helices B and C along an unfolding path. Previously,

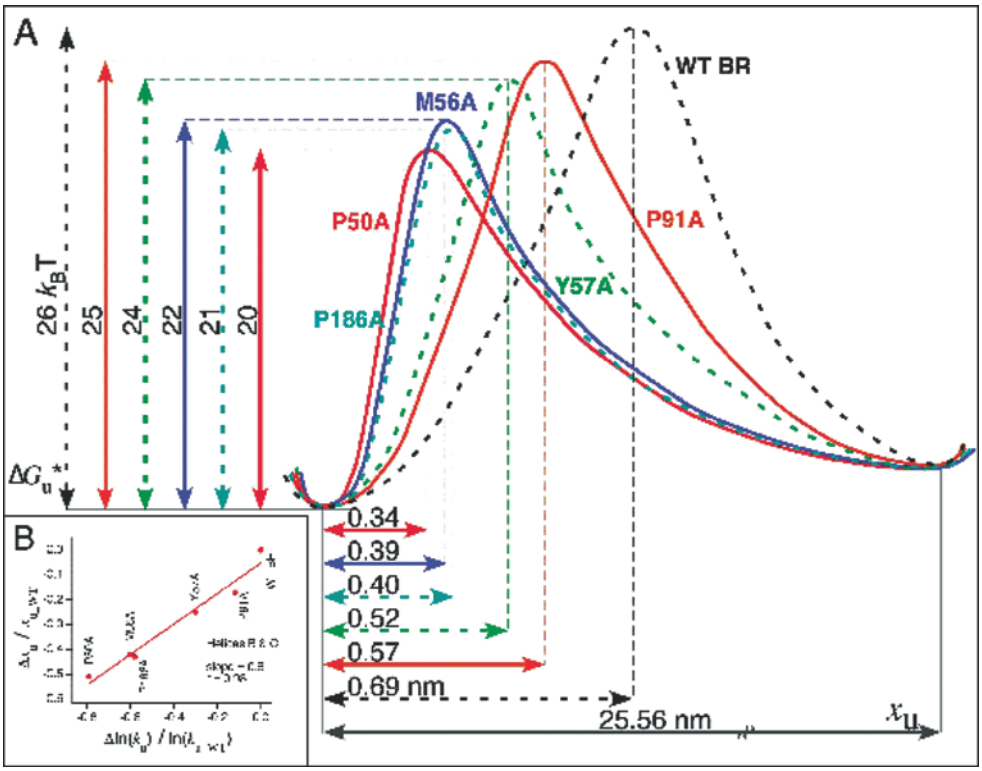


Figure 16.17. Free energy diagram for the mechanical unfolding of transmembrane α-helices B and C in wild-type (WT) and mutant bacteriorhodopsin (BR). **A.** Unfolding barrier heights, ΔG_u^*, and distance to the transition states, x_u, were estimated from the dynamic force spectroscopy data. For the bacteriorhodopsin mutants, P50A, P91A, P186A, M56A, and Y57A, the transition states are significantly shifted toward the folded states of α-helices B and C. All mutations decrease the heights of the unfolding energy barriers. **B.** Quantitative relation between the shift of energy barriers and the decrease in kinetic stability. The plot shows that in the structural segment constituted by transmembrane α-helices B and C of the mutants, x_u increases with increasing ΔG_u^*. Here x_{u_WT} and k_{u_WT} represent parameters from wild-type bacteriorhodopsin, and Δx_u and $\Delta \ln (k_u)$ denote the respective differences between the values of these parameters between wild-type and mutant bacteriorhodopsin.

a similar relationship between x_u and ΔG_u^* was reported for force measurements on single DNA molecules (Strunz et al., 1999) and single antibody–antigen complexes (Schwesinger et al., 2000). Thus, all bacteriorhodopsin mutants exhibited a so-called Hammond behavior. To our knowledge this is the first experimental proof showing that point mutations force single proteins to populate certain unfolding pathways over others (Sapra et al., 2008a).

Taking into account that the difference in the thermodynamic stabilities between the wild-type and mutant proteins is due to changes in molecular interactions, it is safe to speculate from the results that in some cases the populations of major unfolding pathways change with a change in the ΔG_u or ΔG_u^* of the mutants. However, for P50A, which exhibits a thermodynamic stability similar to that of wild-type bacteriorhodopsin, a change in occurrence probabilities of unfolding pathways was still noted, suggesting a possible kinetic role. It is of crucial interest to note in this regard that the presence or absence of proline changes helix–bilayer interactions (Chia et al., 2000; Kobayashi et al., 2000) and has been suggested to influence bacteriorhodopsin folding (Lu et al., 2001). Regardless of whether the balance is under thermodynamic or kinetic control, the probability distribution of the unfolding pathways clearly points to a change in the ratios of many unfolding pathways for the mutant. The origin of this phenomenon may be attributed to mutations altering the transition barrier heights in the energy landscape such that certain unfolding pathways are observed more or less frequently than others.

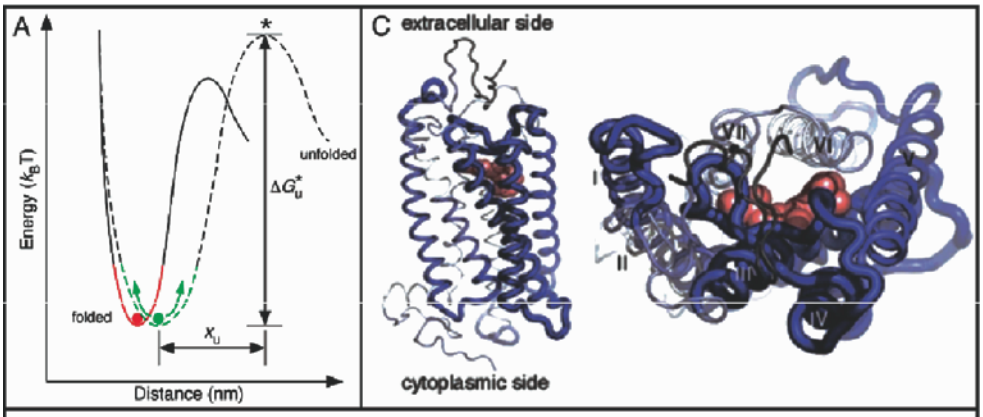
16.4.12. Protein Rigidity, Function, and Energy Landscape

The bovine rhodopsin structure provides opportunities to identify regions critical for proper folding by computational approaches (Jacobs et al., 2001; Rader et al., 2004). The extracellular domain, which is also the folding core of rhodopsin, is tightly coupled functionally and structurally to the transmembrane domain of the molecule (Rader et al., 2004). The identification of amino acids in the rigid and soft cores of rhodopsin has allowed deeper study into the folding and functional mechanisms of this important class of membrane proteins (Tastan et al., 2007).

The parameters x_u and k_u obtained from DFS measurements can be used to determine the stiffness of individual structural segments (Figure 16.18). Recently, the stiffness of all structural segments of bovine rhodopsin and bacteriorhodopsin were estimated and compared (Sapra et al., 2008c). Whereas bovine rhodopsin shows the presence of rigid cores, these cores are absent in bacteriorhodopsin. Computational studies of Klein-Seetharaman and coworkers (Rader et al., 2004; Tastan et al., 2007) predict similar trends in the rigid cores of rhodopsin and bacteriorhodopsin. The presence of rigid cores in rhodopsin points to an interplay of long-range interactions in a cooperative folding mechanism, in contrast to bacteriorhodopsin, where short-range interactions are supposed to drive the individual folding and insertion of helices such as described by the two-stage folding process (Popot and Engelman, 1990).

The signaling mechanism of a GPCR involves the transfer of stimulus from its extracellular to the cytoplasmic domain through an event of well-coordinated conformational changes or changes in dynamic interactions. The conformational changes demand a precise orchestration of helix movement. The presence of both rigid and flexible structural segments (Figure 16.18) suggests that this may help the protein to work as an efficient signaling molecule. The evolution of different regions of a protein to different rigidities may help a protein toward optimization of a smooth synergistic working mechanism.

To test the hypothesis that mechanical rigidity is crucial for a protein's function, we compared the stiffness of all structural segments of a triple mutant of bacteriorhodopsin



B

secondary structure	x_u (±SD) (nm)	k_u (±SD) (s^{-1})	$\Delta G_u^{\ddagger}$ (k_BT)	κ (N/m)
N-terminal region	0.29 ± 0.06	(4.1 ± 7.5) × 10^{-1}	22	2.1
N-terminal region	0.21 ± 0.03	1.7 ± 1.6	20	3.8
α-helix I	0.24 ± 0.02	(5.0 ± 3.7) × 10^{-1}	21	3.0
loop C-I	0.22 ± 0.04	1.3 ± 1.3	20	3.5
α-helix II	0.25 ± 0.03	1.6 ± 1.2	20	2.7
α-helix II	0.47 ± 0.09	(2.1 ± 4.1) × 10^{-2}	25	0.9
loop E-I	0.33 ± 0.03	(7.9 ± 5.9) × 10^{-2}	23	1.8
α-helices III and IV, loops C-II and E-II	0.21 ± 0.02	(14.8 ± 9.3) × 10^{-2}	23	4.2
α-helix V and loop C-III	0.25 ± 0.03	(9.0 ± 8.2) × 10^{-2}	23	3.0
α-helix VI	0.34 ± 0.05	(1.0 ± 1.2) × 10^{-1}	23	1.6
α-helix VI	0.42 ± 0.07	(1.3 ± 1.6) × 10^{-1}	23	1.1
loop E-III	0.34 ± 0.04	(1.8 ± 1.6) × 10^{-1}	22	1.6
α-helix VII	0.36 ± 0.04	(7.4 ± 7.0) × 10^{-2}	23	1.5
α-helix 8	0.44 ± 0.14	(6.6 ± 22.1) × 10^{-3}	26	1.1
C-terminal region	0.33 ± 0.06	(4.5 ± 7.0) × 10^{-2}	24	1.8

Figure 16.18. Mechanical and energetic properties of rhodopsin. **A.** A two-state unfolding energy profile of a structural segment in a membrane protein. The energy profile is assumed to have a parabolic potential well and a sharp energy barrier. The rigidity of a structural segment depends on the curvature of the potential well (*red*, *green*). **B.** The values of x_u and ΔG_u^* obtained from dynamic force spectroscopy experiments were used to calculate the spring constant κ of each structural segment of rhodopsin. The dashed trace in panel A shows that an increase in x_u decreases the rigidity of the structural segment. In general, the height of the energy barrier can increase or decrease with an increase in x_u. **C.** The rigidity κ of structural segments is mapped onto the crystal structure of bovine rhodopsin (Protein Data Base code 1U19) (Okada et al., 2004). Segments are colored from white to blue, with white representing the most flexible region ($\kappa = 0.9$ N/m) and blue representing the most rigid region ($\kappa = 4.2$ N/m). The rigidity of the segments is also indicated by the thickness of the chain, with thicker lines corresponding to higher rigidity.

(E9Q/E194Q/E204Q, 3Glu), with that of wild-type bacteriorhodopsin (Sapra et al., 2008b). 3Glu bacteriorhodopsin is known to have altered functional characteristics compared to those of wild-type bacteriorhodopsin. Glutamate residues E9, E194, and E204 located in the extracellular region of bacteriorhodopsin are involved in a network of hydrogen bonds and are suggested to form part of the proton release pathway (Lanyi, 1997; Luecke et al., 1998). The importance of the extracellular Glu residues and their photoelectrochemical role in bacteriorhodopsin (Bondar et al., 2004; Lazarova et al., 2000; Sanz et al., 2001) provided the rationale for mutating Glu to Gln. Of interest, the stiffness of the structural segments hosting the mutations E194Q and E204Q in 3Glu bacteriorhodopsin decreased as compared to those in wild-type bacteriorhodopsin. Moreover, the same structural segment in 3Glu bacteriorhodopsin unfolded via a new unfolding pathway, with an occurrence probability of 15%.

These results show that the functional capacity of a protein is dictated by its mechanical integrity. Moreover, the results give some of the first experimental evidence that the energy landscape of a protein is not only a component of its energetic properties but also a manifestation of its mechanical characteristics.

16.4.13. Screening Membrane Proteins for Small-Molecule Binding

Secondary active transport is mediated by the uniport, symport, or antiport of ions or small molecules. These transporters play a central role in human health and disease by controlling the amount of solutes and toxins inside a cell. Regulation of proton (H^+) and sodium (Na^+) gradients is involved in virtually every physiological process. Na^+/H^+ antiporters regulate intracellular pH, cellular Na^+ concentration, and cell volume of eukaryotic and prokaryotic organisms. In *Escherichia coli,* two antiporters, NhaA and NhaB, specifically exchange Na^+ and Li^+ ions for H^+, helping the cell to adapt to high environmental salinity and to grow at an alkaline pH (Padan et al., 2001).

Na^+ is a natural ligand of NhaA and binds to helix V bridging amino acids Asp163 and Asp164 at pH 7 (Figure 16.19). This binding manifests itself as an extra force peak in the F-D spectrum of NhaA (Figure 16.19). The appearance of a new force peak is suggestive of new interactions on ligand binding to the protein. The conformity of these results was shown by the reversible nature of the force peak, that is, the force peak disappeared on reducing the pH from 7.7 to 3.8 in the same experiment. The specific nature of these interactions was confirmed by SMFS experiments done in the absence of Na^+ ions at pH 7.7. When unfolded without Na^+ ions the new force peak was not detected in the F-D curves. Based on the frequency of the new peak appearance as a function of pH, the activation pH of NhaA was determined to be between 6 and 7, in excellent agreement with biochemical measurements (Taglicht et al., 1991). These studies were further extended by SMFS experiments for

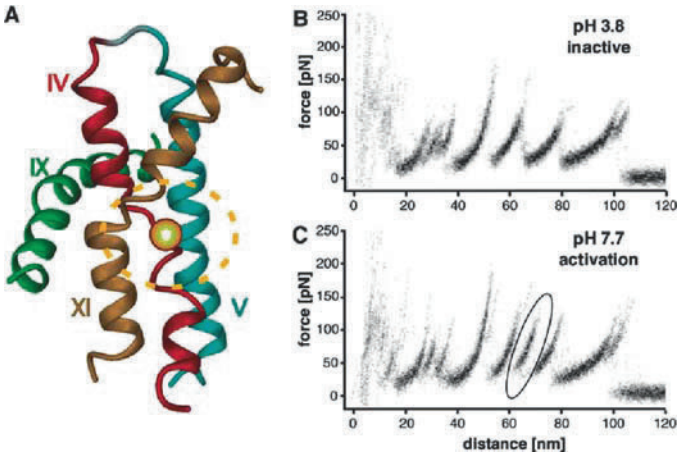


Figure 16.19. Activation of NhaA establishes molecular interactions in the ligand-binding pocket. **A.** The ligand (Na^+ ion) binds to the ligand-binding pocket and establishes interactions at Asp163 and Asp164 at helix V. For clarity, only transmembrane NhaA helices IV, V, IX, and XI are shown. **B.** Single-molecule force spectroscopy (SMFS) data recording interactions of inactive NhaA at pH 3.8. **C.** SMFS data detecting the additional interaction (*ellipse*) established on Na^+ binding in helix V of the ligand-binding pocket. NhaA was activated by shifting the pH of the buffer solution containing NaCl (Kedrov et al., 2005).

detecting inhibitor binding, differentiating ligand and inhibitor binding (Kedrov et al., 2006b), estimating binding constants, and even elucidating the inhibition mechanism (Kedrov et al., 2006b). Recently, Kedrov et al. (2008) compared the dynamics of the energy landscapes of NhaA and NhaA–inhibitor complex. The binding of the inhibitor to α-helix IX stabilized the helix kinetically and entropically by setting it in a narrow potential well, $\sim 9.3 k_B T$ deeper than the potential well of helix IX in the free NhaA state. This consequently led to a decrease in the functionally important structural flexibility of α-helix IX in the inhibitor complex as compared to that in the free NhaA state.

16.5. Outlook

AFM imaging in membrane biology covers a resolution range that bridges high-resolution structure determination techniques (X-ray crystallography, electron crystallography, and nuclear magnetic resonance) and modern light microscopy. The high-resolution techniques can resolve structures of individual proteins but cannot analyze the functional assemblies of several proteins working together. On the other hand, optical microscopy can analyze the global distribution of membrane proteins in a cell but cannot depict detailed supramolecular assemblies due to its limitations in resolution. The AFM, with its outstanding signal-to-noise ratio, occupies a unique position in being able to directly visualize functional assemblies in native membranes. As AFM imaging becomes faster and more sensitive, the information output will become better. We have only been able to reveal high-resolution AFM images of membranes extracted from cells. AFM is also restricted to rather specialized membranes that contain a limited number of protein species. However, the potential for imaging membrane protein ensemble machines from prokaryotes (Scheuring and Sturgis, 2005) and eukaryotes (Buzhynskyy et al., 2007a) is clear. Furthermore, AFM imaging can help to visualize how the supramolecular assemblies of membrane proteins adapt to fulfill the functional needs of the cell (Scheuring and Sturgis, 2005). Many questions remain to be answered: How do cells control membrane protein assembly? What are the underlying interaction mechanisms and physical laws? How does a membrane protein assembly modulate the functional state of its constituting membrane proteins? To answer these questions, high-resolution AFM imaging of living cells must be developed. Certainly, AFM will have to be combined with multifunctional measurements to explain how membrane protein assemblies form to fulfill their functional tasks. Similarly, information complementing structural insights is required to understand the structure–function and assembly–function relationships. In the future, high-resolution AFM imaging may be combined with molecular recognition (Hinterdorfer et al., 1996), single-molecule force spectroscopy (Kedrov et al., 2005), and/or nanomechanical sensing (Sahin et al., 2007) to reveal valuable insights into how membrane proteins work in the complexity of the cellular membrane.

16.5.1. Characterizing Factors That Sculpt the Energy Landscape

The SMFS and DFS experiments reviewed here may have significant physiological relevance. Protein folding/misfolding in the ER happens via multiple steps involving both chemical reactions and multiple protein molecules. Different combinations of these steps give rise to multiple folding and misfolding pathways, as observed *in vitro* with SMFS. The modest folding efficiency for wild-type eukaryotic integral membrane proteins (Kopito, 1999; Seibert et al., 1997) indicates that the energetic balance between folding and misassembly

pathways for many proteins is delicate enough to be tipped in a pathological direction by various mutations (Sanders and Myers, 2004). It has been shown that a number of proteins not implicated in diseases could form amyloid-like plaques if given the optimum environmental factors of pH, temperature, and solvent (Dobson, 2003; Fandrich et al., 2003). How do mutational changes and changes in physical conditions sculpt the energy landscape can be studied by SMFS. This includes characterizing the effect of small molecules or other environmental factors that shift the unfolding landscape of the mutated protein to that of the native one. An example of such an SMFS-based approach is Zn^{2+} binding to rhodopsin, which stabilizes the native Cys110–Cys187 disulfide bridge in the protein, increasing the population of native molecules as compared to molecules without the disulfide bond, which is the implicated diseased state (Park et al., 2007).

Because the (mis)folded (Oberhauser et al., 1999) and functional states of individual proteins can be determined with SMFS (Kedrov et al., 2007), the presented approach may be applied in future to determine which energetic pathways favor membrane protein malfunction, destabilization, and misfolding. More elaborate experimental work in conjunction with molecular dynamic simulations is needed to shed light on the contribution of each pathway in the folding, misfolding, and misassembly of a protein.

16.5.2. Approaches to Screening Drug Targets with Molecular Compounds

GPCRs form a sizeable population of all drug targets (Overington et al., 2006). Their ubiquitous physiological importance has necessitated the development of assays for the screening of drugs, ligands, or other small-molecule libraries. Screening potential lead compounds from a library of molecules is indeed a challenging task. The hurdle is not only to get the right compound, but also to expedite the process. To overcome this difficulty, an ingenious AFM design was constructed recently (Struckmeier et al., 2008). This SMFS robot is a fully automated machine that integrates a buffer exchange, autoalignment and autocalibration, drift compensation, and temperature control system with the advantage of easy analysis in a user-friendly interface. We foresee such SMFS robots as potential tools for screening small libraries of compounds to identify potential drug targets for GPCRs and ion channels in the near future.

References

Agre, P., and Kozono, D. (2003). Aquaporin water channels: molecular mechanisms for human diseases. FEBS Lett *555*, 72–78.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*, 4th ed. (New York, Garland Science).

Alcala, J., Lieska, N., and Maisel, H. (1975). Protein composition of bovine lens cortical fiber cell membranes. Exp Eye Res *21*, 581–595.

Andley, U. (2007). Crystallins in the eye: function and pathology. Prog Retin Eye Res *25*, 78–98.

Ando, T., Kodera, N., Naito, Y., Kinoshita, T., Furuta, K., and Toyoshima, Y. Y. (2003). A high-speed atomic force microscope for studying biological macromolecules in action. Chemphyschem *4*, 1196–1202.

Ando, T., Uchihashi, T., Kodera, N., Yamamoto, D., Taniguch, M., Miyagi, A., and Yamashita, H. (2008). Invited review: high-speed AFM and nano-visualization of biomolecular processes. Eur J Physiol *456*, 211–225.

Ando, T., Uchihashi, T., Kodera, N., Yamamoto, D., Taniguchi, M., Miyagi, A., and Yamashita, H. (2007). High-speed atomic force microscopy for observing dynamic biomolecular processes. J Mol Rec *20*, 448–458.

Aridor, M., and Hannan, M. L. (2000). Traffic jam: a compendium of human diseases that affect intracellular transport processes. Traffic *1*, 836–851.

Aridor, M., and Hannan, M. L. (2002). Traffic jams II: an update of diseases of intracellular transport. Traffic *3*, 781–790.

Bell, G. I. (1978). Models for the specific adhesion of cells to cells. Science *200*, 618–627.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235–242.

Best, R. B., Li, B., Steward, A., Daggett, V., and Clarke, J. (2001). Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation. Biophys J *81*, 2344–2356.

Binnig, G., Quate, C. F., and Gerber, C. (1986). Atomic force microscope. Phys Rev Lett *56*, 930–933.

Binnig, G., Rohrer, H., Gerber, C., and Weibel, E. (1982). Tunneling through a controllable vacuum gap. Appl Phys Lett *40*, 178.

Bondar, A. N., Elstner, M., Suhai, S., Smith, J. C., and Fischer, S. (2004). Mechanism of primary proton transfer in bacteriorhodopsin. Structure *12*, 1281–1288.

Booth, P. J. (2000). Unravelling the folding of bacteriorhodopsin. Biochim Biophys Acta *1460*, 4–14.

Booth, P. J. (2003). The trials and tribulations of membrane protein folding in vitro. Biochim Biophys Acta *1610*, 51–56.

Booth, P. J., and Curran, A. R. (1999). Membrane protein folding. Curr Opin Struct Biol *9*, 115–121.

Borgia, A., Williams, P. M., and Clarke, J. (2008). Single-molecule studies of protein folding. Annu Rev Biochem *77*, 101–125.

Bowie, J. U. (2001). Stabilizing membrane proteins. Curr Opin Struct Biol *11*, 397–402.

Breyton, C., Haase, W., Rapoport, T. A., Kuhlbrandt, W., and Collinson, I. (2002). Three-dimensional structure of the bacterial protein-translocation complex SecYEG. Nature *418*, 662–665.

Brockwell, D. J., Beddard, G. S., Clarkson, J., Zinober, R. C., Blake, A. W., Trinick, J., Olmsted, P. D., Smith, D. A., and Radford, S. E. (2002). The effect of core destabilization on the mechanical resistance of I27. Biophys J *83*, 458–472.

Brodsky, J. L., and McCracken, A. A. (1999). ER protein quality control and proteasome-mediated protein degradation. Semin Cell Dev Biol *10*, 507–513.

Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins *21*, 167–195.

Butt, H.-J., Downing, K. H., and Hansma, P. K. (1990). Imaging the membrane protein bacteriorhodopsin with the atomic force microscope. Biophys J *58*, 1473–1480.

Buzhynskyy, N., Girmens, J.-F., Faigle, W., and Scheuring, S. (2007a). Human cataract lens membrane at sub-nanometer resolution. J Mol Biol *374*, 162–169.

Buzhynskyy, N., Hite, R. K., Walz, T., and Scheuring, S. (2007b). The supramolecular architecture of junctional microdomains in native lens membranes. EMBO Rep *8*, 51–55.

Buzhynskyy, N., Sens, P., Prima, V., Sturgis, J. N., and Scheuring, S. (2007c). Rows of ATP synthase dimers in native mitochondrial inner membranes. Biophys J *93*, 2870–2876.

Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999). Mechanical and chemical unfolding of a single protein: a comparison. Proc Natl Acad Sci USA *96*, 3694–3699.

Chia, Carver, B. C. S., Mulhern, J. A., and Bowie, J. H. (2000). Maculatin 1.1, an anti-microbial peptide from the Australian tree frog, *Litoria genimaculata*—solution structure and biological activity. Eur J Biochem *267*, 1894–1908.

Cisneros, D. A., Oberbarnscheidt, L., Pannier, A., Klare, J. P., Helenius, J., Engelhard, M., Oesterhelt, F., and Muller, D. J. (2008). Transducer binding establishes localized interactions to tune sensory rhodopsin II. Structure, *16*, 1206–1213.

Cisneros, D. A., Oesterhelt, D., and Müller, D. J. (2005). Probing origins of molecular interactions stabilizing the membrane proteins halorhodopsin and bacteriorhodopsin. Structure *13*, 235–242.

Cotton, R. G., and Horaitis, O. (2002). The HUGO Mutation Database Initiative. Human Genome Organization. Pharmacogenomics J *2*, 16–19.

Curnow, P., and Booth, P. J. (2007). Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. Proc Natl Acad Sci USA *104*, 18970–18975.

Czajkowsky, D. M., Hotze, E. M., Shao, Z., and Tweten, R. K. (2004). Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. EMBO J *23*, 3206–3215.

Dill, K. A., and Chan, H. S. (1997). From Levinthal to pathways to funnels. Nat Struct Biol *4*, 10–19.

Dobson, C. M. (2003). Protein folding and misfolding. Nature *426*, 884–890.

Donaldson, P., Kistler, J., and Mathias, R. T. (2001). Molecular solutions to mammalian lens transparency. News Physiol Sci *16*, 118–123.

Drake, B., Prater, C. B., Weisenhorn, A. L., Gould, S. A., Albrecht, T. R., Quate, C. F., Cannell, D. S., Hansma, H. G., and Hansma, P. K. (1989). Imaging crystals, polymers, and processes in water with the atomic force microscope. Science *243*, 1586–1589.

Dupuy, A. D., and Engelman, D. M. (2008). Protein area occupancy at the center of the red blood cell membrane. Proc Natl Acad Sci USA *105*, 2848–2852.

Elie-Caille, C., Severin, F., Helenius, J., Howard, J., Muller, D. J., and Hyman, A. A. (2007). Straight GDP-tubulin protofilaments form in the presence of taxol. Curr Biol *17*, 1765–1770.

Engel, A., and Gaub, H. E. (2008). Structure and mechanics of membrane proteins. Annu Rev Biochem *77*, 127–148.

Engel, A., and Müller, D. J. (2000). Observing single biomolecules at work with the atomic force microscope. Nat Struct Biol *7*, 715–718.

Engelman, D. M. (2005). Membranes are more mosaic than fluid. Nature *438*, 578–580.

Engelman, D. M., Chen, Y., Chin, C. N., Curran, A. R., Dixon, A. M., Dupuy, A. D., Lee, A. S., Lehnert, U., Matthews, E. E., Reshetnyak, Y. K., et al. (2003). Membrane protein folding: beyond the two stage model. FEBS Lett *555*, 122–125.

Engelman, D. M., and Steitz, T. A. (1981). The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell *23*, 411–422.

Evans, E. (1998). Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. Faraday Discuss *111*, 1–16.

Evans, E. (1999). Looking inside molecular bonds at biological interfaces with dynamic force spectroscopy. Biophys Chem *82*, 83–97.

Evans, E. (2001). Probing the relation between force-lifetime and chemistry in single molecular bonds. Annu Rev Biophys Biomol Struct *30*, 105–128.

Evans, E., and Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. Biophys J *72*, 1541–1555.

Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J. P., and Bowie, J.U. (2004). Side-chain contributions to membrane protein structure and stability. J Mol Biol *335*, 297–305.

Fandrich, M., Forge, V., Buder, K., Kittler, M., Dobson, C. M., and Diekmann, S. (2003). Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments. Proc Natl Acad Sci USA *100*, 15463–15468.

Fleschner, C., and Cenedella, R. (1991). Lipid composition of lens plasma membrane fractions enriched in fiber junctions. J Lipid Res *32*, 45–53.

Fotiadis, D., Jastrzebska, B., Philippsen, A., Muller, D. J., Palczewski, K., and Engel, A. (2006). Structure of the rhodopsin dimer: a working model for G-protein–coupled receptors. Curr Opin Struct Biol *16*, 252–259.

Fotiadis, D., Liang, Y., Filipek, S., Saperstein, D. A., Engel, A., and Palczewski, K. (2003). Atomic-force microscopy: rhodopsin dimers in native disc membranes. Nature *421*, 127–128.

Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991). The energy landscapes and motions of proteins. Science *254*, 1598–1603.

Gerber, C., and Lang, H. P. (2006). How the doors to the nanoworld were opened. Nat Nanotechnol *1*, 3–5.

Gonçalves, R. P., Bernadac, A., Sturgis, J. N., and Scheuring, S. (2005). Architecture of the native photosynthetic apparatus of *Phaeospirillum molischianum*. J Struct Biol *152*, 221–228.

Gonçalves, R. P., Buzhynskyy, N., Prima, V., Sturgis, J. N., and Scheuring, S. (2007). Supramolecular assembly of VDAC in native mitochondrial outer membranes. J Mol Biol *369*, 413–418.

Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N. H., Kasas, S., Hansma, H. G., Smith, B., Hansma, P. K., and Bustamante, C. (1999). Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase. Biophys J *77*, 2284–2294.

Hansma, H. G., and Laney, D. E. (1996). DNA binding to mica correlates with cationic radius: assay by atomic force microscopy. Biophys J *70*, 1933–1939.

Hansma, P. K., Cleveland, J. P., Radmacher, M., Walters, D. A., Hillner, P. E., Bezanilla, M., Fritz, M., Vie, D., Hansma, H. G., Prater, C. B., et al. (1994). Tapping mode atomic force microscopy in liquids. Appl Phys Lett *64*, 1738–1740.

Hinterdorfer, P., Baumgartner, W., Gruber, H. J., Schilcher, K., and Schindler, H. (1996). Detection and localization of individual antibody–antigen recognition events by atomic force microscopy. Proc Natl Acad Sci USA *93*, 3477–3481.

Humphris, A. D., Miles, M., and Hobbs, J. K. (2005). A mechanical microscope: high-speed atomic force microscopy. Appl Phys Lett *86*, 34106–34109.

Hunt, J. F., Earnest, T. N., Bousche, O., Kalghatgi, K., Reilly, K., Horvath, C., Rothschild, K. J., and Engelman, D. M. (1997a). A biophysical study of integral membrane protein folding. Biochemistry *36*, 15156–15176.

Hunt, J. F., Rath, P., Rothschild, K. J., and Engelman, D. M. (1997b). Spontaneous, pH-dependent membrane insertion of a transbilayer alpha-helix. Biochemistry *36*, 15177–15192.

Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H. (2005). Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH. Nature *435*, 1197–1202.

Hyeon, C., and Thirumalai, D. (2003). Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? Proc Natl Acad Sci USA *100*, 10249–10253.

Jacobs, D. J., Rader, A. J., Kuhn, L. A., and Thorpe, M. F. (2001). Protein flexibility predictions using graph theory. Proteins *44*, 150–165.

Janovjak, H., Kedrov, A., Cisneros, D. A., Sapra, K. T., and Müller, D. J. (2006). Imaging and detecting molecular interactions of single transmembrane proteins. Neurobiol Aging *27*, 546–561.

Janovjak, H., Kessler, M., Oesterhelt, D., Gaub, H. E., and Müller, D. J. (2003). Unfolding pathways of native bacteriorhodopsin depend on temperature. EMBO J *22*, 5220–5229.

Janovjak, H., Knaus, H., and Muller, D. J. (2007). Transmembrane helices have rough energy surfaces. J Am Chem Soc *129*, 246–247.

Janovjak, H., Struckmeier, J., Hubain, M., Kedrov, A., Kessler, M., and Müller, D. J. (2004). Probing the energy landscape of the membrane protein bacteriorhodopsin. Structure *12*, 871–879.

Jaroslawski, S., Zadek, B., Ashcroft, F., Venien-Bryan, C., and Scheuring, S. (2007a). Direct visualization of KirBac3.1 potassium channel gating by atomic force microscopy. J Mol Biol *374*, 500–505.

Jaroslawski, S., Zadek, B., Ashcroft, F. M., Venien-Bryan, C., and Scheuring, S. (2007b). Direct visualization of KirBac3.1 potassium channel gating by atomic force microscopy. J Mol Biol *374*, 500–505.

Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Woods, V. L., and Bowie, J. U. (2008). Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. Nature *453*, 1266–1270.

Kedrov, A., Appel, M., Baumann, H., Ziegler, C., and Muller, D. J. (2008). Examining the dynamic energy landscape of an antiporter upon inhibitor binding. J Mol Biol *375*, 1258–1266.

Kedrov, A., Janovjak, H., Sapra, K. T., and Muller, D. J. (2007). Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu Rev Biophys Biomol Struct *36*, 233–260.

Kedrov, A., Janovjak, H., Ziegler, C., Kühlbrandt, W., and Müller, D. J. (2006a). Observing folding kinetics and pathways of single antiporters. J Mol Biol *355*, 2–8.

Kedrov, A., Krieg, M., Ziegler, C., Kuhlbrandt, W., and Müller, D. J. (2005). Locating ligand binding and activation of a single antiporter. EMBO Rep *6*, 668–674.

Kedrov, A., Ziegler, C., Janovjak, H., Kuhlbrandt, W., and Müller, D. J. (2004). Controlled unfolding and refolding of a single sodium-proton antiporter using atomic force microscopy. J Mol Biol *340*, 1143–1152.

Kedrov, A., Ziegler, C., and Müller, D. J. (2006b). Differentiating ligand and inhibitor interactions of a single antiporter. J Mol Biol *362*, 925–932.

Kessler, M., and Gaub, H. (2006). Unfolding barriers in bacteriorhodopsin probed from the cytoplasmic and the extracellular side by AFM. Structure *14*, 521–527.

Kessler, M., Gottschalk, K. E., Janovjak, H., Muller, D. J., and Gaub, H. E. (2006). Bacteriorhodopsin folds into the membrane against an external force. J Mol Biol *357*, 644–654.

Khutorsky, V. (2003). Alpha-hairpin stability and folding of transmembrane segments. Biochem Biophys Res Commun *301*, 31–34.

Kihara, A., Akiyama, Y., and Ito, K. (1999). Dislocation of membrane proteins in FtsH-mediated proteolysis. EMBO J *18*, 2970–2981.

Klein-Seetharaman, J. (2005). Dual role of interactions between membranous and soluble portions of helical membrane receptors for folding and signaling. Trends Pharmacol Sci *26*, 183–189.

Klyszejko, A. L., Shastri, S., Mari, S. A., Grubmuller, H., Muller, D. J., and Glaubitz, C. (2008). Folding and assembly of proteorhodopsin. J Mol Biol *376*, 35–41.

Kobayashi, S., Takeshima, K., Park, C. B., Kim, S. C., and Matsuzaki, K. (2000). Interactions of the novel antimicrobial peptide buforin 2 with lipid bilayers: proline as a translocation promoting factor. Biochemistry *39*, 8648–8654.

Kopito, R. R. (1999). Biosynthesis and degradation of CFTR. Physiol Rev *79*, S167–S173.

Ladokhin, A. S., and White, S. H. (1999). Folding of amphipathic alpha-helices on membranes: energetics of helix formation by melittin. J Mol Biol *285*, 1363–1369.

Ladokhin, A. S., and White, S. H. (2004). Interfacial folding and membrane insertion of a designed helical peptide. Biochemistry *43*, 5782–5791.

Lanyi, J. K. (1997). Mechanisms of ion transport across membranes. J Biol Chem *272*, 31209–31212.

Lazarova, T., Sanz, C., Querol, E., and Padros, E. (2000). Fourier transform infrared evidence for early deprotonation of Asp85 at alkaline pH in the photocycle of bacteriorhodopsin mutants containing E194Q. Biophys J *78*, 2022–2030.

Leonhard, K., Guiard, B., Pellecchia, G., Tzagoloff, A., Neupert, W., and Langer, T. (2000). Membrane protein degradation by AAA proteases in mitochondria: extraction of substrates from either membrane surface. Mol Cell *5*, 629–638.

Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Marszalek, P. E., and Fernandez, J. M. (2000). Point mutations alter the mechanical stability of immunoglobulin modules. Nat Struct Biol *7*, 1117–1120.

Lu, H., Marti, T., and Booth, P. J. (2001). Proline residues in transmembrane alpha helices affect the folding of bacteriorhodopsin. J Mol Biol *308*, 437–446.

Luecke, H., Richter, H.-T., and Lanyi, J. K. (1998). Proton transfer pathways in bacteriorhodopsin at 2.3 Angstrom resolution. Science *280*, 1934–1937.

Mangenot, S., Buzhynskyy, N., Girmens, J.-F., and Scheuring, S. (2008). Malformation of junctional microdomains in type II diabetic cataract lens membranes. pflugers Arch *457*, 1265–1274.

McKibbin, C., Farmer, N. A., Jeans, C., Reeves, P. J., Khorana, H. G., Wallace, B. A., Edwards, P. C., Villa, C., and Booth, P. J. (2007). Opsin stability and folding: modulation by phospholipid bicelles. J Mol Biol *374*, 1319–1332.

Mitsuoka, K., Hirai, T., Murata, K., Miyazawa, A., Kidera, A., Kimura, Y., and Fujiyoshi, Y. (1999). The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: implication of the charge distribution. J Mol Biol *286*, 861–882.

Möller, C., Allen, M., Elings, V., Engel, A., and Müller, D. J. (1999). Tapping-mode atomic force microscopy produces faithful high-resolution images of protein surfaces. Biophys J *77*, 1150–1158.

Möller, C., Fotiadis, D., Suda, K., Engel, A., Kessler, M., and Müller, D. J. (2003). Determining molecular forces that stabilize human aquaporin-1. J Struct Biol *142*, 369–378.

Müller, D. J., Amrein, M., and Engel, A. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. J Struct Biol *119*, 172–188.

Müller, D. J., Baumeister, W., and Engel, A. (1996). Conformational change of the hexagonally packed intermediate layer of *Deinococcus radiodurans* monitored by atomic force microscopy. J Bacteriol *178*, 3025–3030.

Muller, D. J., and Dufrene, Y. F. (2008). Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. Nat Nanotechnol *3*, 261–269.

Muller, D. J., and Engel, A. (2007). Atomic force microscopy and spectroscopy of native membrane proteins. Nature Protocols *2*, 2191–2197.

Müller, D. J., and Engel, A. (1999). pH and voltage induced structural changes of porin OmpF explain channel closure. J Mol Biol *285*, 1347–1351.

Müller, D. J., Engel, A., Matthey, U., Meier, T., Dimroth, P., and Suda, K. (2003). Observing membrane protein diffusion at subnanometer resolution. J Mol Biol *327*, 925–930.

Müller, D. J., Fotiadis, D., Scheuring, S., Müller, S. A., and Engel, A. (1999). Electrostatically balanced subnanometer imaging of biological specimens by atomic force microscopy. Biophys J *76*, 1101–1111.

Müller, D. J., Hand, G. M., Engel, A., and Sosinsky, G. (2002). Conformational changes in surface structures of isolated connexin26 gap junctions. EMBO J *21*, 3598–3607.

Müller, D. J., Kessler, M., Oesterhelt, F., Möller, C., Oesterhelt, D., and Gaub, H. (2002c). Stability of bacteriorhodopsin alpha-helices and loops analyzed by single-molecule force spectroscopy. Biophys J *83*, 3578–3588.

Muller, D. J., Sapra, K. T., Scheuring, S., Kedrov, A., Frederix, P. L., Fotiadis, D., and Engel, A. (2006). Single-molecule studies of membrane proteins. Curr Opin Struct Biol *16*, 489–495.

Neuman, K. C., and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nat Methods *5*, 491–505.

Oberhauser, A. F., Marszalek, P. E., Carrion-Vazquez, M., and Fernandez, J. M. (1999). Single protein misfolding events captured by atomic force microscopy. Nat Struct Biol *6*, 1025–1028.

Oesterhelt, F., Oesterhelt, D., Pfeiffer, M., Engel, A., Gaub, H. E., and Müller, D. J. (2000). Unfolding pathways of individual bacteriorhodopsins. Science *288*, 143–146.

Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., and Buss, V. (2004). The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. J Mol Biol *342*, 571–583.

Oliveberg, M., and Wolynes, P.G. (2006). The experimental survey of protein–folding energy landscapes. Q Rev Biophys, 1–44.

Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006). How many drug targets are there? Nat Rev Drug Discov *5*, 993–996.

Padan, E., Venturi, M., Gerchman, Y., and Dover, N. (2001). Na(+)/H(+) antiporters. Biochim Biophys Acta *1505*, 144–157.

Park, P. S., Sapra, K. T., Kolinski, M., Filipek, S., Palczewski, K., and Muller, D. J. (2007). Stabilizing effect of Zn^{2+} in native bovine rhodopsin. J Biol Chem *282*, 11377–11385.

Park, P. S.-H., and Palczewski, K. (2005). Diversifying the repertoire of G protein–coupled receptors through oligomerization. Proc Natl Acad Sci USA *102*, 8793–8794.

Popot, J. L., and Engelman, D. M. (1990). Membrane protein folding and oligomerization: the two-stage model. Biochemistry *29*, 4031–4037.

Rader, A. J., Anderson, G., Isin, B., Khorana, H. G., Bahar, I., and Klein-Seetharaman, J. (2004). Identification of core amino acids stabilizing rhodopsin. Proc Natl Acad Sci USA *101*, 7246–7251.

Rajendran, L., and Simons, K. (2005). Lipid rafts and membrane dynamics. J Cell Sci *118*, 1099–1102.

Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663–669.

Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., and Gaub, H. E. (1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. Science *276*, 1109–1112.

Riley, M. L., Wallace, B. A., Flitsch, S. L., and Booth, P. J. (1997). Slow alpha helix formation during folding of a membrane protein. Biochemistry *36*, 192–196.

Sadlish, H., Pitonzo, D., Johnson, A. E., and Skach, W. R. (2005). Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. Nat Struct Mol Biol *12*, 870–878.

Sahin, O., Magonov, S., Su, C., Quate, C., and Solgaard, O. (2007). An atomic force microscope tip designed to measure time-varying nanomechanical forces. Nat Nanotechnol *2*, 507–514.

Sanders, C. R., and Myers, J. K. (2004). Disease-related misassembly of membrane proteins. Annu Rev Biophys Biomol Struct *33*, 25–51.

Sanders, C. R., and Nagy, J. K. (2000). Misfolding of membrane proteins in health and disease: the lady or the tiger? Curr Opin Struct Biol *10*, 438–442.

Sanz, C., Marquez, M., Peralvarez, A., Elouatik, S., Sepulcre, F., Querol, E., Lazarova, T., and Padros, E. (2001). Contribution of extracellular Glu residues to the structure and function of bacteriorhodopsin. Presence of specific cation-binding sites. J Biol Chem *276*, 40788–40794.

Sapra, K. T., Balasubramanian, G. P., Labudde, D., Bowie, J. U., and Muller, D. J. (2008a). Point mutations in membrane proteins reshape energy landscape and populate different unfolding pathways. J Mol Biol *376*, 1076–1090.

Sapra, K. T., Besir, H., Oesterhelt, D., and Muller, D. J. (2006a). Characterizing molecular interactions in different bacteriorhodopsin assemblies by single-molecule force spectroscopy. J Mol Biol *355*, 640–650.

Sapra, K. T., Doehner, J., Renugopalakrishnan, V., Padros, E., and Muller, D. J. (2008b). Role of extracellular glutamic acids in the stability and energy landscape of bacteriorhodopsin. Biophys J, *95*, 3407–3418.

Sapra, K. T., Park, P. S., Filipek, S., Engel, A., Müller, D. J., and Palczewski, K. (2006b). Detecting molecular interactions that stabilize native bovine rhodopsin. J Mol Biol *358*, 255–269.

Sapra, K. T., Park, P. S., Palczewski, K., and Muller, D. J. (2008c). Mechanical properties of bovine rhodopsin and bacteriorhodopsin: possible roles in folding and function. Langmuir *24*, 1330–1337.

Scheuring, S. (2006). AFM studies of the supramolecular assembly of bacterial photosynthetic core-complexes. Curr Opinion Chem Biol *10*, 387–393.

Scheuring, S., Boudier, T., and Sturgis, J. N. (2007a). From high-resolution AFM topographs to atomic models of supramolecular assemblies. J Struct Biol *159*, 268–276.

Scheuring, S., Buzhynskyy, N., Jaroslawski, S., Gonçalves, R. P., Hite, R. K., and Walz, T. (2007b). Structural models of the supramolecular organization of AQP0 and connexons in junctional microdomains J Struct Biol *160*, 385–394.

Scheuring, S., Buzhysnskyy, N., Gonçalves, R., and Jaroslawski, S. (2007c). Atomic force microscopy: high-resolution imaging of structure and assembly of membrane proteins. In *Biophysical Analysis of Membrane Proteins*, E. Pebay-Peyroula, ed. (Wiley, Chichester, UK) 141–158.

Scheuring, S., Fotiadis, D., Möller, C., Müller, S. A., Engel, A., and Müller, D. J. (2001). Single proteins observed by atomic force microscopy. Single Molecules *2*, 59–67.

Scheuring, S., Gonçalves, R. P., Prima, V., and Sturgis, J. N. (2006). The photosynthetic apparatus of *Rhodopseudomonas palustris*: structures and organization. J Mol Biol *358*, 83–96.

Scheuring, S., Levy, D., and Rigaud, J.-L. (2005). Watching the components of photosynthetic bacterial membranes and their "in situ" organization by atomic force microscopy. Biochim Biophys Acta *1712*, 109–127.

Scheuring, S., Rigaud, J.-L., and Sturgis, J. N. (2004a). Variable LH2 stoichiometry and core clustering in native membranes of *Rhodospirillum photometricum*. EMBO J *23*, 4127–4133.

Scheuring, S., Ringler, P., Borgina, M., Stahlberg, H., Müller, D. J., Agre, P., and Engel, A. (1999). High resolution topographs of the *Escherichia coli* waterchannel aquaporin Z. EMBO J *18*, 4981–4987.

Scheuring, S., and Sturgis, J. N. (2005). Chromatic adaptation of photosynthetic membranes. Science *309*, 484–487.

Scheuring, S., and Sturgis, J. N. (2006). Dynamics and diffusion in photosynthetic membranes from *Rhodospirillum photometricum*. Biophys J *91*, 3707–3717.

Scheuring, S., Sturgis, J. N., Prima, V., Bernadac, A., Lévy, D., and Rigaud, J.-L. (2004b). Watching the photosynthetic apparatus in native membranes. Proc Natl Acad Sci USA *101*, 11293–11297.

Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Guntherodt, H. J., Honegger, A., Jermutus, L., Tiefenauer, L., and Pluckthun, A. (2000). Unbinding forces of single antibody–antigen complexes correlate with their thermal dissociation rates. Proc Natl Acad Sci USA *97*, 9972–9977.

Seddon, A. M., Curnow, P., and Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta *1666*, 105–117.

Seibert, F. S., Loo, T. W., Clarke, D. M., and Riordan, J. R. (1997). Cystic fibrosis: channel, catalytic, and folding properties of the CFTR protein. J Bioenerg Biomembr *29*, 429–442.

Singer, S. J., and Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. Science *175*, 720–731.

Soekarjo, M., Eisenhawer, M., Kuhn, A., and Vogel, H. (1996). Thermodynamics of the membrane insertion process of the M13 procoat protein, a lipid bilayer traversing protein containing a leader sequence. Biochemistry *35*, 1232–1241.

Stahlberg, H., Müller, D. J., Suda, K., Fotiadis, D., Engel, A., Matthey, U., Meier, T., and Dimroth, P. (2001). Bacterial ATP synthase has an undacemeric rotor. EMBO Rep *2*, 229–235.

Stoffler, D., Goldie, K. N., Feja, B., and Aebi, U. (1999). Calcium-mediated structural changes of native nuclear pore complexes monitored by time-lapse atomic force microscopy. J Mol Biol *287*, 741–752.

Struckmeier, J., Wahl, R., Leuschner, M., Nunes, J., Janovjak, H., Geisler, U., Hofmann, G., Jahnke, T., and Müller, D. J. (2008). Fully automated single-molecule force spectroscopy for screening applications. Nanotechnology, *19*, 384020–384030.

Strunz, T., Oroszlan, K., Schafer, R., and Guntherodt, H. J. (1999). Dynamic force spectroscopy of single DNA molecules. Proc Natl Acad Sci USA *96*, 11277–11282.

Sudo, Y., and Spudich, J. L. (2006). Three strategically placed hydrogen-bonding residues convert a proton pump into a sensory receptor. Proc Natl Acad Sci USA *103*, 16129–16134.

Taglicht, D., Padan, E., and Schuldiner, S. (1991). Overproduction and purification of a functional Na+/H+ antiporter coded by nhaA (ant) from *Escherichia coli*. J Biol Chem *266*, 11289–11294.

Tastan, O., Yu, E., Ganapathiraju, M., Aref, A., Rader, A. J., and Klein-Seetharaman, J. (2007). Comparison of stability predictions and simulated unfolding of rhodopsin structures. Photochem Photobiol *83*, 351–362.

Tsai, B., Ye, Y., and Rapoport, T. A. (2002). Retrotranslocation of proteins from the endoplasmic reticulum into the cytosol. Nat Rev Mol Cell Biol *3*, 246–255.

Van den Berg, B., Clemons, W. M. J., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. Nature *427*, 36–44.

Vendruscolo, M., and Dobson, C. M. (2005). A glimpse at the organization of the protein universe. Proc Natl Acad Sci USA *102*, 5641–5642.

Viani, M. B., Pietrasanta, L. I., Thompson, J. B., Chand, A., Gebeshuber, I. C., Kindt, J. H., Richter, M., Hansma, H. G., and Hansma, P. K. (2000). Probing protein–protein interactions in real time. Nat Struct Biol *7*, 644–647.

Viani, M. B., Schäffer, T. E., Paloczi, G. T., Pietrasanta, L. I., Smith, B. L., Thompson, J. B., Richter, M., Rief, M., Gaub, H. E., Plaxco, K. W., et al. (1999). Fast imaging and fast force spectroscopy of single biopolymers with a new atomic force microscope designed for small cantilevers. Rev Sci Instrum *70*, 4300–4303.

White, S. (2008). Membrane proteins of known 3D structure. Available at: http://blanco.biomol.uci.edu/ Membrane_Proteins_xtal.html.

Wieprecht, T., Beyermann, M., and Seelig, J. (2002). Thermodynamics of the coil–alpha-helix transition of amphipathic peptides in a membrane environment: the role of vesicle curvature. Biophys Chem *96*, 191–201.

Williams, P. M., Fowler, S. B., Best, R. B., Toca-Herrera, J. L., Scott, K. A., Steward, A., and Clarke, J. (2003). Hidden complexity in the mechanical properties of titin. Nature *422*, 446–449.

Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995). Navigating the folding routes. Science *267*, 1619–1620.

Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G., and Dalbey, R. E. (2007). Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry *46*, 15153–15161.

Yamamoto, D., Uchihashi, T., Kodera, N., and Ando, T. (2008). Anisotropic diffusion of point defects in two-dimensional crystal of streptavidin observed by high-speed atomic force microscopy. Nanotechnology, *19*, 384009–384018.

Ye, Y., Meyer, H. H., and Rapoport, T. A. (2001). The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol. Nature *414*, 652–656.

Yohannan, S., Faham, S., Yang, D., Whitelegge, J. P., and Bowie, J. U. (2004). The evolution of transmembrane helix kinks and the structural diversity of G protein–coupled receptors. Proc Natl Acad Sci USA *101*, 959–963.

Yu, J., Bippes, C. A., Hand, G. M., Muller, D. J., and Sosinsky, G. E. (2007). Aminosulfonate modulated pH-induced conformational changes in connexin26 hemichannels. J Biol Chem *282*, 8895–8904.

17

High-Speed Atomic Force Microscopy

Toshio Ando and Takayuki Uchihashi

Abstract Biological macromolecules are responsible for the vital activities of life. Among the various approaches to understanding their functional mechanisms, the most straightforward approach is to directly visualize the structure and dynamic action of biological macromolecules at high spatial and temporal resolution. However, the microscopy needed to enable such visualization was not available until the recent development of high-speed atomic force microscopy (AFM). This allows the recording of images of biological samples at 30–60 ms/frame without disturbing delicate biomolecular interactions and hence the delineation of time-series events that occur in biomolecules at work. This chapter describes various devices and techniques developed for high-speed AFM and imaging studies performed on several protein systems.

17.1. Introduction

The dynamic aspect of biological macromolecules is one of the essential attributes. Their functions are produced through their dynamic structural changes and dynamic interactions with other molecules. Furthermore, their functions are produced at the single-molecule level. Therefore, it is important to observe their dynamic behaviors straightforwardly at the single-molecule level to understand their functional mechanisms. As seen in other chapters in this book, single-molecule fluorescence microscopy has been widely exploited for this type of observation. It enables the measurement of translational or rotational motions of individual fluorophores attached to biological molecules and in some cases the measurement of association and dissociation reactions between biological molecules. However, it requires bridging the gap between the recorded fluorescence images and the actual behavior of the labeled biological molecules.

The atomic force microscope (AFM) is capable of directly visualizing unstained biological samples in liquids at nanometer resolution. However, unlike fluorescence microscopy, its imaging rate is low, and hence it cannot trace the fast dynamic processes within the sample. The low imaging rate mainly arises from the fact that AFM uses mechanical sensing and mechanical scanning to detect the sample height at each pixel. Another weak point of

T. Ando and T. Uchihashi • Department of Physics, Kanazawa University, Kakuma-machi, Kanazawa 920-1192 Japan, and CREST, JST, Sanban-cho, Chiyoda-ku, Tokyo 102-0075, Japan

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_17,

AFM also results from the mechanical sensing: The mechanical tip–sample interaction tends to disturb delicate samples and, in the worst-case situation, disrupts fragile samples. In the last decade, various efforts have been made to increase the scan speed, although little attention has been paid to reconciling the fast imaging capability with low-invasiveness imaging. Recent efforts toward this reconciliation have given rise to high-speed AFM, which enables the observation of dynamic biomolecular processes at an imaging rate of 30–60 ms/frame without significantly disturbing delicate biomolecular interactions.

Because this new microscopy has appeared only recently, the user population and the number of successful imaging studies are still limited. We hope that this situation will change in a few years in line with the increase in the availability of commercial high-speed AFM. In this chapter, we first describe the imaging rate as a function of various parameters and also discuss key devices and techniques for high-speed imaging. Next, we give some examples of the successful imaging of dynamic biomolecular processes and describe some techniques and problems associated with the imaging. For the history of high-speed AFM development and the future prospects of high-speed AFM studies in biological research, see recent review articles (e.g., Ando et al., 2007; 2008a, b).

17.2. Basic Principle of AFM and Various Imaging Modes

A typical AFM setup is depicted in Figure 17.1. The sample surface is touched with a sharp tip attached to the free end of a soft cantilever while the sample stage is scanned horizontally (x and y directions). On touching the sample, the cantilever is deflected. Among several methods of sensing this deflection, optical beam deflection (OBD) sensing is often used because of its simplicity; a collimated laser beam is focused onto the cantilever and reflected back into closely spaced photodiodes (a position-sensitive photodetector [PSPD]) whose photocurrents are fed into a differential amplifier. The output of the differential amplifier is proportional to the cantilever deflection. During the raster scan of the sample stage, the detected deflection is compared with the target value (set-point deflection), and then the stage is moved in the z direction to minimize the error signal (the difference between the detected and set-point deflections). This closed-loop feedback operation can maintain the cantilever deflection (hence, the tip–sample interaction force) at the set-point value. The resulting three-dimensional (3D) movement of the sample stage approximately traces the sample surface, and hence a topographic image can be constructed, by using a computer, from the electric signals that are used to drive the sample stage scanner in the z direction. In the operation mode just described (constant-force mode; one of the direct current [DC] modes or contact modes), the cantilever tip, which is always in contact with the sample, exerts relatively large lateral forces on the sample because the spring constant of the cantilever is large in the lateral direction.

To avoid the foregoing problem, tapping-mode AFM (one of the dynamic modes) was invented (Zhong et al., 1993), in which the cantilever is oscillated in the z direction at (or near) its resonant frequency. In this mode, the tip intermittently taps the sample surface at the end of bottom swings. Therefore, little lateral tip force is exerted on the sample, provided that the velocity of the sample stage lateral movement is not too high compared with that of cantilever movement in the z direction. The cantilever oscillation amplitude is reduced by the repulsive interaction between the tip and the sample. Therefore, this mode is also called the amplitude modulation (AM) mode. The amplitude signal is usually generated by a root mean square (RMS)-to-DC converter and is maintained at a constant level (set-point amplitude) by a feedback operation. In AM-AFM, the cantilever oscillation amplitude decreases not only

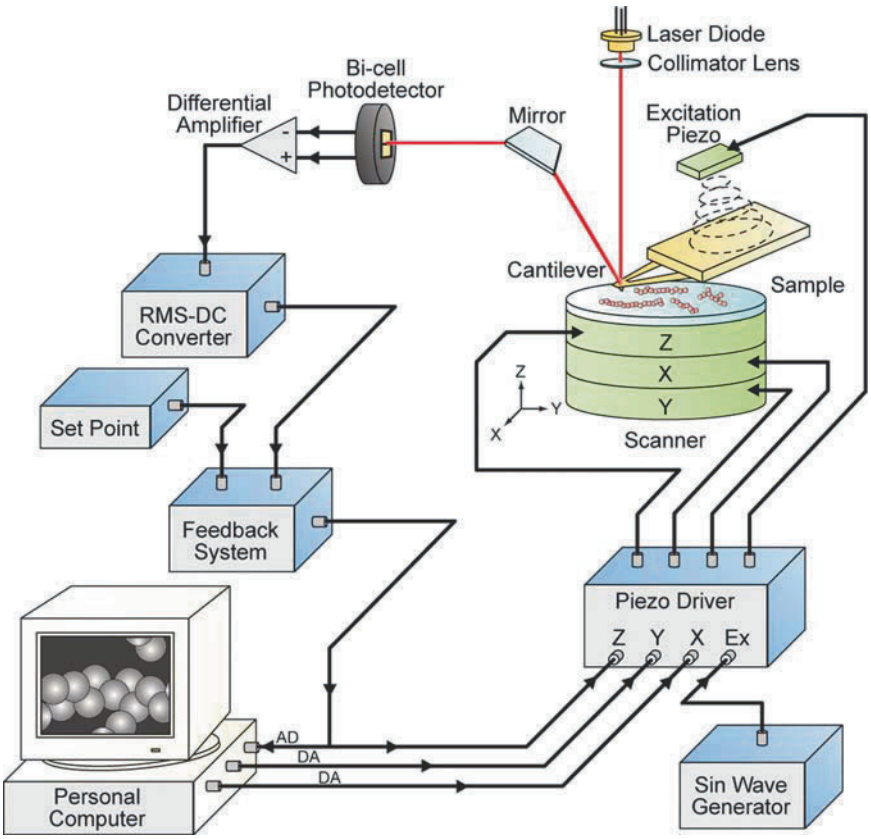


Figure 17.1. Schematic presentation of the tapping-mode AFM system. In the constant-force mode, the excitation piezoelectric actuator and the root mean square-to-direct current (RMS-to-DC) converter are omitted.

by energy dissipation due to the tip–sample interaction, but also by a shift in the cantilever resonant frequency caused by the interaction (Tamayo and García, 1996; Bar et al., 1997; Cleveland et al., 1998). The repulsive interaction produces a positive shift in the resonant frequency because the apparent spring constant k_c of the cantilever is changed by the gradient of the interaction force ($k \equiv \partial f/\partial z$; k is negative for the repulsive interaction). As the excitation frequency is fixed at (or near) the resonant frequency, this frequency shift produces a phase shift of the cantilever oscillation relative to the excitation signal. When this phase shift is maintained by a feedback operation and an image is constructed from the electric signals used for driving the z-scanner, the imaging mode is called the phase modulation (PM) mode. Alternatively, one can construct a phase-contrast image from the phase signal while maintaining the amplitude at a constant level by a feedback operation. More details of the PM mode and phase-contrast imaging are given in Sections 17.4.7 and 17.5.5. Instead of using a fixed frequency, it is possible to set the excitation frequency automatically to the varying resonant frequency of the cantilever by using a self-oscillation circuit (Albrecht et al., 1991; Giessibl, 1995). In this case, the phase of the cantilever oscillation relative to the excitation signal is always maintained at –90°, and the resonant frequency shift is maintained at a constant level by feedback operation. This mode is called the frequency modulation (FM) mode.

17.3. Imaging Rate and Feedback Bandwidth

AFM users do not seem to well understand how a high imaging rate is possible with a given AFM instrument. In this section, on the basis of an idea previously presented (Kodera et al., 2006), we derive the quantitative relationship between the feedback bandwidth and the various factors involved in AFM devices and the scanning conditions in the tapping mode.

17.3.1. Image Acquisition Time and Feedback Bandwidth

Instrumental factors that determine the maximum possible imaging rate are the feedback bandwidth and the maximum frequency at which the scanner can operate without producing unwanted vibrations. As the latter factor is obvious, we only describe the feedback bandwidth. Supposing that a frame is acquired in time T over the scan range $W \times W$ with N scan lines, then the scan velocity V_s in the x direction is given by $V_s = 2WN/T$. Assuming that the sample has a sinusoidal shape with periodicity λ, we see that the scan velocity V_s requires feedback operation at frequency $f = V_s/\lambda$ to maintain the tip–sample distance. The feedback bandwidth f_B should be greater than or equal to f and can therefore be expressed as

$$f_B \geq 2WN/\lambda T \tag{17.1}$$

Equation (17.1) gives the relationship between the frame acquisition time T and the feedback bandwidth f_B. For example, for $T = 30$ ms with $W = 240$ nm and $N = 100$, the scan velocity is 1.6 mm/s. When λ is 10 nm, $f_B \geq 160$ kHz is required to attain this scan velocity. Note that the maximum scan velocity achievable under a given feedback bandwidth depends on the spatial frequency contained in the sample topography, and therefore it is not an appropriate index for evaluating the instrument speed performance.

17.3.2. Feedback Bandwidth as a Function of Various Factors

Various devices are contained in the feedback loop (closed loop) (see Figure 17.1). The sum of the time delays ($\Delta\tau$) that occur with these devices is called "open-loop time delay." Note that the time delay (or phase delay θ) of the closed-loop feedback control is different from the open-loop time delay (or phase delay φ). There is the relationship $\theta \sim 2\varphi$ when the feedback gain is maintained at $\sim$1 (Ando et al., 2008b). Thus, $\theta \sim 2\pi f(2\Delta\tau) = 4\pi f\Delta\tau$, where f is the feedback frequency. In tapping-mode AFM, the main delays are the time (τ_a) required to measure the cantilever oscillation amplitude, the cantilever response time (τ_c), the z-scanner response time (τ_s), the integral time (τ_I) of error signals in the feedback controller, and the parachuting time (τ_p). Here, "parachuting" means that the cantilever tip completely detaches from the sample surface at a steeply inclined region of the sample and cannot quickly land on the surface again. The minimum τ_a is given by $1/(2f_c)$, where f_c is the cantilever's fundamental resonant frequency. Cantilevers and the z-scanner are second-order resonant systems. Therefore, τ_c and τ_s are expressed by $Q_c/(\pi f_c)$ and $Q_s/(\pi f_s)$, respectively, where f_s is the resonant frequency of the z-scanner and Q_c and Q_s are, respectively, the quality factors of the cantilever and the z-scanner. The τ_I and τ_p are functions of various parameters, the approximate analytical expressions of which will be given later. The feedback bandwidth f_B is usually defined by the feedback frequency that results in a phase delay of $\pi/4$. On the basis of this definition, f_B is approximately expressed as

$$f_B = \alpha \frac{f_c}{8} / \left(1 + \frac{2Q_c}{\pi} + \frac{2Q_s f_c}{\pi f_s} + 2f_c\left(\tau_p + \tau_I + \delta\right)\right) \tag{17.2}$$

where δ represents the sum of other time delays and α represents a factor related to the phase compensation effect given by the D component in the proportional integral derivative (PID) feedback controller or in an additional phase compensator. According to our experience, α ~ 2.8. Thus, from Eqs. (17.1) and (17.2), we can estimate the maximum possible imaging rate in a given tapping-mode AFM setup by examining the open-loop time delay $\Delta\tau$. However, this estimation must be modified depending on the sample to be imaged because the allowable maximum phase delay depends on the strength or fragility of the sample. In the next section, we determine the allowable maximum phase delay.

17.3.3. Feedback Operation and Parachuting

To maintain the amplitude of an oscillating cantilever at a constant level while the sample stage is being raster-scanned in the *xy* directions, the detected amplitude is compared with the set-point amplitude. Their difference (error signal) is fed to a PID feedback controller. The PID output is fed to a voltage amplifier to drive the *z*-piezoactuator. This is repeated until the error signal is minimized. To reduce the tapping force exerted from the oscillating tip to the sample, the set-point amplitude should be set close to the cantilever's free oscillation amplitude. However, under this condition, the tip tends to detach completely from the sample surface, particularly at a steep downhill region of the sample. Once detached, the error signal is constant (i.e., saturated at a small level) irrespective of how far the tip is separated from the sample surface. The gain parameters of the PID controller can be increased to reduce the parachuting time. However, such large gains in turn produce an overshoot in the uphill region of the sample, which promotes parachuting around the top region of the sample and leads to instability in the feedback operation.

As has been seen, parachuting is problematic, particularly for high-speed bioAFM in which the tapping force must be minimized. During parachuting, the information of sample topography is completely lost. Here, we determine the conditions that cause parachuting, and obtain a rough estimate of the parachuting time and its effect on the feedback bandwidth (Kodera et al., 2006). The theoretical results obtained here are compared with experimental data to refine the analytical expression for the parachuting time.

When a sample having a sinusoidal shape with periodicity λ and maximum height h_0 is scanned at velocity V_s in the *x* direction, the sample height $S(t)$ under the cantilever tip varies as

$$S(t) = \frac{h_0}{2} \sin(2\pi f t) \tag{17.3}$$

where $f = V_s/\lambda$. When no parachuting occurs, the *z*-scanner moves as

$$Z(t) = -\frac{h_0}{2} \sin(2\pi f t - \theta) \tag{17.4}$$

The feedback error ("residual topography" ΔS) is thus expressed as

$$\Delta S(t) = S(t) + Z(t) = h_0 \sin\frac{\theta}{2} \cos\left(2\pi f t - \frac{\theta}{2}\right) \tag{17.5}$$

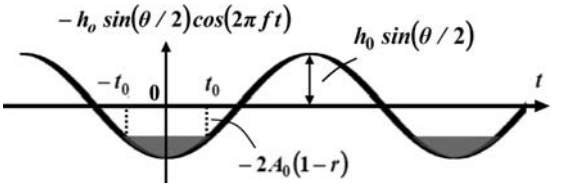


Figure 17.2. The residual topography to be sensed by a cantilever tip under feedback control. When the maximum height of the residual topography is greater than the difference $2A_0(1-r)$, the tip completely detaches from the surface at an area around the bottom. The untouched areas are shown in gray. The average tip-surface separation <*d*> at the end of the cantilever's bottom swing is given by $< d >= \frac{1}{2t_0}\int_{-t_0}^{t_0}\left[-2A_0\,(1-r)+h_0\sin(\theta/2)\,\cos(2\pi ft)\right]dt$, where $t_0 = \beta/2\pi f$ (see text). This integral results in $< d >= 2A_0\,(1-r)\;\;(\tan\beta/\beta - 1)$.

The cantilever tip feels this residual topography (Figure 17.2) in addition to a constant height of $2A_0(1-r)$, where A_0 is the free oscillation amplitude of the cantilever and r is the dimensionless peak-to-peak amplitude set-point. When the set-point peak-to-peak amplitude is denoted as A_s, we have $r = A_s/(2A_0)$.

Because of feedback error, an extra force is exerted onto the sample, the maximum value of which corresponds to a distance of $h_0 \sin(\theta/2)$. Therefore, an allowable maximum phase delay θ_a, which depends on the sample strength, is determined by this distance. The amplitude set-point r is usually determined by compromising between two factors: (1) decrease in tapping force with increasing r, and (2) decrease in the feedback bandwidth with increasing r owing to parachuting. Therefore, the allowable maximum extra force approximately corresponds to $\sim 2A_0(1-r)$, which gives the relationship $\sin(\theta_a/2) \sim (2A_0/h_0)(1-r)$.

When $\Delta S(t) + 2A_0\,(1-r) > 0$, no parachuting occurs. Therefore, the maximum set-point $r_{\max}$ for which parachuting does not occur is given by

$$r_{max} = 1 - \frac{h_0}{2A_0}\,sin\frac{\theta}{2} \tag{17.6}$$

Equation (17.6) indicates that $r_{\max}$ decreases with increasing $h_0/2A_0$ and with increasing phase delay in the feedback operation.

The parachuting time is a function of various parameters, such as the sample height h_0, the free oscillation amplitude A_0 of the cantilever, the set-point r, the phase delay θ, and the cantilever resonant frequency f_c. Its analytical expression cannot be obtained exactly. As a first approximation, we assume that during parachuting, Eq. (17.5) holds and the z position of the sample stage does not move. During parachuting, the average separation between the sample surface and the tip at the end of the bottom swing is given by $2A_0(1-r)(\tan\beta/\beta - 1)$ (see Figure 17.3), where β is given by

$$\beta = cos^{-1}\{2A_0(1-r)/[h_0 sin(\theta/2)]\} \tag{17.7}$$

The feedback gain is usually set to a level at which the separation distance of $2A_0(1-r)$ decreases to approximately zero in a single period of the cantilever oscillation. Therefore, the parachuting time τ_p is expressed as

$$\tau_p = (\tan\beta/\beta - 1)/f_c \tag{17.8}$$

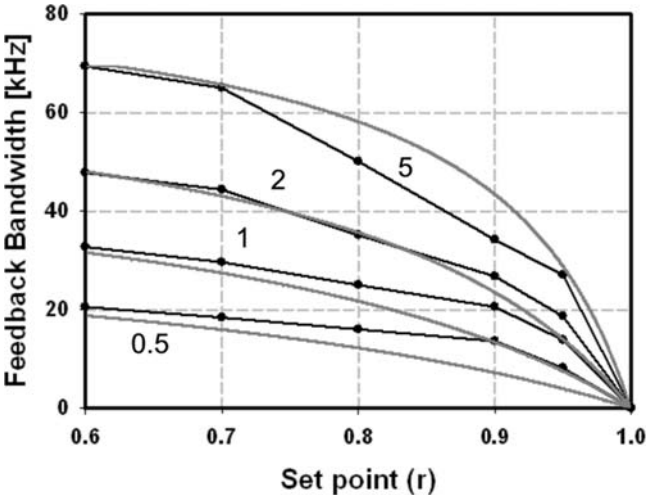


Figure 17.3. Feedback bandwidth as a function of the set-point r and the ratio $2A_0/h_0$ of the free oscillation peak-to-peak amplitude to the sample height. The number attached to each curve indicates the ratio $2A_0/h_0$. The feedback bandwidths were obtained under the following conditions: cantilever resonant frequency, 1.2 MHz; Q factor of cantilever oscillation, 3; resonant frequency of z-scanner, 150 kHz; Q factor of z-scanner, 0.5. The black lines are the experimentally obtained feedback bandwidths using a mock atomic force microscope; the gray lines are the theoretically derived feedback bandwidths.

However, the assumptions under which the average separation during parachuting was derived are different from the reality. As mentioned later (Section 17.3.4), the analytical expression for τ_p should be modified in light of the experimentally obtained feedback bandwidth as a function of r and h_0/A_0.

The main component of PID control is the integral operation. It is difficult to theoretically estimate the integral time constant τ_I with which the optimum feedback control is attained. Intuitively, τ_I should be longer when a greater phase delay exists in the feedback loop. In other words, when a long phase delay exists, the gain parameters of the PID controller cannot be increased. Therefore, τ_I must be proportional to the height of residual topography relative to the free oscillation amplitude of the cantilever. Because the error signals fed into the PID controller are renewed every (half) cycle of the cantilever oscillation, τ_I must be inversely proportional to the resonant frequency of the cantilever. The feedback gain should be independent of parachuting because the gain is maximized so that optimum feedback control is performed for a nonparachuting regime. Thus, τ_I is approximately expressed as $\tau_I = \kappa h_0 \sin(\theta/2)/(A_0 f_c)$, where κ is a proportion coefficient.

17.3.4. Refinement of Analytical Expressions for τ_p and τ_I

We experimentally measured the feedback bandwidth as a function of $2A_0/h_0$ and r using a mock AFM system containing a mock cantilever and z-scanner (Kodera et al., 2006) (Figure 17.3). The mock cantilever and z-scanner are second-order low-pass filters whose resonant frequencies and quality factors are adjusted to have the corresponding values of a real cantilever and z-scanner. The experimentally obtained feedback bandwidths are shown by the black lines in Figure 17.3. Feedback bandwidths are theoretically calculated using

Eq. (17.2) with κ and β as variables and known values of the other parameters. From this analysis, we obtained the following refined expressions for β and τ_I:

$$\beta = \cos^{-1}\{A_0(1 - r)/[5h_0 \sin(\theta/2)]\} \tag{17.9}$$

$$\tau_I = 4h_0 \sin(\theta/2)/(A_0 f_c) \tag{17.10}$$

Feedback bandwidths calculated using these refined expressions are shown by the gray lines in Figure 17.3. They approximately coincide with the experimental data.

17.4. Devices for High-Speed AFM

17.4.1. Small Cantilevers and Related Devices

The cantilever resonant frequency affects the feedback bandwidth via two terms: the amplitude detection time and the cantilever response time [see Eq. (17.2)]. Therefore, cantilevers are the most important devices in high-speed AFM. The resonant frequency f_c and the spring constant k_c of a rectangular cantilever with thickness d, width w, and length L are expressed as

$$f_c = 0.56\frac{d}{L^2}\sqrt{\frac{E}{12\rho}} \tag{17.11}$$

and

$$k_c = \frac{wd^3}{4L^3}E \tag{17.12}$$

where E and ρ are Young's modulus and the density of the material used, respectively. Young's modulus and the density of silicon nitride (Si_3N_4), which is often used as a material for soft cantilevers, are $E = 1.46 \times 10^{11}$ N/m^2 and $\rho = 3{,}087$ kg/m^3, respectively. To attain a high resonant frequency and a small spring constant simultaneously, cantilevers with small dimensions must be fabricated.

The small cantilevers recently developed by Olympus are made of silicon nitride and are coated with gold of ~20 nm thickness. They have a length of 6–7 μm, a width of 2 μm, and a thickness of ~90 nm, which result in resonant frequencies of ~3.5 MHz in air and ~1.2 MHz in water, a spring constant of ~0.2 N/m, and $Q \sim 2.5$ in water. These values in water give the minimum $\tau_a = 0.42$ μs and $\tau_c = 0.66$ μs. We are currently using this type of cantilever, although it is not yet commercially available.

The tip apex radius of the small cantilevers developed by Olympus is ~17 nm (Kitazawa et al., 2003), which is too large for the high-resolution imaging of biological samples. We usually use electron-beam deposition (EBD) to form a sharp tip extending from the original tip. A piece of phenol crystal (sublimate) is placed in a small container, the top of which is perforated with holes ~0.1 mm in diameter. The container is placed in a scanning electron microscope (SEM) chamber, and cantilevers are placed immediately above the container's holes. Spot-mode electron beam irradiation onto the cantilever tip produces a needle on the original tip at a growth rate of ~50 nm/sec. The newly formed tip has an apex radius of

~25 nm and is sharpened by plasma etching in argon or oxygen gas, which decreases the apex radius to ~4 nm.

We developed an OBD detector for small cantilevers (Ando et al., 2001) (Figure 17.4); a laser beam reflected back from the rear side of a cantilever is collected and collimated using the same objective lens as that used for focusing the incident laser beam onto the cantilever. The focused spot is 3–4 μm in diameter. The incident and reflected laser beams are separated using a quarter-wavelength plate and a polarization splitter. Our recent high-speed AFM is integrated with a laboratory-made inverted optical microscope with robust mechanics. The focusing objective lens is also used to view the cantilever and the focused laser spot with the optical microscope. The laser driver is equipped with a radio-frequency (RF) power modulator to reduce noise originating in the optics (Fukuma et al., 2005). The photosensor consists of a four-segment Si PIN photodiode (3 pF, 40 MHz) and a custom-made fast amplifier/signal conditioner (~20 MHz).

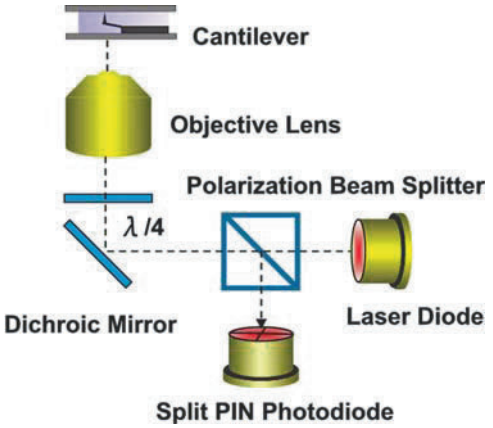


Figure 17.4. Schematic drawing of the objective-lens-type optical beam deflection detection system. The collimated laser beam is reflected up by the dichroic mirror and is incident on the objective lens. The beam reflected at the cantilever is collimated by the objective lens, separated from the incident beam by the polarization beam splitter and the λ/4 wave plate, and reflected onto the split photodiode.

In addition to the advantage of achieving a high imaging rate, small cantilevers have other advantages. The total thermal noise depends only on the spring constant and the temperature and is given by $\sqrt{k_B T/k_c}$, where k_B is Boltzmann's constant and T is the temperature in kelvins. Therefore, a cantilever with a higher resonant frequency has a lower noise density. In the tapping mode, the frequency region used for imaging is approximately the imaging bandwidth (its maximum is the feedback frequency) centered on the resonant frequency. Thus, a cantilever with a higher resonant frequency is less affected by thermal noise. In addition, shorter cantilevers have higher OBD detection sensitivity because the sensitivity is $\Delta\phi/\Delta z = 3/2L$, where Δz is the displacement and $\Delta\phi$ is the change in the angle of a free cantilever end. A high resonant frequency and a small spring constant result in a large ratio f_c/k_c, which gives the cantilever high sensitivity to the gradient k of the force exerted between the tip and the sample. The gradient of the force shifts the cantilever resonant frequency by approximately $-0.5kf_c/k_c$. Therefore, small cantilevers with large values of f_c/k_c are useful for phase-contrast imaging. The practice of phase-contrast imaging using small cantilevers is described in Section 17.5.5.

17.4.2. Tip–Sample Interaction Detection

Tip–sample interactions change the amplitude, phase, and resonant frequency of the oscillating cantilever. They also produce higher harmonic oscillations. In this section, we describe methods for detecting the amplitude and the interaction force. A fast detection method of phase shifts is described in Section 17.4.7.

Amplitude Detectors

Conventional RMS-to-DC converters require at least several oscillation cycles to output an accurate RMS value. To detect the cantilever oscillation amplitude at the periodicity of half the oscillation cycle, we developed a peak-hold method; the peak and bottom voltages are captured and then their difference is output as the amplitude (Figure 17.5) (Ando et al., 2001). The sample/hold timing signals are usually produced using the input signals [shown at (i) in Figure 17.5] (i.e., sensor output signals) themselves. Alternatively, external signals [shown at (ii) in Figure 17.5] that are synchronized with the cantilever excitation signals can be used to produce the timing signals. This is sometimes useful for maximizing the detection sensitivity of the tip–sample interaction because the detected signal is affected by both the amplitude change and the phase shift. This is the fastest amplitude detector, and the phase delay has a minimum value of π. A drawback of this amplitude detector seems to be the detection of noise because the sample/hold circuits capture the sensor signal only at two timing positions. However, the electric noise detected in this peak-hold method is less than that produced by the thermal fluctuations of the cantilever oscillation amplitude.

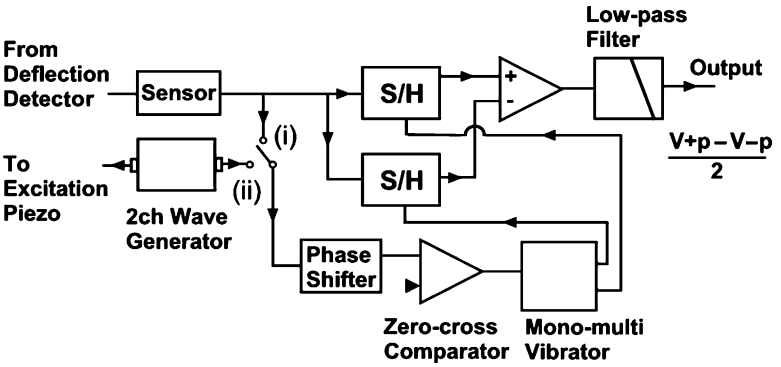


Figure 17.5. Circuit for fast amplitude measurement. The output sinusoidal signal from the split-photodiode amplifier is fed to this circuit. The output of this circuit provides the amplitude of the sinusoidal input signal at half the periodicity of the oscillation signal. 2ch; two-channel.

A different type of amplitude (plus phase) detector can be simply constructed using an analog multiplier and a low-pass filter. The sensor signal $s(t) \sim A_m(t) \sin[\omega_0 t + \varphi(t)]$ is multiplied by a reference signal $2 \sin(\omega_0 t + \phi)$ that is synchronized with the excitation signal. This multiplication produces a signal given by

$$A_m(t)\{\cos[\varphi(t) - \phi] - \cos[2\omega_0 t + \varphi(t) + \phi]\} \tag{17.13}$$

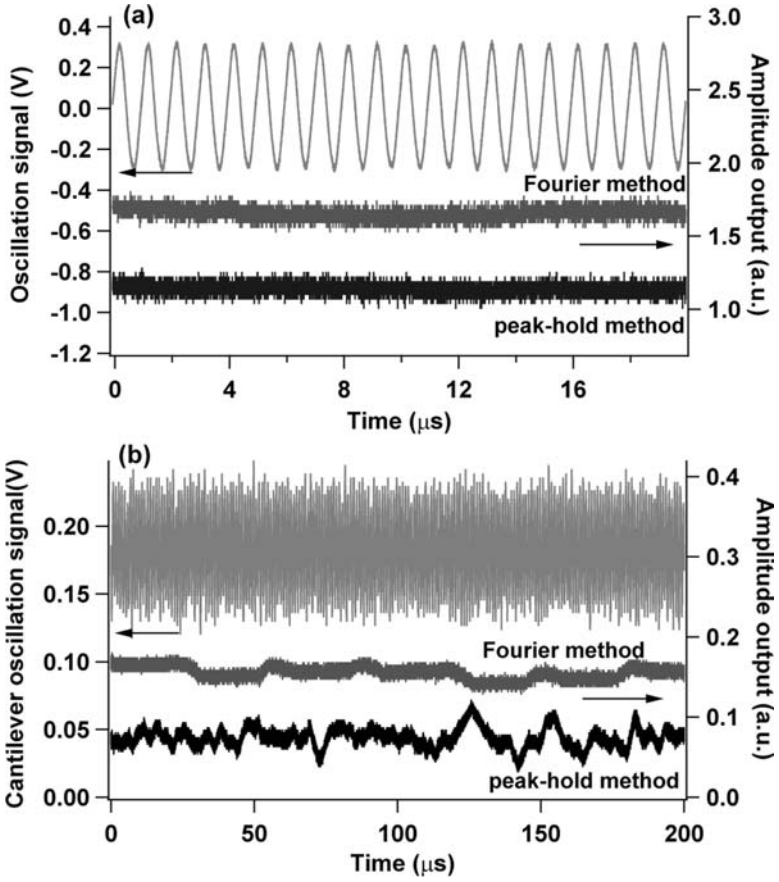


Figure 17.6. Noise levels of two amplitude detection methods (peak-hold method and Fourier method). The upper trace represents the input signal. **(a)** Electric noise. A clean sinusoidal signal mixed with white noise was input to the detectors. The root mean square voltage of white noise was adjusted to be the same magnitude as that of the optical beam deflection photosensor output. **(b)** Variations in the detected cantilever oscillation amplitude.

By adjusting the phase of the reference signal and placing a low-pass filter after the multiplier output, we can obtain a DC signal of $\sim A_m(t)\cos[\Delta\varphi(t)]$, where $\Delta\varphi\,(t)$ is a phase shift produced by the tip–sample interaction. In this method, the delay in the amplitude detection is determined mostly by the low-pass filter. In addition, electric noise is effectively removed by the low-pass filter.

A Fourier method for generating the amplitude signal at the periodicity of a single oscillation cycle has been proposed (Kokavecz et al., 2006). In this method, the Fourier sine and cosine coefficients A and B are calculated for the fundamental frequency from the deflection signal to produce $\sqrt{A^2 + B^2}$. The electric noise level in the Fourier method is similar to that in the peak-hold method (Figure 17.6a). However, regarding the accuracy of amplitude detection, the performance of the Fourier method is better because the cantilever's thermal deflection fluctuations can be averaged in this method (Figure 17.6b).

Force Detectors

A physical quantity that affects the cantilever oscillation is not the force itself but the impulse (∼peak force × time over which the tip interacts with the sample). In the tapping mode, the interaction time is short (∼one-tenth the oscillation period). Consequently, the peak force is relatively large and therefore must be a more sensitive indicator of the interaction than the amplitude change. The impulsive force $F(t)$ consists of harmonic components and therefore cannot be detected directly because the cantilever's flexural oscillation gain is lower at higher harmonic frequencies. $F(t)$ can be calculated from the cantilever's oscillation wave $z(t)$ by substituting $z(t)$ into the equation of cantilever motion and then subtracting the excitation signal (i.e., inverse determination problem) (Stark et al., 2002; Chang et al., 2004; Legleiter et al., 2006). To ensure that this method is effective, the cantilever oscillation signal with a wide bandwidth (at least up to $\sim 5f_c$) must be detected, and fast analog or digital calculation systems are necessary for converting $z(t)$ to $F(t)$. In addition, a fast peak-hold system is necessary to capture the peak force. We are now attempting to build a peak-force detection circuit with a real-time calculation capability.

Recently, a method of directly detecting the impulsive force was presented (Sahin, 2007; Sahin et al., 2007). The torsional vibrations of a cantilever have a higher fundamental resonant frequency (f_t) than that of flexural oscillations (f_c). Therefore, the gain of torsional vibrations excited by impulsive tip–sample interaction is maintained at ∼1 over frequencies higher than f_c. Here, we assume that flexural oscillations are excited at a frequency of $\sim f_c$. To excite torsional vibrations effectively, "torsional harmonic cantilevers" with an off-axis tip have been introduced (Sahin, 2007; Sahin et al., 2007). Oscilloscope traces of torsional vibration signals indicated a time-resolved tip–sample force. We recently observed similar force signals using our small cantilevers with an EBD tip at an off-axis position near the free beam end. After we filtered out the f_c component from the sensor output, periodic force signals appeared clearly (Figure 17.7). To use the sensitive force signals for high-speed imaging, we again need a means of capturing the peak force or a real-time calculation system to obtain it.

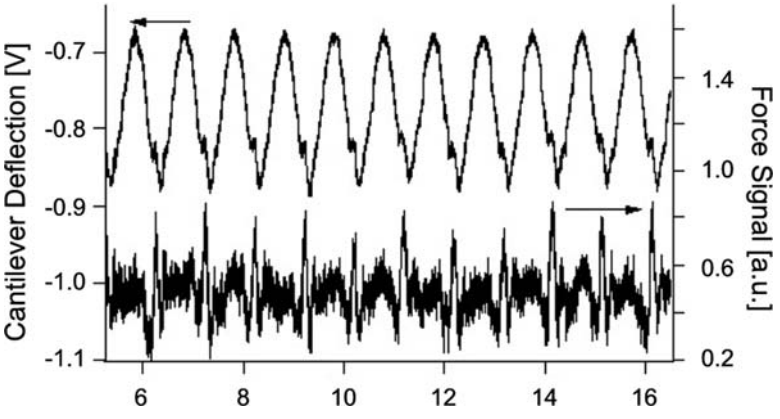


Figure 17.7. Force signal directly obtained from the torsional signal of a small cantilever with an off-axis tip. The cantilever was excited at its first flexural resonant frequency (∼1 MHz) in water. The off-axis tip was intermittently contacted with a mica surface in water. The upper trace shows the torsional vibrations of the cantilever. The lower trace shows the force signal obtained by filtering the torsional signal using a low-pass filter to remove the carrier wave (1 MHz). The torsional signal appears even under free oscillation owing to cross-talk between flexural and torsional vibrations.

17.4.3. High-Speed Scanner

The high-speed driving of mechanical devices having macroscopic dimensions tends to produce unwanted vibrations. Therefore, among the devices used in high-speed AFM, the scanner is the most difficult to optimize for high-speed scanning. Several techniques are required to realize high-speed scanners: (1) a technique to suppress the structural vibrations, (2) a technique to increase the resonant frequencies, (3) an active damping technique to reduce the resonant vibrations of the piezoelectric actuators, and (4) a technique to attain small cross-talk between the three axes.

Counterbalance

The quick displacements of a piezoelectric actuator exert impulsive forces onto the supporting base, which cause vibrations of the base and the surrounding framework and, in turn, of the actuator itself. To alleviate the vibrations, a counterbalance method was introduced for the *z*-scanner (Ando et al., 2001); impulsive forces are countered by the simultaneous displacements of two *z*-piezoelectric actuators of the same length in the counter direction (Figure 17.8a). In this arrangement, which we are currently using, the counterbalance works effectively below the first resonant frequency of the piezoelectric actuators. Because the resonant frequencies originating in the scanner structure are generally lower than those of piezoelectric actuators, structural vibrations are almost completely suppressed. In the arrangement depicted in Figure 17.8a, we can use the maximum displacement of the piezoelectric actuator. However, the resonant frequency becomes approximately half the resonant frequency of the piezoelectric actuator under free oscillation.

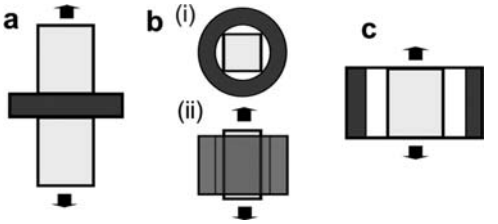


Figure 17.8. Various configurations of holding piezoelectric actuator for suppressing unwanted vibrations. The piezoelectric actuators are shown in light gray, and the holders are shown in black. **(a)** Two actuators are attached to the base. **(b)** An actuator is glued to the solid base at the rims parallel to the displacement direction, top (i) and side (ii) views. **(c)** The two ends of an actuator are held with flexures in the displacement direction.

Figure 17.8b shows a different counterbalance method that we recently developed (Fukuma et al., 2008). A piezoelectric actuator is glued to (or pushed into) a circular hole of a solid base so that the four side rims parallel to the displacement direction are held. Even when held in this way, the actuator can be displaced almost up to the maximum length attained under the load-free condition. In this method, the available maximum displacement becomes approximately half the maximum displacement of the piezoelectric actuator. However, the resonant frequency is almost unchanged from that under free oscillation. This method can also be used for the *x*- and *y*-scanners.

Figure 17.8c shows a counterbalance method that we are using for the *x*-scanner (although we will change it to the rim-holding method described earlier). A piezoelectric actuator is sandwiched between two flexures in the displacement directions. To ensure that

this method is effective, the flexure resonant frequency must be increased, and hence a large spring constant of the flexures must be used, which results in the reduction of the maximum displacement of the scanner.

Mechanical Scanner Design

The structural resonant frequency can be enhanced by adopting a compact structure and a material that has a large Young's modulus–to-density ratio. However, a compact structure tends to produce interference (cross-talk) among the three scan axes. To achieve small cross-talk, we can use flexures (blade springs) that are sufficiently flexible to be displaced but sufficiently rigid in the directions perpendicular to the displacement axis (Kindt et al., 2004; Ando et al., 2005). The scanner mechanism, except for the piezoelectric actuators, must be produced by monolithic processing to minimize the number of resonant elements. We have used an asymmetric structure in the x and y directions; the slowest y-scanner displaces the x-scanner, and the x-scanner displaces the z-scanner, as in our currently used scanner (Figure 17.9). With this scanner, the maximum displacements (at 100 V) of the x- and y-scanners are 1 and 3 μm, respectively. Two z-piezoelectric actuators (maximum displacement, 2 μm at 100 V; self-resonant frequency, 360 kHz) are used in the configuration shown in Figure 17.8a. The gaps in the scanner are filled with an elastomer to passively damp the vibrations. This passive damping is effective in suppressing low-frequency vibrations. To minimize the hydrodynamic pressure generated by the quick displacement of the z-scanner (Ando et al., 2002), we have been using glass sample stages with a circular-trapezoid (or pillar) shape and a small top surface of 1–2 mm diameter.

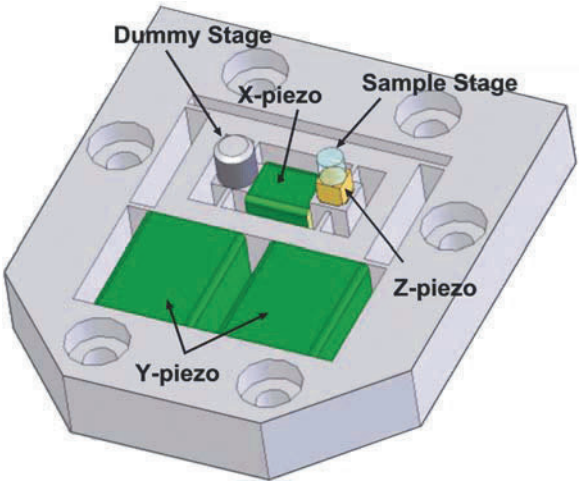


Figure 17.9. Sketch of the high-speed scanner currently used for imaging studies. A sample stage is attached at the top of the upper z-piezoelectric actuator (the lower z-piezoelectric actuator used for counterbalancing is hidden). The dimensions ($W \times L \times H$) of the z-actuators are $3 \times 3 \times 2$ mm^3. The gaps are filled with an elastomer for passive damping.

17.4.4. Active Damping

As mentioned later, the most advanced high-speed scanner has the lowest resonant frequency of ~500 kHz and the maximum displacement of ~1 μm in the z direction. Therefore, we do not need to expand much effort to actively damp the z-scanner vibrations because its

feedback scan (∼150 kHz for video-rate imaging) would not excite the z-scanner resonance of such a high frequency. Here, we describe simple active damping techniques among those developed previously. Because notch filtering, which can be effectively used for eliminating a single clean resonance, is simple, it is not described here.

Feedback Q Control

In the feedback damping method (Figure 17.10a), the feedback operator $H(s)$ converts the resonant system $G(s)$ to a target system $R(s)$ expressed as

$$R(s) = \frac{G(s)}{1 - G(s)H(s)} \tag{17.14}$$

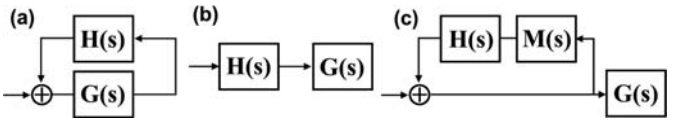


Figure 17.10. Active damping methods for suppressing the unwanted vibrations of the scanner. **(a)** Feedback control method. **(b)** Feedforward control method. **(c)** Feedback Q-control system with a mock scanner. $G(s)$ represents the transfer function of the scanner to be controlled, and $M(s)$ represents the transfer function of the mock scanner. $M(s)$ is similar to the transfer function $G(s)$. $H(s)$ represents the transfer function of the controller for active damping.

In the first step, let us consider the simplest case, where $G(s)$ consists of a single resonant element with resonant frequency ω_1 and quality factor Q_1. The target system is expressed as a single resonator with resonant frequency ω_0 and quality factor Q_0. For these systems, $H(s)$ is expressed as

$$H(s) = \frac{1}{G(s)} - \frac{1}{R(s)} = \left(\frac{s^2}{\omega_1^2} + \frac{s}{Q_1\omega_1} + 1\right) - \left(\frac{s^2}{\omega_0^2} + \frac{s}{Q_0\omega_0} + 1\right) \tag{17.15}$$

To eliminate the second-order term from $H(s)$, ω_0 should equal ω_1, which results in

$$H(s) = -\left(\frac{1}{Q_0} - \frac{1}{Q_1}\right)\frac{s}{\omega_1} \tag{17.16}$$

This $H(s)$ is identical to a derivative operator with a gain of -1 at the frequency $\hat{\omega} = Q_0Q_1\omega_1/(Q_1 - Q_0)$. By adjusting the gain parameter of the derivative operator, we can arbitrarily change the target quality factor Q_0. This method is known as "Q-control" and is often used for controlling the quality factor of cantilevers (Anczykowski et al., 1998; Gao et al., 2001; Tamayo et al., 2001). Here, note that when the Q-controller is applied to the z-scanner, we must measure the displacement or velocity of the z-scanner. However, this is difficult. This problem is solved by using a mock scanner $M(s)$ (a second-order low-pass filter) characterized by the same transfer function as the z-scanner (Figure 17.10c) (Kodera et al., 2005). The practical use of this technique is described in the next section.

Real scanners often exhibit multiple resonant peaks, and therefore the simple Q control just described does not work well. When elemental resonators are connected in series, we must use mock scanners $M_1(s), M_2(s),\ldots$, each of which is characterized by a transfer function representing each elemental resonator. Each mock scanner is controlled by a corresponding

Q-controller. For example, when the scanner consists of two resonators connected in series, the composite transfer function $T(s)$ of the total system is expressed as

$$T(s) = \frac{1}{1 - M_1(s)H_1(s)} \times \frac{1}{1 - M_2(s)H_2(s)} \times G_1(s)G_2(s) \tag{17.17}$$

Because $M_1(s)$ and $M_2(s)$ are the same as (or similar to) $G_1(s)$ and $G_2(s)$, respectively, Eq. (17.17) represents a target system consisting of damped resonators connected in series.

When the scanner consists of elemental resonators connected in parallel, active damping becomes more difficult. Although it does not work perfectly, we can use a mock scanner with two resonators connected in parallel. This gives a better result compared with the case in which the mock scanner with a single resonator is used. More sophisticated methods are described elsewhere (Ando et al., 2008b).

Feedforward Active Damping

The feedforward control type of active damping (Figure 17.10b) is based on inverse compensation [i.e., $H(s) \sim 1/G(s)$]. Generally, inverse compensation–based damping has an advantage, in that we can extend the scanner bandwidth. This damping method is much easier to apply for the x-scanner than for the z-scanner because for the former, the scan waves are known beforehand and are periodic (hence, the frequencies used are discrete, integral multiples of the fundamental frequency) (Schitter et al., 2004; Hung et al., 2006). Here, we only describe the damping of the x-scanner vibrations. The waveforms of the x-scan (as a function of time) are isosceles triangles characterized by amplitude X_0 and fundamental angular frequency ω_0. Their Fourier transform is given by

$$F(\omega) = 2\pi X_0 \left[\frac{1}{2}\delta(\omega) - \frac{2}{\pi^2} \sum_{k=-\infty}^{+\infty} \frac{1}{k^2}\delta(\omega - k\omega_0) \right] \quad (k{:}odd) \tag{17.18}$$

To move the x-scanner in the isosceles triangle waveforms, the signal $X(t)$, which drives the x-scanner characterized by the transfer function $G(s)$, is given by the inverse Fourier transform of $F(\omega)/G(i\omega)$, and is expressed as

$$X(t) = \frac{X_0}{2} - \frac{4X_0}{\pi^2} \sum_{k=1}^{+\infty} \frac{1}{k^2} \frac{1}{G(ik\omega_0)} \cos(k\omega_0 t) \quad (k{:}odd) \tag{17.19}$$

In practice, the sum of the first ~10 terms in the series of Eq. (17.19) is sufficient. We can calculate Eq. (17.19) in advance to obtain numerical values of $X(t)$ and output them in succession from a computer through a D/A converter.

Practice of Active Damping of the Scanner

In this section, we describe the active damping applied to the scanners that we are currently using (one type is shown in Figure 17.9) and to another scanner under development. Here, we first show the effect of feedforward active damping applied to the x-scanner having a large resonance at ~60 kHz. When the line scan was performed at 3.3 kHz without damping, its displacement exhibited vibrations (Figure 17.11a). When it was driven by a waveform

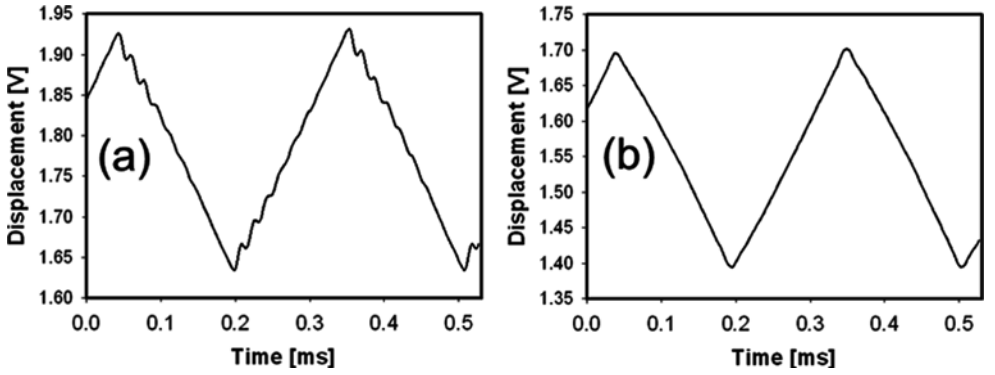


Figure 17.11. Effect of feedforward active-damping control on the *x*-scanner vibrations. **(a)** The *x*-scanner displacement driven by triangular waves without damping. **(b)** The *x*-scanner displacement driven by triangular waves compensated by feedforward control.

calculated using Eq. (17.19) with the maximum term $k = 17$ in the series, the *x*-scanner moved approximately in the isosceles triangle waveform (Figure 17.11b).

In the *z*-scanner that we are currently using for imaging studies, the *z*-piezoelectric actuators have a resonant frequency of 410 kHz under free oscillation. The counterbalance method employed is depicted in Figure 17.8a. With this method of counterbalancing, no structural resonances appear at frequencies lower than the first resonant frequency (171 kHz) of the two identical *z*-piezoelectric actuators (gray lines in Figure 17.12). The resonant frequency of 171 kHz is derived from the fact that one end of each piezoelectric actuator is attached to the base (410/2 = 205 kHz) and the other end is attached to a sample stage (or a dummy stage). Higher resonant frequencies also appear at around 350 kHz. Judging from the phase spectrum (gray line in Figure 17.12b), the two main resonant elements are connected in parallel. For the active damping of these resonant vibrations, we employ the simple *Q*-control method (Figure 17.10c) described in the subsection Feedback *Q* Control. In addition, to compensate for the phase delay produced by active damping, a (1+ derivative) circuit is inserted between the *Q*-controller and the piezodriver. Although this method does not work perfectly, we use a mock *z*-scanner with two resonator circuits connected in parallel. With this method of *Q* control, the first resonance completely disappears (the black line in Figure 17.12a), and the phase delay is not significantly deteriorated (the black line in Figure 17.12b). Because the quality factor is reduced significantly, the *z*-scanner response time is markedly improved, from 17.1 μs to 0.93 μs (Figure 17.13). The slight ringing observed at the rising and falling edges are due to the remaining resonances at around 350 kHz. These resonances could be removed using a notch filter or a low-pass filter, but we do not remove them because such filtering increases the phase delay. In practice, these remaining resonances negligibly disturb the imaging of protein molecules, even at an imaging rate of 30 ms/frame over a scan range of 240 nm with 100 scan lines.

We recently attempted to develop other types of scanners in which a *z*-piezoelectric actuator is held at the four side rims parallel to the displacement direction (Fukuma et al., 2008) (method depicted in Figure 17.8b). The gain and phase spectra of the mechanical response of the *z*-scanner are shown by the black lines in Figures 17.14a, b. The *z*-scanner exhibited large resonant peaks at 440 and 550 kHz. The resonant frequency of 440 kHz is very similar to that of the free oscillation of the piezoelectric actuator. Judging from the

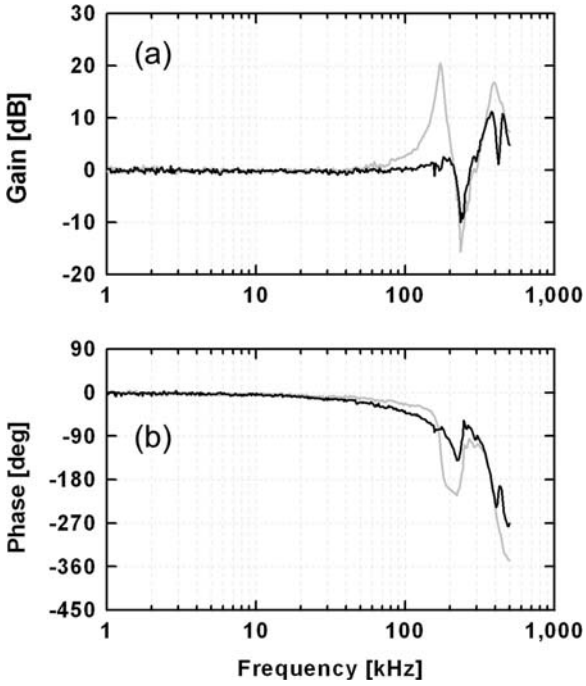


Figure 17.12. Frequency spectra of mechanical response of the z-scanner. **(a)** Gain spectra. **(b)** Phase spectra. Gray lines represent the response without feedback Q control, and black lines represent the response with feedback Q control.

phase spectrum, the two resonators are connected in series. We used two active Q-control circuits connected in series [see Eq. (17.17)]. The resulting gain and phase spectra are shown by the gray lines in Figures 14a, b, respectively. The peak at 440 kHz is almost completely removed, and the frequency that gives a 90° phase delay reaches 250 kHz.

17.4.5. Dynamic PID Control

Dynamic PID Controller

Various efforts have been made to increase the AFM scan speed. However, little attention has been directed toward reducing the tip–sample interaction force. This reduction is quite important for biological AFM imaging. The most ideal scheme is the use of noncontact AFM (nc-AFM), but, to date, high-speed nc-AFM has not been exploited. It is unknown whether the high-speed and noncontact conditions can be reconciled with each other. We previously discussed this matter (Ando et al., 2008a, b). There are several methods of reducing the force in tapping-mode AFM: (1) use of softer cantilevers, (2) enhancement of the quality factor of small cantilevers, and (3) use of a shallower amplitude set-point (i.e., r is close to 1). However, none of these methods appears to be compatible with high-speed scanning. The most advanced small cantilevers developed seem to have reached their limit of balancing high resonant frequency with a small spring constant, and, hence, softer cantilevers can

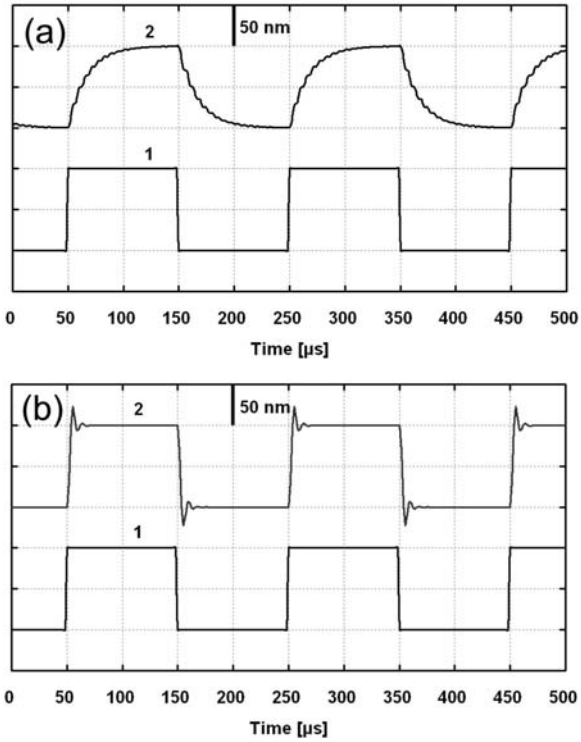


Figure 17.13. Cantilever deflection responses to rectangular variation of the set-point of the proportional integral derivative controller. **(a)** Trace 2 is obtained without active damping of the z-scanner vibrations. **(b)** Trace 2 is obtained with active damping. Traces 1 represent the rectangular set-point variation.

be obtained only by sacrificing the resonant frequency. Although the tapping force decreases with increasing Q in the cantilever, so does its response speed.

The last possibility—a shallower amplitude set-point—promotes "parachuting" during which the error signal is saturated at $2A_0(1 - r)$, and therefore the parachuting time is prolonged with increasing r, resulting in a decrease in the feedback bandwidth. This difficult issue was resolved by the invention of a new PID controller, the "dynamic PID controller," whose gains are automatically changed depending on the cantilever's oscillation amplitude (Kodera et al., 2006). Briefly, a threshold level A_{upper} is set at or slightly above the set-point amplitude $A_s = 2A_0r$. When the cantilever peak-to-peak oscillation amplitude $A_{\text{p-p}}$ exceeds this threshold level, the feedback gain is increased, which either shortens the parachuting time or avoids parachuting entirely. As mentioned in the next section, the dynamic PID controller can avoid parachuting even when r is increased to $\sim$0.9. Note that to ensure that this dynamic PID controller is effective even when the set-point r is very close to 1, the amplitude signal should not contain noise greater than $\sim 2A_0(1 - r)/3$.

Similar manipulation of the error signal can also be performed when $A_{\text{p-p}}$ is smaller than A_s. In this case, a new threshold level A_{lower} is set with a value sufficiently lower than A_s. When $A_{\text{p-p}}$ becomes lower than A_{lower}, the feedback gain is increased. This can prevent the cantilever tip from pushing into the sample too strongly, particularly at steep uphill regions of the sample.

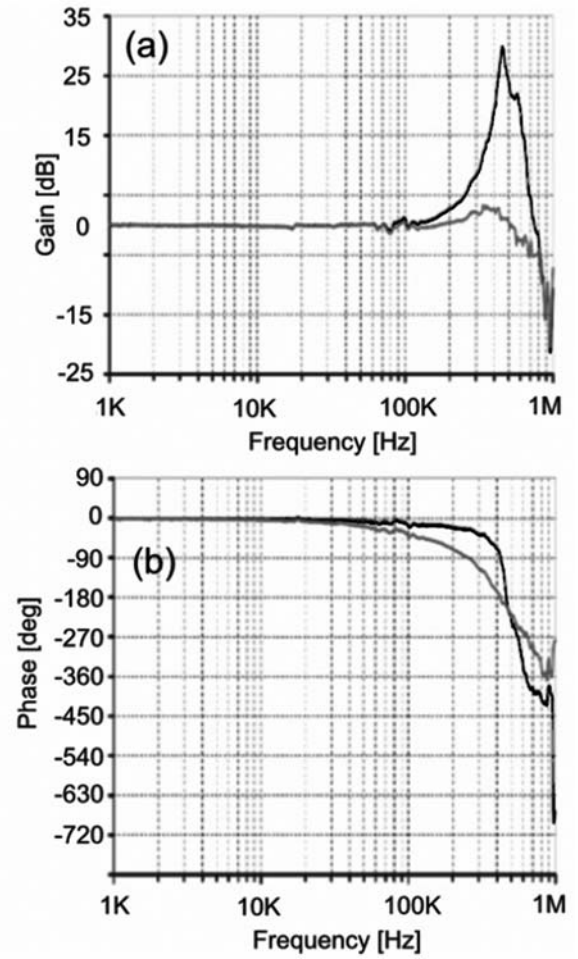


Figure 17.14. Frequency spectra of the mechanical response of the z-scanner. **(a), (b).** The gain and phase spectra, respectively, of a z-scanner whose piezoelectric actuator is held at the four rims parallel to the displacement direction (see Figure 17.8b). Black lines represent the responses without feedforward active damping, and gray lines represent the responses with feedback Q control.

Performance of Dynamic PID Control

Dynamic PID control significantly enhances the feedback bandwidth, particularly when the set-point r is close to 1 (dotted curves in Figure 17.15). The feedback bandwidth becomes independent of r, provided r is less than $\sim$0.9, indicating that parachuting does not occur. The superiority of the dynamic PID control is also clear from captured images of a mock sample with steep slopes (Figure 17.16a). The images were obtained using a mock AFM. Here, a mock cantilever with $Q = 3$ oscillating at its resonant frequency of 1.2 MHz is scanned over a mock sample surface (rectangular shapes with two different heights) from left to right at a scan speed of 1 mm/sec (frame rate of 100 ms/frame). Here, the height of the taller rectangle is $2A_0$, and A_s is set at $0.9 \times 2A_0$. When the conventional PID controller was used,

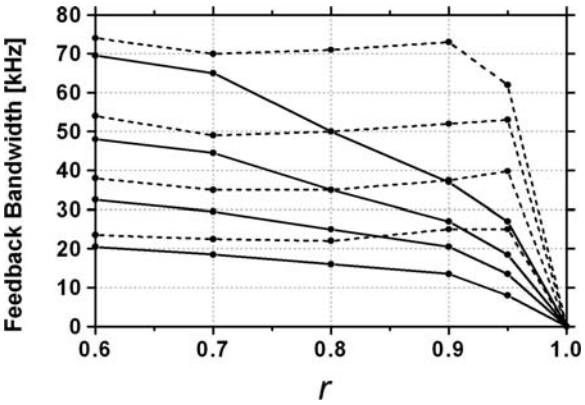


Figure 17.15. Feedback bandwidth as a function of the set-point r, measured using the mock atomic force microscope (AFM) system. The solid curves show the feedback bandwidths measured using the mock AFM system with a conventional proportional integral derivative (PID) controller. The dotted curves show the feedback bandwidths measured using the mock AFM system with the dynamic PID controller. The solid curves and the dotted curves are aligned from top to bottom according to the ratio $2A_0/h_0 = 5$, 2, 1, and 0.5, respectively.

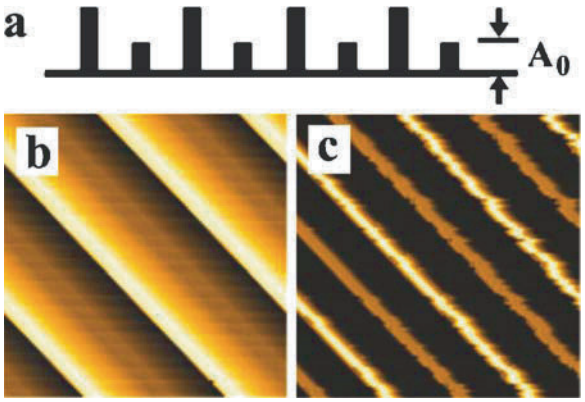


Figure 17.16. Pseudo–atomic force microscope (AFM) images of a sample with rectangles of two different heights. **(a)** Mock AFM sample. **(b), (c)** The images were obtained using a conventional proportional integral derivative (PID) controller and the dynamic PID controller, respectively. The simulations with the mock AFM system were performed under the following conditions: cantilever resonant frequency, 1.2 MHz; quality factor of cantilever oscillation, 3; resonant frequency of z-scanner, 150 kHz; quality factor of z-scanner, 0.5; line scan speed, 1 mm/sec; line scan frequency, 1 kHz; frame rate, 100 ms/frame; ratio of $2A_0$ to total sample height, 1; and $r = 0.9$

the topographic image became blurred (Figure 17.16b) due to significant parachuting at steep downhill regions. On the other hand, the use of the dynamic PID controller resulted in a clear image (Figure 17.16c), indicating that almost no parachuting occurred. Dynamic PID control improves both the feedback bandwidth and the tapping force simultaneously and is therefore indispensable in the high-speed AFM imaging of delicate samples.

17.4.6. Drift Compensator

The drift in the cantilever excitation efficiency poses a problem, particularly when r is close to 1 (i.e., A_s is set close to $2A_0$). For example, as the efficiency is lowered, $2A_0$ decreases and, concomitantly, $A_{\text{p-p}}$ decreases. The feedback system interprets this decrease in $A_{\text{p-p}}$ as an overly strong interaction of the tip with the sample and therefore withdraws the sample stage from the cantilever; this is an incorrect operation. This withdrawal tends to dissociate the cantilever tip completely from the sample surface. For example, when $2A_0 = 5$ nm and $A_s = 4.5$ nm ($r = 0.9$), their difference is only 0.5 nm. The satisfactory operation of dynamic PID control under such a set-point condition requires high stability of the excitation efficiency. The stabilization of the excitation efficiency was previously attempted (Schiener et al., 2004) by using the second-harmonic amplitude of the cantilever oscillation to detect drifts. The second-harmonic amplitude is sensitive to the tip–sample interaction, and therefore the drift in A_0 is reflected in the second-harmonic amplitude averaged over a period longer than the image-acquisition time. To compensate for drift in the cantilever excitation efficiency, we also used the second-harmonic amplitude of cantilever oscillation, but instead of controlling A_s, we controlled the output gain of a wave generator (WF-1946A, NF Corp., Osaka, Japan) connected to the excitation piezoelectric actuator (Kodera et al., 2006). We only used an I-controller whose time constant was adjusted to 1–2 sec (about ten times longer than the image acquisition time). By this drift compensation method, we achieved very stable imaging, even with a small difference ($2A_0 - A_s$) of 0.4 nm.

17.4.7. High-Speed Phase Detector

The material property map is obtained by measuring the phase difference between the excitation signal and the cantilever oscillation (Tamayo and Garcia, 1996; Bar et al., 1997; Cleveland et al., 1998). The phase difference is related to several material properties, such as viscoelasticity, elasticity, adhesion, hydrophobicity/hydrophilicity (Tamayo and Garcia, 1996), and surface charge (Czajkowsky et al., 1998). The energy dissipation of an oscillating cantilever due to inelastic tip–sample interactions is considered to be the main mechanism in the generation of phase contrast (Cleveland et al., 1998). Combining the information on these properties with a topographic video image with unprecedented temporal resolution would provide a deeper insight into biomolecular functional mechanisms and dynamic processes.

For high-speed phase-contrast imaging, we need a fast phase detector. As mentioned in Section 17.4.1, the cantilever resonant frequency shifts by approximately $-0.5kf_c/k_c$ by energy-conservative tip–sample interaction. The frequency shift results in a phase shift as the excitation frequency is fixed. For a given frequency shift, the phase shift increases with Q. With conventional cantilevers, the frequency shift is generally around 50 Hz. Therefore, phase-contrast imaging had been possible only with a large Q (hence only at a low imaging rate). Because the ratio f_c/k_c with the most advanced small cantilevers is ~1,000 times larger than that with conventional cantilevers, we can expect a large shift of ~50 kHz. Therefore, even with a small Q, a relatively large phase shift occurs. Even in the energy-dissipative tip–sample interaction, we can expect a large phase shift with small cantilevers because of a relatively small damping effect of the surrounding medium. Thus, we do not need to detect the phase shift using a very sensitive yet slow phase detector such as a lock-in amplifier. To explore the possibility of high-speed phase-contrast imaging, a fast phase detector was developed (Uchihashi et al., 2006) on the basis of a previous design (Stark and Guckenberger, 1999). Figure 17.17 shows the operational principle of the fast phase detection. A two-channel

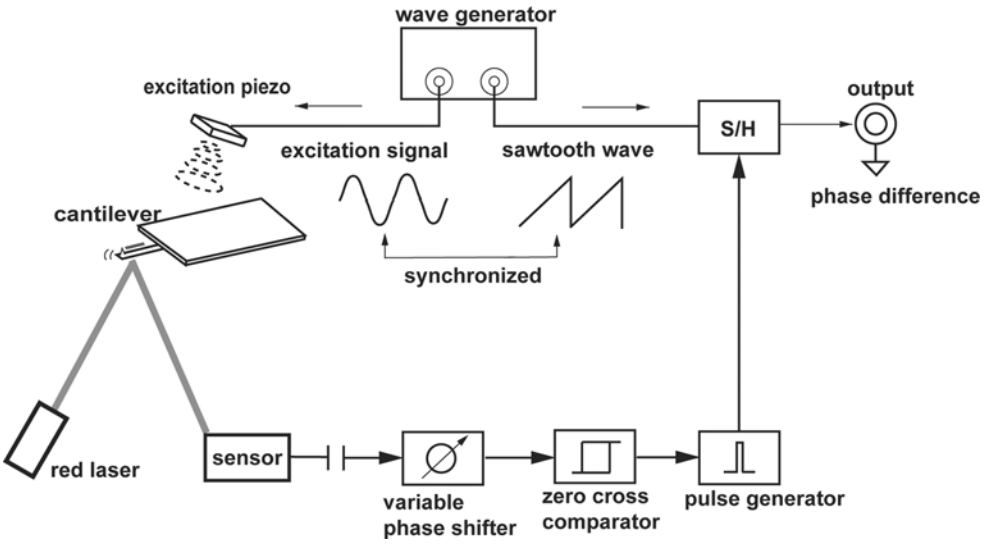


Figure 17.17. Schematic diagram of the fast phase-detection system for the atomic force microscope. S/H, sample-and-hold.

waveform generator produces sinusoidal and sawtooth waves with a given phase difference. The sinusoidal wave is used to oscillate the cantilever. The cantilever oscillation signal is fed to a phase detector composed of a high-pass filter, a variable phase shifter, a zero-crossing comparator, a pulse generator, and a sample-and-hold (S/H) circuit. A pulse signal is generated at either the rising or falling edge of the output signal of the zero-crossing comparator. The pulse signal acts as a trigger for the S/H circuit. When the trigger is generated, the amplitude voltage of the sawtooth signal is captured and retained by the S/H circuit. Thus, the sawtooth signal acts as a phase-voltage converter. The variable phase shifter enables us to vary the timing of the trigger signal generated within the cantilever oscillation cycle. This function is essential for obtaining the optimum phase contrast because the phase contrast markedly depends on the trigger timing, as will be described in Section 17.5.5.

The bandwidth of the fast phase detector reached 1.3 MHz. However, its sensitivity was not as high as that of a lock-in amplifier. The main noise source of the fast phase detector is jitter originating from voltage noise contained in the PSPD output. The RMS value of jitter Δt_{RMS} for a sinusoidal wave containing broadband white noise V_{RMS} is given by

$$\Delta t_{\mathrm{RMS}} = \frac{V_{\mathrm{RMS}}}{2\pi A f_{\mathrm{osc}}} \tag{17.20}$$

Here, A and f_{osc} are the amplitude and frequency of the sinusoidal wave, respectively. The jitter gives rise to timing fluctuations in the trigger pulse for the S/H circuit, and therefore the phase error ΔP_{RMS} (in degrees) is given by

$$\Delta P_{\mathrm{RMS}} = \frac{V_{\mathrm{RMS}}}{2\pi A} \times 360^{\circ} = \frac{360^{\circ}}{2\pi \times \mathrm{SNR}} \tag{17.21}$$

where SNR is the signal-to-noise ratio of the sensor output (i.e., A/V_{RMS}). In the actual experiment, SNR < 10 due to the thermal fluctuations of the cantilever under a low-Q condition. Therefore, the intrinsic phase noise is expected to be larger than 5.7°. However, because the bandwidth of 1.3 MHz is much higher than the imaging bandwidth (approximately 100 kHz), the phase noise can be reduced by using a low-pass filter.

17.5. High-Speed Bioimaging

High-speed AFM imaging studies performed thus far are classified into three categories: early, middle period, and recent studies. In the early stage, imaging studies were carried out to find devices and techniques to be improved. The most serious obstacle we encountered in this stage was that the tip–sample interaction was too strong. Fragile samples such as microtubules and actin filaments were destroyed during imaging. This was due not only to the insufficient feedback speed, but also to the large tapping force exerted by the oscillating tip on the sample. A prototype dynamic PID controller (Ando et al., 2003) and an active damping technique for the z-scanner (Kodera et al., 2005) were developed during this period. Using the dynamic PID controller, the tapping force was greatly reduced and the feedback bandwidth was enhanced. For example, the unidirectional movement of individual kinesin molecules along microtubules was observed without disassembling the microtubules (Ando et al., 2003).

In the middle period, imaging studies were performed to confirm whether biological processes that had been known or predicted to occur would indeed be visualized. For example, the gliding movement of actin filaments over a surface densely coated with myosin V was captured on video (Ando et al., 2005; 2006). However, when the myosin V density was lowered, actin filaments did not appear in the images, suggesting that they were removed by the scanning and oscillating tip. Another example is the stem movement of dynein C in the presence of adenosine triphosphate (ATP). The video image revealed that the stem moved back and forth between two positions while the stalk and head remained stationary (Ando et al., 2005; 2006). The two positions approximately corresponded to the nucleotide-free and adenosine diphosphate (ADP)-vanadate bound states (Burgess et al., 2003). In this period, we developed a method of combining flash photolysis of caged compounds with high-speed AFM. Attenuated high-frequency laser pulses (∼50 kHz, 355 nm) were applied while the y-scan was performed toward the starting point after the completion of one frame acquisition. During this y-scan period, the sample stage was withdrawn from the cantilever tip. This method allowed us to observe the rotational movement of the myosin V head around the head–neck junction that occurred immediately after ultraviolet application to the caged ATP-containing solution (Ando et al., 2006). We also applied this method to observe the height changes in chaperonin GroEL on binding to ATP and GroES (Ando et al., 2005; 2006).

More recently, full-scale imaging studies have been carried out to explore the potential of high-speed AFM, some of which have proved that the current state of our high-speed AFM can be used to reveal functional mechanisms of proteins, although further improvements are still necessary, particularly on the reduction in the tip–sample interaction force. In the following sections, we describe some recent imaging studies.

17.5.1. Chaperonin GroEL

Here, we show that a long-lasting controversial question regarding a biomolecular reaction was quickly solved by directly observing the reaction dynamics. Chaperonin GroEL

consists of 14 identical ATPase subunits that form two heptameric rings stacked back to back (Braig et al., 1994; Xu et al., 1997). A series of biochemical studies (e.g., Burston et al., 1995; Yifrach and Horovitz, 1995) showed that there is positive cooperativity regarding the binding and hydrolysis of ATP in the same ring, whereas there is negative cooperativity between the two rings. Owing to this negative cooperativity, it has been presumed that GroEL binds to GroES at one ring while releasing GroES from the other (Lorimer, 1997; Rye et al., 1997; 1999). This alternate on–off switching appears to exclude the concomitant binding of GroES to both rings of GroEL. However, this issue had remained controversial (e.g., Azem et al., 1994; Grallert and Buchner, 2001) because some electron micrographs showed a complex of GroES-GroEL-GroES with a football shape. To place GroEL on a substratum in a side-on orientation so that both rings are accessible to GroES, we prepared GroEL biotinylated at the equatorial domains (Taguchi et al., 2001). GroEL was attached to streptavidin two-dimensional (2D) crystal sheets prepared on planar lipid bilayers containing a biotin lipid. We were able to capture the GroES alternate association and dissociation at the two GroEL rings. Surprisingly, before the alternate switching took place, a football structure appeared with a high probability (unpublished data). Thus, the high-speed AFM observation clearly solved the controversial issue.

17.5.2. Lattice Defect Diffusion in Two-Dimensional Protein Crystals

For protein crystal formation, the protein–protein association energy must be in an appropriate range. However, the association energy at each contact point had not been assessed experimentally. Here, we show that high-speed AFM imaging can enable its estimation. During protein crystal growth, vacancy point defects are sometimes formed. When their size is small, particularly in the case of monovacancy point defects, they cannot be easily filled with protein molecules floating in the bulk solution. Consequently, they must remain in the crystal. However, in reality, they are mostly removed. Here, we show that high-speed AFM imaging can reveal the defect-removal mechanism.

As mentioned in the previous section, to observe GroEL-GroES association/dissociation dynamics, we prepared streptavidin 2D crystals on supported planar lipid bilayers (Darst et al., 1991). When we observed the crystal surfaces with orthorhombic *C222* symmetry, we noticed that lattice defects sometimes formed at a few places and moved in the crystal (Figure 17.18). This movement is caused by exchanges between the defect (empty) site and one of the surrounding filled sites, because no free streptavidin exists in the bulk solution. To study the vacancy defect mobility and its participation in the crystal growth, monovacancy defects in the streptavidin 2D crystals were systematically produced by increasing the tapping

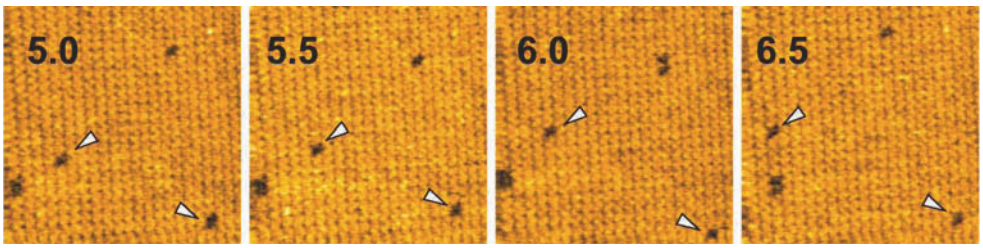


Figure 17.18. Movement of point defects in two-dimensional streptavidin crystal formed on a supported lipid planar bilayer containing biotinylated lipid. The point defects indicated by arrowheads are maintained as point defects during the observation. Scan range, 150 nm; imaging rate, 0.5 sec/frame.

force onto the sample from the oscillating tip. Unexpectedly, the movement of the created monovacancy defects had a preference for one direction over the other (Yamamoto et al., 2008). The movement projected onto each lattice axis showed a time–displacement relationship characteristic of Brownian motion, but the diffusion coefficients differed, depending on the axes.

Streptavidin is a homotetramer with subunits organized in dihedral D_2 symmetry. Therefore, in the 2D crystals, biotin only binds to the two subunits facing the planar lipid bilayers (Figure 17.19a). There are two types of subunit–subunit interactions between adjacent streptavidin molecules: interactions between biotin-bound subunits and interactions between biotin-unbound subunits (Figure 17.19b). One crystallographic axis (*a* axis) is comprised of contiguous biotin-bound subunit pairs, whereas the other axis (*b* axis) is comprised of contiguous biotin-unbound subunit pairs. A previous study on the formation of streptavidin 2D crystal on supported planar lipid bilayers suggested that the interaction between biotin-unbound subunits is stronger than that between biotin-bound subunits (Ku et al., 1993), whereas another study suggested the reverse relationship (Wang et al., 1999). For a streptavidin molecule to move to the adjacent monovacancy defect site along the *a* axis, two biotin-free subunit–subunit contacts and one biotin-bound subunit–subunit contact must be broken. On the other hand, to move along the *b* axis, two biotin-bound subunit–subunit contacts and one biotin-free subunit–subunit contact must be broken. Therefore, the preferential diffusion of the monovacancy defects along the *b* axis indicates that the association between biotin-bound subunits is weaker than that between biotin-free subunits. In addition, from the diffusion constant ratio ($D_b/D_a \sim 2.4$), the difference in the association free energy of the two types of subunit–subunit contacts ($G_{u\text{-}u}$ and $G_{b\text{-}b}$ for the biotin-unbound and biotin-bound pairs, respectively) are quantified to be $G_{u\text{-}u} - G_{b\text{-}b} \sim -0.88 k_B T$ ($T \sim 300$ K), which corresponds to –0.52 kcal/mol (Yamamoto et al., 2008). The aspect ratio of the streptavidin *C222* crystal is known to be ~2 at neutral pH (Yatcilla et al., 1998). Supposing that the aspect ratio of a crystal is proportional to the ratio of the free energies of attractive interactions that occur along the crystal axes [Wolff's rule (Wulff, 1901)], we find that $G_{u\text{-}u}/G_{b\text{-}b}$ is approximately 2. This relationship and $G_{u\text{-}u} - G_{b\text{-}b} \sim -0.88 k_B T$ result in $G_{b\text{-}b} \sim -0.88 k_B T$ and $G_{u\text{-}u} \sim -1.76 k_B T$.

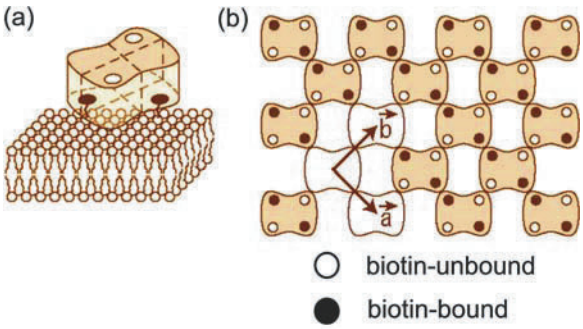


Figure 17.19. **(a)** Schematic of a streptavidin molecule on biotinylated lipid bilayer. Two biotin-binding sites occupied by biotin are indicated by the closed circles. The open circles indicate biotin-binding sites facing the aqueous solution and are biotin-free. **(b)** Schematic of streptavidin arrays in *C222* crystal. Unit lattice vectors are indicated. The *a* axis includes rows of contiguous biotin-bound subunits, whereas the *b* axis includes rows of contiguous biotin-unbound subunits.

The fusion of two point defects into a larger point defect was often observed. On the other hand, the fission of a multivacancy point defect into smaller point defects was rarely observed (Yamamoto et al., 2008). We believe that fission often occurs but cannot easily be observed because the two point defects formed immediately after fission are quickly fused again. During the formation of streptavidin 2D crystals in the presence of free streptavidin in the bulk solution, small point defects such as mono- and divacancy defects do not have easy access to the free streptavidin molecules and hence have a tendency to remain in the crystals (Figure 17.20a). However, the fusion of small point defects into a larger point defect facilitates its access to the free streptavidin molecules and thereby promotes the removal of small point defects from the crystals (Figure 17.20b). Of interest, the defect mobility increases with increasing defect size (Yamamoto et al., 2008). The higher mobility of larger point defects increases their probability to encounter other point defects to form yet larger point defects with further higher mobility. During the crystal growth in the presence of free streptavidin in the bulk solution, this acceleration effect also promotes the removal of point defects from the crystalline regions.

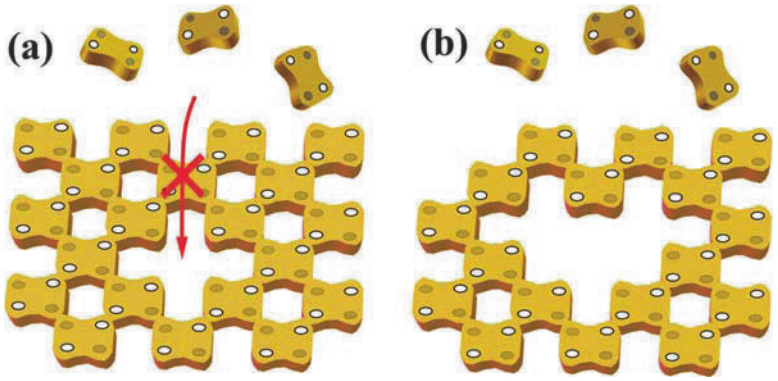


Figure 17.20. Filling vacancy defects in streptavidin two-dimensional crystal with free streptavidin molecules in the bulk solution. Small point defects **(a)** such as mono- and divacancy defects are not easily accessible the free streptavidin molecules, whereas larger point defects **(b)** are easily accessible to them and hence are removed from the crystal.

17.5.3. Myosin V

Here, we show that lively dynamic behavior of a motor protein can be captured by high-speed AFM. Myosin V is a double-headed, actin-based molecular motor that functions as an organelle transporter in various cells and moves processively along actin tracks (Sakamoto et al., 2000). After attaching to an actin filament, it can move on the actin filament over a long distance without dissociation. By single-molecule fluorescence microscopy studies, it has been established that the movement proceeds in 36-nm steps and in a hand-over-hand fashion (Yildiz et al., 2003; Forkey et al., 2003; Warshaw et al., 2005; Syed et al., 2006); the two heads alternate between the leading and trailing positions.

In physiological-ionic-strength solutions, myosin V tended to attach to the mica surface. In higher-ionic-strength solutions, it mostly remained free from the surface, but the affinity of myosin V for actin was lowered. To circumvent this problem, we reduced the cantilever free oscillation amplitude from the usual ~5 nm to ~1 nm, thereby reducing the tip–sample

interaction force but sacrificing the feedback bandwidth. With such small amplitude and in a high-ionic-strength solution containing a low concentration of ATP, the processive movement of myosin V was captured on video at 0.1 sec/frame. The two heads of a myosin V molecule alternated between leading and trailing positions with a walking stride of ~72 nm, that is, in a hand-over-hand manner (unpublished data). High-speed AFM revealed the movement in greater detail than fluorescence microscopy. The lead lever arm bent immediately before the rear head detached from the actin filament. This bent form was similar to that shown previously by electron microscopy (Burgess et al., 2002). The detached rear head rotated around the junction between the two lever arms and then attached to a frontward actin. Immediately after the attachment, the new lead head hopped and sometimes moved forward or backward by 5–10 nm along the actin filament. In the absence of nucleotide or in the presence of ADP, only one head of a myosin V molecule attached to an actin filament. From the orientation of the bound head, it was clear that the bound head was in the trailing position. On the other hand, in the presence of a medium concentration of adenylyl-imidodiphosphate (AMP-PNP), both heads were associated with an actin filament for a long time. This indicates that on binding to AMP-PNP, the lead head conformation adapts itself so that it is able to bind to actin by rotating around the head–neck junction. Therefore, when myosin V is bound to an actin filament at both heads in an ATP-containing solution, the trailing head must contain ADP (or no nucleotide at a low ATP concentration) and the lead head must contain ATP or ADP–inorganic phosphate. After the phosphate is released, the lead head rotates back to the previous orientation, which is unfavorable for actin binding. However, the lead head cannot dissociate from the actin because, in the ADP-bound state, the head has a high affinity for actin. This energetically unfavorable conformation causes the lead neck to bend forward, thereby pulling the trailing head, so that it detaches from the actin and is then conveyed forward. By these mechanical processes along the chemical reaction pathway, the two heads alternate their positions to walk along an actin filament.

17.5.4. Intrinsically Disordered Regions of Proteins

Here, we show that high-speed AFM imaging also is useful for identifying structurally flexible regions of proteins.

Biomolecular binding specificity had been described for a quite while by the famous "lock and key" mechanism, which postulated that a protein must be folded to give a high degree of geometrical precision in molecular binding. This concept has recently being overturned by a surprising yet credible finding that many proteins in the cell appear to be unfolded most of the time. Unstructured segments mostly comprise flexible linkers that play important roles in the assembly of macromolecular complexes and in the recognition of smaller biomolecules. The functional importance of intrinsically disordered (ID) regions has recently been recognized, particularly in transcription, translation, and cellular signal transduction (Demarest et al., 2002; Minezaki et al., 2006). However, there are no useful techniques for analyzing unstructured segments of this sort at the single-molecule level. Indeed, X-ray crystallography and electron microscopy do not allow the direct observation of the ID regions. Conventional slow AFM also cannot reveal such a thin and flexible entity. Under ambient conditions, thin and flexible unstructured polypeptides tend to form lumps or become flattened due to their strong attachment to a surface. Such flattened polypeptides are difficult to visualize by AFM. In addition, they are not easily adsorbed onto substrate surfaces in solution, even with the assistance of adhesive chemicals. Recently, we observed the facilitates chromatin transcription (FACT) protein in solution by high-speed AFM (Miyagi et al., 2008).

FACT is a heterodimer and displaces histone H2A/H2B dimers from nucleosomes, thereby facilitating RNA polymerase II transcription (Belotserkovskaya et al., 2003; Reinberg and Sims, 2006) and chromatin remodeling (Shimojima et al, 2003). The image of oligomerized FACT clearly revealed undulating tail-like structures protruding from the main body of FACT (Figure 17.21). The image of a FACT monomer indicated that FACT contains two tail-like structures of different lengths. This finding coincided with the ID regions predicted from the amino acid sequences. In fact, the AFM images of deletion mutants lacking either of the two predicted ID regions showed either the shorter or the longer tail-like structure. The macroscopic contour lengths of the shorter and longer ID regions showed wide distributions and were, on average, 17.8 and 26.2 nm, respectively. We analyzed mechanical properties (stiffness) of the observed tail-like structures by estimating the macroscopic persistence length and Young's modulus. The persistence length was approximately 11 nm for both tail-like structures. Young's modulus was approximately 9–58 MPa. This Young's modulus is two to

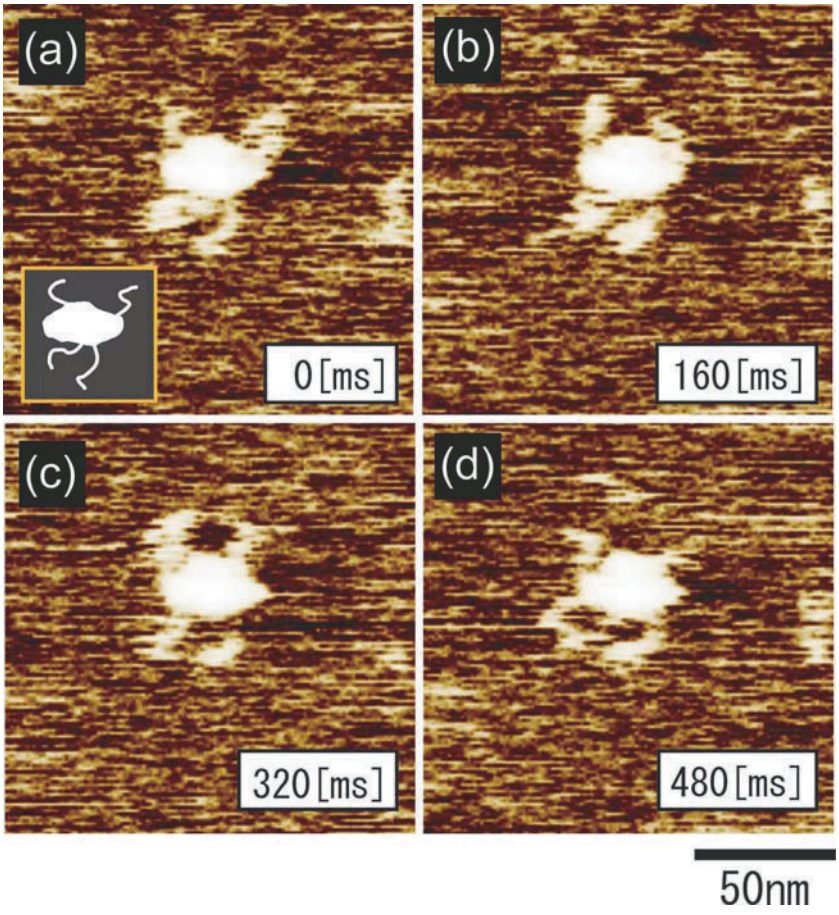


Figure 17.21. Successive atomic force microscope images of oligomerized facilitates chromatin transcription (FACT) protein captured at an imaging rate of 6 frames/sec. The scan range was 150 nm ×150 nm. A lumpy shape and four tail-like structures can be seen (see the schematic inset in panel (**a**). The position of the tail-like structures fluctuated markedly over time. The images were processed using brightness-equalizing software.

three orders of magnitude smaller that those reported for globular proteins but similar to that reported for carbonate anhydrase II incubated in the presence of 2–6 M guanidine HCl (Afrin et al., 2005), suggesting that the structures of the ID regions are similar to those of denatured proteins.

17.5.5. High-Speed Phase-Contrast Imaging

Compositional Mapping on Blended Polymers

To demonstrate that high-speed phase-contrast imaging is possible using small cantilevers and the fast phase detector described in Section 17.4.7, we first imaged poly(styrene-butadiene-styrene) (SBS) block copolymers as a test sample because SBS films are often used for phase imaging by tapping-mode AFM. Figure 17.22 shows typical (a) topographic and (b) phase images obtained simultaneously at a rate of 84 ms/frame in distilled water. Here, the phase detection timing was set at a regime in which the tip approached the surface. Clear phase-contrast images were obtained even at such a high imaging rate. In the phase-contrast images shown in Figures 17.22b, d, darker regions correspond to an advanced phase with an average shift of approximately +6° relative to the phase at brighter regions. Because the image contrast in the error signal is faint, the phase contrast predominantly reflects the compositional heterogeneity of the SBS film. To identify the compositions of different regions in the topographic image, the peak-to-peak set-point amplitude was reduced from 11 nm to 7 nm during imaging. Figures 17.22c, d show topographic and phase-contrast images obtained with the reduced set-point amplitude, respectively. The higher topographic region shown in Figure 17.22a was dented by the increased loading force, as shown in Figure 17.22c. The

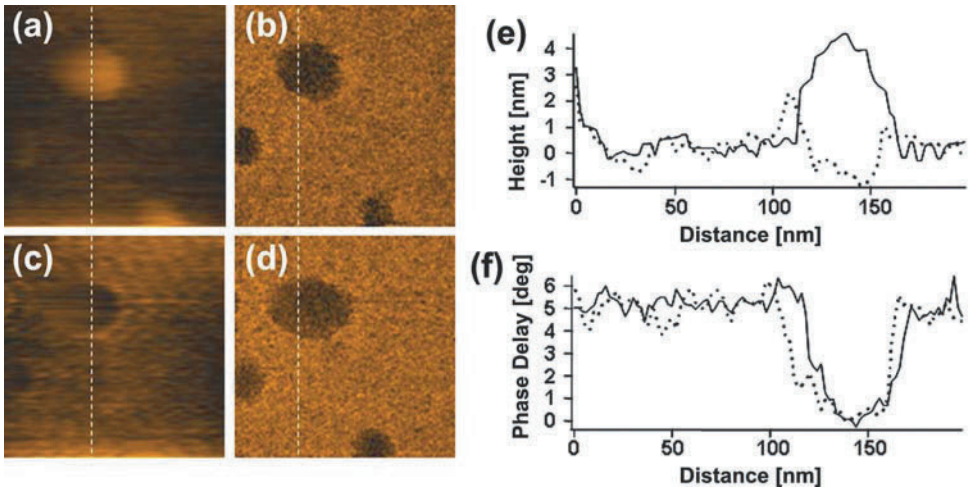


Figure 17.22. (**a, c**) Topographic and (**b, d**) phase-contrast images of a poly(styrene-butadiene-styrene) film captured at 84 ms/frame in pure water. The scan area is 200 × 200 nm^2 with 100 × 100 pixels. (**a**), (**b**). The amplitude set-point r is 0.85, and the peak-to-peak free amplitude is 13 nm. (**c**), (**d**). The amplitude set-point is reduced to 0.54. Solid and broken lines correspond to the cross-sectional profiles (**e**), (**f**) obtained before and after reducing the amplitude set-point, respectively. Broken lines in panels a–d indicate the positions at which the cross sections were obtained.

cross-section height profiles shown in Figure 17.22e reveal that it was dented by more than 5 nm. Thus, the compositions of the higher and lower regions in Figure 17.22a are identified to be the less stiff polybutadiene (PB) and the stiff polystyrene (PS) domains, respectively. The darker (advanced phase) area in the phase-contrast image shown in Figure 17.22d is wider than that in Figure 17.22b. This is because the contact area between the tip and the surface was increased by the increased loading force. However, the magnitude of the phase shift is independent of the loading force, as can be seen in the cross-sectional phase profiles shown in Figure 17.22f. Therefore, the observed phase contrast is caused by only compositional heterogeneities and is independent of the surface topography.

We found that phase contrast largely depends on the phase-detection timing within a cantilever oscillation cycle (Uchihashi et al., 2006). A large relative phase advance was observed in the PB region when the cantilever tip approached the surface. On the other hand, only a faint phase contrast was observed when the cantilever tip was withdrawing from the surface. After analyzing this detection-timing dependence, we reached the following conclusion. A phase advance produced by energy-conservative repulsive interaction decays within one oscillating cycle due to the low Q factor of the cantilever, whereas a phase delay produced by the adhesive force does not decay quickly within a single oscillation cycle because of the energy-dissipative interaction. For more details, see the literature (Uchihashi et al., 2006; Ando et al., 2008b).

Phase-Contrast Imaging of Myosin Filaments

To demonstrate the performance of the simultaneous topographic and phase-contrast imaging of biological systems, here we show images obtained using myosin filaments in a buffer solution. Figures 17.23a, b show typical topographic and phase-contrast images, respectively, obtained at an imaging rate of 4 sec/frame. The scan area was 2 μm $\times$ 2 μm, which was relatively large for high-speed imaging, and therefore the imaging was performed at relatively slow speed. The surfaces of the myosin filaments appear smooth in the topographic image. On the other hand, the phase-contrast image shows the structures of the filaments in more detail. Here, the trigger timing was tuned so that the maximum phase contrast was

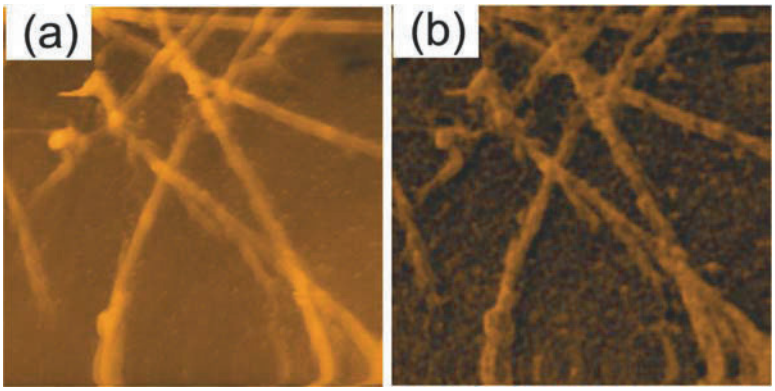


Figure 17.23. (**a**), (**b**) Topographic and phase-contrast images, respectively, of myosin filaments captured at 4 sec/frame in buffer solution. The scan area is 2 $\times$ 2 μm^2 with 256 $\times$ 256 pixels.

obtained. Note that the phase on the myosin filaments is delayed relative to that on the mica surface. This is due to a large difference in the mechanical properties between the biological sample and the mica. The phase contrast on the myosin filaments indicates that the subtle variations in the mechanical properties such as elasticity and viscoelasticity exist in the myosin filaments.

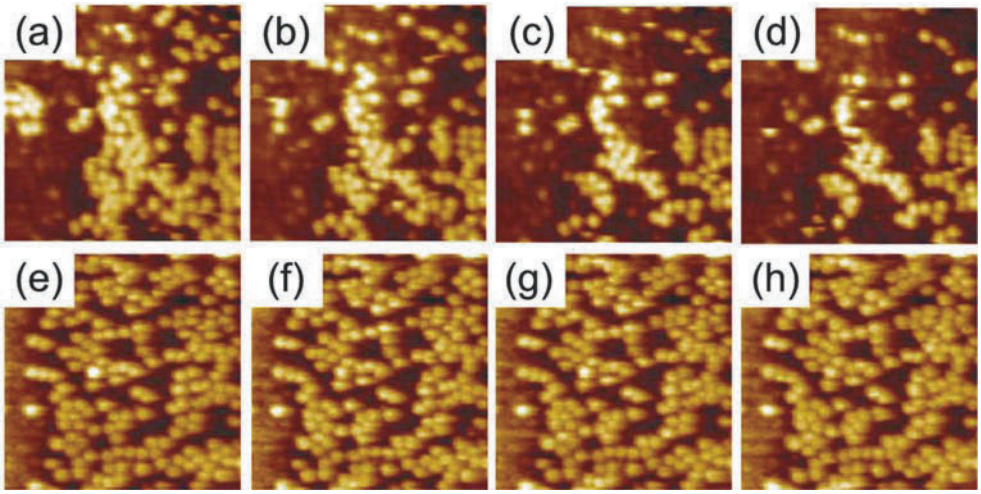


Figure 17.24. Successive atomic force microscope images of GroEL taken in (**a–d**) the amplitude modulation mode and (**e–h**) the phase modulation mode. Scan area is 400 × 400 nm^2. Imaging rate is 800 ms/frame.

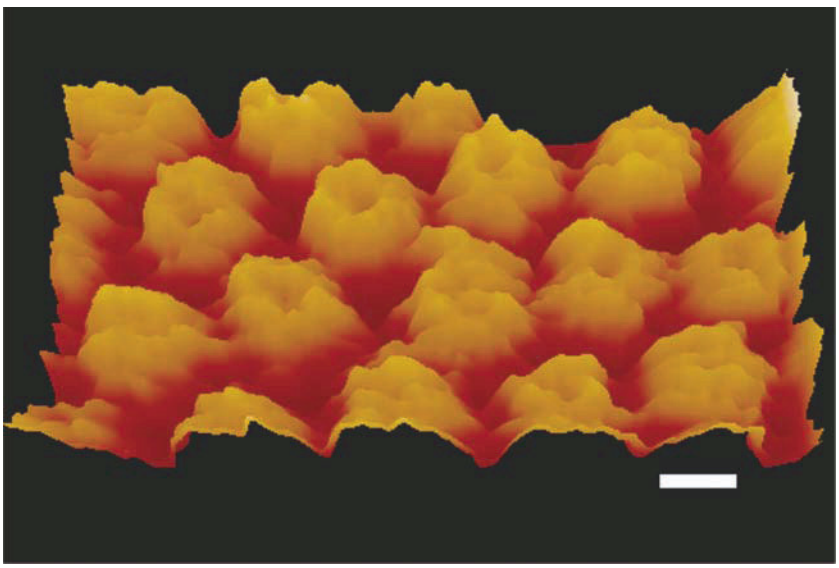


Figure 17.25. High-resolution image of GroEL taken in the phase modulation mode. Scale bar, 10 nm; imaging rate, 480 ms/frame.

Phase-Contrast Imaging of GroEL

Chaperonin GroEL was used as a test sample to show that the phase shift is more sensitive to the tip-sample interaction than the amplitude change, as its stacked ring structure is fragile and is often bisected by strong tip forces. Figures 17.24a–d show successive topographic images taken in the AM mode at 800 ms/frame. The set-point amplitude was approximately 90% of the peak-to-peak free amplitude of 7.5 nm. In these images, a large fraction of GroEL molecules are bisected and are sometimes dislodged by the scanning tip. On the other hand, in the PM mode (phase signal is used for feedback control of the tip–sample distance), the bisection of GroEL rings rarely occurred, as shown in Figures 17.24e–h. In addition, the hole centered on the ring structure is more clearly observed in the PM mode than in the AM mode, and the subunits are resolved in the PM mode (Figure 17.25).

17.6. Substrata for Observing Dynamic Biomolecular Processes

To visualize dynamic biomolecular processes, the substratum surfaces on which a sample is placed should meet various requirements; for example, (1) the surface must have an appropriate association affinity for the sample so that the sample does not move too rapidly on it, (2) the functional activity of the sample should be retained on the surface, (3) the surface should selectively attach a specific component in a multicomponent sample, and (4) the sample should attach to the surface in a desired orientation. Mica (natural muscovite or synthetic fluorophlogopite) has frequently been used as a source substratum owing to its surface flatness at the atomic level over a large area. It has net negative charge and is therefore quite hydrophilic. A bare mica surface adsorbs various proteins by electrostatic interactions. Except in some cases (such as GroEL attachment in an end-up orientation), the orientation of adsorbed proteins is not unique, and the selective attachment of a specific species is not expected. When the dynamic process of a single species of protein is to be observed without other proteins, a mica surface is useful. We can control the affinity for a specific protein by varying the ionic strength or pH or by varying the concentration of divalent cations such as Mg^{2+}. However, interesting biomolecular processes often occur in multicomponent systems.

It is desired to have, in the first step, a surface that does not initially attach to the sample. Then, this surface is modified so that only a specific species of molecules can attach to it via a linker in a desired orientation. A membrane surface with zwitterionic polar head groups such as phosphatidyl choline (PC) and phosphatidyl ethanolamine (PE) is known to resist protein adsorption (Zhang et al., 1998; Vadgama, 2005). Streptavidin is also useful for the specific attachment of a biotinylated protein because it also resists nonspecific protein adsorption. Mica-surface-supported planar lipid bilayers (Sackmann, 1996) can be easily formed from liposomes because their structure is disrupted by the strong hydrophilic interaction with the mica surface (Reviakine and Brisson, 2000). Various lipids with functional groups that are attached to polar groups [e.g., biotin attached to PE, Ni-nitrilotriacetic acid attached to phosphatidylserine] are commercially available. They enable the specific attachment of proteins labeled with biotin or his-tags onto planar lipid bilayers.

Dioleoyl-phosphatidyl-choline (DOPC) is useful for the preparation of 2D streptavidin crystals when it is used together with biotinylated lipids (Scheuring et al., 1999; Reviakine and Brisson, 2001). It contains an unsaturated hydrocarbon in each of the two alkyl chains, which causes the bending of the chains that weakens the interaction between neighboring DOPCs. This weak interaction lowers the phase transition temperature of DOPC lipid to

about −20°C, thereby affording considerable fluidity to the planar bilayer at room temperature and thus facilitating the 2D crystal formation of streptavidin (Figure 17.26). The densely packed streptavidin does not diffuse easily. If less diffusibility is necessary, the packed streptavidin can be cross-linked using glutaraldehyde, which does not affect its ability to bind to biotin.

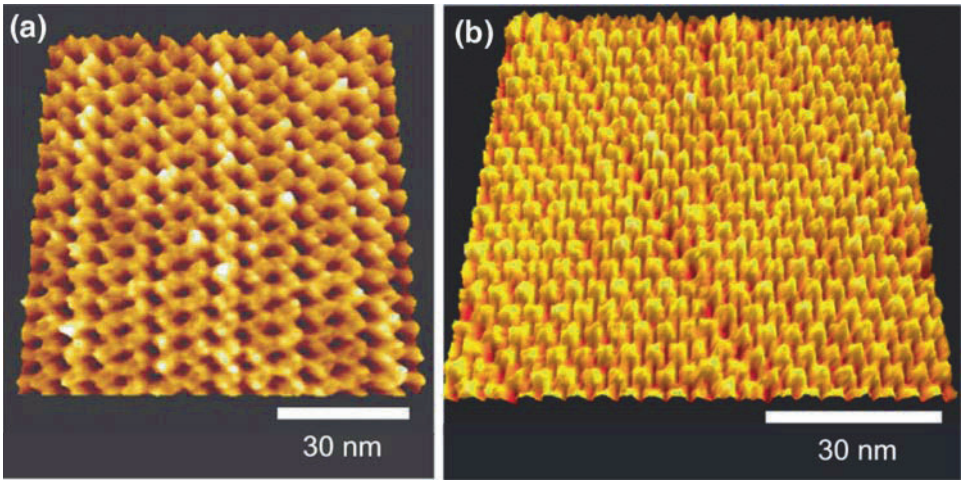


Figure 17.26. AFM images of streptavidin 2D crystals with (**a**) P2 (**b**) C222 symmetries formed at pH 5.8 and 7.4, respectively, on supported planar lipid bilayers containing biotin-lipid. The lipid compositions are DOPC, DOPS and 1,2-dioleoyl-sn-glycero-3-phosphoethanol-amine-N-(cap-biotinyl) at 7:2:1 molar ratio.

Dipalmitoyl-phosphatidyl-choline (DPPC) contains no unsaturated hydrocarbons in the alkyl chains; therefore, its phase transition temperature is high (∼41°C), and it is suitable for preparing planar lipid bilayers with low fluidity. For example, when planar bilayers are formed using DPPC at a high temperature (∼60°C) together with a suitable fraction of DPPE-biotin, streptavidin, which is sparsely attached to the surface, negligibly diffuses at room temperature (Ando et al., 2008a). We can anchor a biotinylated protein sparsely on this lipid bilayer surface and observe the dynamic interaction of the anchored protein with a floating counterpart protein. When the counterpart protein contains a flexible structure, it exhibits very rapid Brownian motion even when it is in contact with the anchored protein. Because such a rapidly moving structure cannot be imaged even by high-speed AFM, it is necessary to find appropriate lipid polar groups to slow the motion. In addition, the anchored protein may rapidly rotate around the biotin–streptavidin connection, which often hampers the observation of the protein–protein interaction events. To solve this problem, the protein to be anchored must be biotinylated at two appropriate sites, or the lipid polar groups should have an appropriate affinity for the anchored protein.

References

Afrin R., M. T. Alam, and A. Ikai, Pretransition and progressive softening of bovine carbonic anhydrase II as probed by single molecule atomic force microscopy, Protein Sci. 14:1447–1457 (2005).

Albrecht T. R., P. Grütter, D. Horne, and D. Rugar, Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity, J. Appl. Phys. 69:668–673 (1991).

Anczykowski B., J. P. Cleveland, D. Krüger, V. Elings, and H. Fuchs, Analysis of the interaction mechanisms in dynamic mode SFM by means of experimental data and computer simulation, Appl. Phys. A 66:S885–S889 (1998).

Ando T., N. Kodera, E. Takai, D. Maruyama, K. Saito, and A. Toda, A high-speed atomic force microscope for studying biological macromolecules, Proc. Natl. Acad. Sci. USA 98:12468–12472 (2001).

Ando T., N. Kodera, D. Maruyama, E. Takai, K. Saito, and A. Toda, A high-speed atomic force microscope for studying biological macromolecules in action, Jpn. J. Appl. Phys. 41:4851–4856 (2002).

Ando T., N. Kodera, Y. Naito, T. Kinoshita, K. Furuta, and Y. Y. Toyoshima, A High-speed atomic force microscope for studying biological macromolecules in action, Chem. Phys. Chem. 4:1196–1202 (2003).

Ando T., T. Uchihashi, N. Kodera, A. Miyagi, R. Nakakita, H. Yamashita, and K. Matada, High-speed atomic force microscopy for capturing the dynamic behavior of protein molecules at work, Surf. Sci. Nanotechnol. 3:384–392 (2005).

Ando T., T. Uchihashi, N. Kodera, A. Miyagi, R. Nakakita, H. Yamashita, and M. Sakashita, High-speed atomic force microscopy for studying the dynamic behavior of protein molecules at work, Jpn. J. Appl. Phys. 45:1897–1903 (2006).

Ando T., T. Uchihashi, N. Kodera, D. Yamamoto, M. Taniguchi, A. Miyagi, and H. Yamashita, High-speed atomic force microscopy for observing dynamic biomolecular processes, J. Mol. Recognit. 20:448–458 (2007).

Ando T., T. Uchihashi, N. Kodera, D. Yamamoto, A. Miyagi, M. Taniguchi, and H. Yamashita, High-speed AFM and nano-visualization of biomolecular processes, Pflugers Arch. Eur. J. Physiol. 456:211–225 (2008a).

Ando T., T. Uchihashi, and T. Fukuma, High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes, Prog. Surf. Sci. 83:337–437 (2008b).

Azem A., M. Kessel, and P. Goloubinoff, Characterization of a functional GroEL14(GroES7)2 chaperonin hetero-oligomer, Science 265:653–656 (1994).

Bar G., Y. Thomann, R. Brandsch, H.-J. Cantow, and M.-H. Whangbo, Factors affecting the height and phase images in tapping mode atomic force microscopy. Study of phase-separated polymer blends of poly(ethene-co-styrene) and poly (2,6-dimethyl- 1,4-phenylene oxide), Langmuir 13:3807–3812 (1997).

Belotserkovskaya R., S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg, FACT facilitates transcription-dependent nucleosome alteration, Science 301:1090–1093 (2003).

Burgess S. A., M. L. Walker, F. Wang, J. P. Sellers, H. D. White, P. J. Knight, and J. Trinick, The prepower stroke conformation of myosin V, J. Cell Biol. 159:983–991 (2002).

Burgess S. A., M. L. Walker, H. Sakakibara, P. J. Knight, and K. Oiwa, Dynein structure and power stroke, Nature 421:715–718 (2003).

Burston S. G., N. A. Ranson, and A. R. Clarke, The origins and consequences of asymmetry in the chaperonin reaction cycle, J. Mol. Biol. 249:138–152 (1995).

Braig K., Z. Otwinowski, R. Hegde, D. C. Boisvert, A. Joachimiak, A. L. Horwich, and P. B. Sigler, The crystal structure of the bacterial chaperonin GroEL at 2.8 Å, Nature 371:578–586 (1994).

Chang W.-J., J.-C. Hsu, and T.-H. Lai, Inverse calculation of the tip–sample interaction force in atomic force microscopy by the conjugate gradient method, J. Phys. D Appl. Phys. 37:1123–1126 (2004).

Cleveland J. P., B. Anczykowski, A. E. Schmid, and V. B. Elings, Energy dissipation in tapping-mode atomic force microscopy, Appl. Phys. Lett. 72:2613–2615 (1998).

Czajkowsky D. M., M. J. Allen, V. Elings, and Z. Shao, Direct visualization of surface charge in aqueous solution, Ultramicroscopy 74:1–5 (1998).

Darst S. A., M. Ahlers, P. H. Meller, E. W. Kubalek, R. Blankenburg, H. O. Ribi, H. Ringsdorf, and R. D. Kornberg, Two-dimensional crystals of streptavidin on biotinylated lipid layers and their interactions with biotinylated macromolecules, Biophys. J. 59:387–396 (1991).

Demarest S. J., M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H. J. Dyson, R. M. Evans, and P. E. Wright, Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators, Nature 415: 549–553 (2002).

Forkey J. N., M. E. Quinlan, M. A. Shaw, J. E. T. Corrie, and Y. E. Goldman, Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization, Nature, 422:399–404 (2003).

Fukuma T., M. Kimura, K. Kobayashi, K. Matsushige, and H. Yamada, Development of low noise cantilever deflection sensor for multienvironment frequency-modulation atomic force microscopy, Rev. Sci. Instrum. 76:053704 (2005).

Fukuma T., Y. Okazaki, N. Kodera, T. Uchihashi, and T. Ando, High resonance frequency force microscope scanner using inertia balance support, Appl. Phys. Lett. 92:243119 (2008).

Gao S., L. F. Chi, S. Lenhert, B. Anczykowski, C. M. Niemeyer, M. Adler, and H. Fuchs, High quality mapping of DNA-protein complexes by dynamic scanning force microscopy, Chem. Phys. Chem. 6:384–388 (2001).

Giessibl F.J., Atomic resolution of the silicon (111)-(7 × 7) surface by atomic force microscopy, Science 267:68–71 (1995).

Grallert H. and J. Buchner, Review: A structural view of the GroE chaperone cycle, J. Struct. Biol. 135:95–103 (2001).

Hung S. K., E.-T. Hwu, I.-S. Hwang, and L.-C. Fu, Postfitting control scheme for periodic piezoscanner driving, Jpn. J. Appl. Phys. 45B:1917–1921 (2006).

Kindt J. H., G. E. Fantner, J. A. Cutroni, and P. K. Hansma, Rigid design of fast scanning probe microscopes using finite element analysis, Ultramicroscopy 100:259–265 (2004).

Kitazawa M., K. Shiotani, and A. Toda, Batch fabrication of sharpened silicon nitride tips, Jpn. J. Appl. Phys. (Pt. 1) 42:4844–4847 (2003).

Kodera N., H. Yamashita, and T. Ando, Active damping of the scanner for high-speed atomic force microscopy, Rev. Sci. Instrum. 76:053708 (2005).

Kodera N., M. Sakashita, and T. Ando, Dynamic proportional-integral-differential controller for high-speed atomic force microscopy, Rev. Sci. Instrum. 77:083704 (2006).

Kokavecz J., Z. Tóth, Z. L. Horváth, P. Heszler, and Á. Mechler, Novel amplitude and frequency demodulation algorithm for a virtual dynamic atomic force microscope, Nanotechnology 17:S173–S177 (2006).

Ku A. C., S. A. Darst, C. R. Robertson, A. P. Gast, and R. D. Kornberg, Molecular analysis of two-dimensional protein crystallization, J. Phys. Chem. 97:3013–3016 (1993).

Legleiter J., M. Park, B. Cusick, and T. Kowalewski, Scanning probe acceleration microscopy (SPAM) in fluids: Mapping mechanical properties of surfaces at the nanoscale, Proc. Natl. Acad. Sci. USA 103:4813–4818 (2006).

Lorimer G. H., Protein folding. Folding with a two-stroke motor, Nature 388:720–722 (1997).

Minezaki Y., K. Homma, A. R. Kinjo, and K. Nishikawa, Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation, J. Mol. Biol. 359:1137–1149 (2006).

Miyagi A., Y. Tsunaka, T. Uchihashi, K. Mayanagi, S. Hirose, K. Morikawa, and T. Ando, Visualization of intrinsically disordered regions of proteins by high-speed atomic force microscopy, Chem. Phys. Chem. 9:1859–1866 (2008).

Reinberg D. and R. J. Sims III, de FACTo nucleosome dynamics, J. Biol. Chem. 281:23297–23301 (2006).

Reviakine I. and A. Brisson, Formation of supported phospholipid bilayers from unilamellar vesicles investigated by atomic force microscopy, Langmuir 16:1806–1815 (2000).

Reviakine I. and A. Brisson, Streptavidin 2D crystals on supported phospholipid bilayers: Toward constructing anchored phospholipid bilayers, Langmuir 17:8293–8299 (2001).

Rye H. S., S. G. Burston, W. A. Fenton, J. M. Beechem, Z. Xu, P. B. Sigler, and A. L. Horwich, Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL, Nature 388:792–798 (1997).

Rye H. S., A. M. Roseman, S. Chen, K. Furtak, W. A. Fenton, H. R. Saibil, and A. L. Horwich, GroEL–GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. Cell 97:325–338 (1999).

Sackmann E., Supported membranes: Scientific and practical applications, Science 271:43–48 (1996).

Sahin O., Harnessing bifurcations in tapping-mode atomic force microscopy to calibrate time-varying tip–sample force measurements, Rev. Sci. Instrum. 78:103707 (2007).

Sahin O., S. Magonov, C. Su, C. F. Quate, and O. Solgaard, An atomic force microscope tip designed to measure time-varying nanomechanical forces, Nat Nanotechnol. 2:507–514 (2007).

Sakamoto T., I. Amitani, E. Yokota, and T. Ando, Direct observation of processive movement by individual myosin V molecules, Biochem. Biophys. Res. Commun. 272:586–590 (2000).

Scheuring S., D. J. Müller, P. Ringler, J. B. Heymann, and A. Engel, Imaging streptavidin 2D crystals on biotinylated lipid monolayers at high resolution with the atomic force microscope, J. Microsc 193:28–35 (1999).

Schiener J., S. Witt, M. Stark, and R. Guckenberger, Stabilized atomic force microscopy imaging in liquids using second harmonic of cantilever motion for setpoint control, Rev. Sci. Instrum. 75:2564–2568 (2004).

Schitter G., F. Allgöwer, and A Stemmer, A new control strategy for high-speed atomic force microscopy, Nanotechnology 15:108–114 (2004).

Shimojima T., M. Okada, T. Nakayama, H. Ueda, K. Okawa, A. Iwamatsu, H. Handa, and S. Hirose, *Drosophila* FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor, Genes Dev. 17:1605–1616 (2003).

Stark M. and R. Guckenberger, Fast low-cost phase detection setup for tapping-mode atomic force microscopy, Rev. Sci. Instrum. 70:3614–3619 (1999).

Stark M., R. W. Stark, W. M. Heck, and R. Guckenberger, Inverting dynamic force microscopy: From signals to time-resolved interaction forces, Proc. Natl. Acad. Sci. USA 99:8473–8478 (2002).

Syed S., G. E. Snyder, C. Franzini-Armstrong, P. R. Selvin, and Y. E. Goldman, Adaptability of myosin V studied by simultaneous detection of position and orientation, EMBO J. 25:1795–1803 (2006).

Taguchi H., T. Ueno, H. Tadakuma, M. Yoshida, and T. Funatsu, Single-molecule observation of protein–protein interactions in the chaperonin system, Nat Biotechnol. 19:861–865 (2001).

Tamayo J., and R. García, Deformation, contact time, and phase contrast in tapping mode scanning force microscopy, Langmuir 12:4430–4435 (1996).

Tamayo J., A. D. L. Humphris, R. J. Owen, and M. J. Miles, High-Q dynamic force microscopy in liquid and its application to living cells, Biophys. J. 81:526–537 (2001).

Uchihashi T., T. Ando, and H. Yamashita, Fast phase imaging in liquids using a rapid scan atomic force microscope, Appl. Phys. Lett. 89:213112 (2006).

Vadgama P., Surface biocompatibility, Annu. Rep. Prog. Chem. C Phys. Chem. 101:14–52 (2005).

Wang S.-W., C. R. Robertson, and A. P. Gast, Molecular arrangement in two-dimensional streptavidin crystals, Langmuir 15:1541–1548 (1999).

Warshaw D. M., G. G. Kennedy, S. S. Work, E. B. Krementsova, S. Beck, and K. M. Trybus, Differential labeling of myosin V heads with quantum dots allows direct visualization of hand-over-hand processivity, Biophys. J. 88:L30–L32 (2005).

Wulff G., Zur frage der geschwindigkeit des wachstums und der auflösung der kristallflächen *Z. Kristallogr.* 34:449–530 (1901).

Xu Z., A. L. Horwich, and P. B. Sigler, The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex, Nature 388:741–750 (1997).

Yamamoto D., T. Uchihashi, N. Kodera, and T. Ando, Anisotropic diffusion of point defects in two-dimensional crystal of streptavidin observed by high-speed atomic force microscopy, Nanotechnology 19:384009 (2008).

Yatcilla M. T., C. R. Robertson, and A. P. Gast, Influence of pH on two-dimensional streptavidin crystals, *Langmuir* 14:497–503 (1998).

Yifrach O. and A. L. Horovitz, Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL, Biochemistry 34:5303–5308 (1995).

Yildiz A., J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization, Science 300:2061–2065 (2003).

Zhang S. F., P. Rolfe, G. Wright, W. Lian, A.J . Milling, S. Tanaka, and K. Ishihara, Physical and biological properties of compound membranes incorporating a copolymer with a phosphorylcholine head group, Biomaterials 19:691–700 (1998).

Zhong Q., D. Inniss, K. Kjoller, and V. B. Elings, Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy, Surf. Sci. 290:L688–L692 (1993).

18

Recognition Imaging Using Atomic Force Microscopy

Andreas Ebner, Lilia Chtcheglova, Jilin Tang, David Alsteens, Vincent Dupres, Yves F. Dufrêne, and Peter Hinterdorfer

Abstract Recognition imaging using atomic force microscopy (AFM) offers a wealth of new opportunities in biophysical research, such as its ability to localize specific chemical groups and biological receptors on biosurfaces and to measure their molecular-scale interactions. By attaching well-defined chemical groups on tips, it is possible to map chemical properties and interactions on cell surfaces on a scale of a few functional groups. Single-molecule force spectroscopy with tips functionalized with relevant bioligands provides a means of localizing individual receptors and measuring their specific binding forces. Alternatively, recognition sites may also be mapped with unprecedented temporal resolution using dynamic recognition imaging, in which molecular recognition signals are detected during dynamic force microscopy imaging. These AFM modalities, which all have functionalization of the tips with specific molecules in common, provide new avenues for understanding the structure–function of cell surfaces in connection with medical and physiological issues.

18.1. Introduction

Molecular recognition involving the specific interaction between receptors and their cognitive ligands plays a pivotal role in life sciences. Signalling cascades, enzymatic activity, genome replication and transcription, cohesion of cellular structures, interaction of antigens and antibodies, and metabolic pathways are just some examples of processes which critically rely on specific recognition. In addition to these specific interactions, nonspecific intermolecular interactions, such as hydrophobic and electrostatic forces, also play essential roles in biological events, such as protein folding and cell adhesion. Force-measuring techniques such as the surface forces apparatus and the optical and magnetic tweezers have been developed to measure these noncovalent intermolecular forces. Although very powerful, these approaches are limited by their poor lateral resolution, meaning that they cannot resolve the spatial distribution of (bio)chemical sites or their interactions. Accordingly, localizing specific chemical

A. Ebner, L. Chtcheglova, J. Tang, and P. Hinterdorfer • Institute for Biophysics, Johannes Kepler University of Linz, Altenbergerstr.69, A-4040 Linz, Austria
D. Alsteens, V. Dupres, and Y. F. Dufrêne • Unité de chimie des interfaces, Université catholique de Louvain, Croix du Sud 2/18, B-1348 Louvain-la-Neuve, Belgium

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_18,

groups and receptor sites on biological samples and measuring their molecular-scale interactions represents a major challenge in current biophysical research.

In recent years, atomic force microscopy (AFM) has become a powerful tool for probing biosurfaces on the nanoscale [1–5]. Whereas AFM imaging provides three-dimensional views of specimens with (sub)molecular resolution and with minimal sample preparation, AFM force spectroscopy measures the forces within or between single molecules, thereby providing new insights into the molecular basis of crucial biological events such as protein folding and cell adhesion. The AFM works by scanning a sharp tip over the surface of the sample while sensing near-field physical interactions between the tip and the sample. This allows three-dimensional images to be directly generated, even in aqueous solution. The sample is mounted on a piezoelectric scanner that ensures three-dimensional positioning with high accuracy. While the tip (or sample) is being scanned in the (x, y) directions, the force interacting between tip and specimen is monitored with a sensitivity on the order of piconewtons. This force is measured by the deflection of a soft cantilever that is detected by a laser beam focused on the free end of the cantilever and reflected into a photodiode. AFM cantilevers and tips are typically made of silicon or silicon nitride using microfabrication techniques. The most widely used imaging mode is contact mode, in which sample topography can be measured in two ways. In constant-height mode, the cantilever deflection is recorded while the sample is scanned at constant height. It is often necessary to minimize large deflections to reduce the forces applied to the sample and thus prevent damage. This is achieved in the constant-force mode, in which the sample height is adjusted to keep the deflection of the cantilever constant; thus, the force applied to the tip is maintained at a constant level using a feedback loop. In dynamic or intermittent mode, also known as tapping mode, a tip is oscillated near its resonance frequency and scanned over the surface. The amplitude and phase of the cantilever are monitored and taken as signals for imaging.

In force spectroscopy modalities, such as chemical force microscopy (CFM) and single-molecule force spectroscopy (SMFS), the cantilever deflection is recorded as a function of the vertical displacement of the scanner, that is, as the sample is pushed toward the tip and retracted. This results in a cantilever deflection versus scanner displacement curve, which can be transformed into a force–distance curve by using the appropriate corrections. To gain accurate knowledge about the measured forces, it is important to experimentally determine cantilever spring constants because they may substantially differ from values quoted by the manufacturer [6]. The characteristic adhesion (or unbinding) force between tip and sample observed during retraction may then be used to detect chemical groups (CFM) and receptor sites (SMFS). A few years ago, a novel molecular recognition imaging mode was introduced in which molecular recognition signals were detected during dynamic force microscopy imaging (topography and recognition imaging). This chapter focuses on available molecular recognition AFM modalities, emphasizing methodologies and applications in biophysical research.

18.2. Chemical Force Microscopy

Noncovalent interactions, such as hydrophobic and electrostatic interactions, play essential roles in nature because they mediate crucial events such as protein folding and cellular interactions. For the first time, CFM enables researchers to measure these interactions on biosurfaces with nanoscale lateral resolution.

18.2.1. Methods

The principle of CFM, which was introduced in 1994, is to use AFM tips with well-defined chemistry for measuring adhesion (or friction) and for imaging surfaces [7,8]. The surface chemistry of commercial tips is poorly controlled and often contaminated with gold and other materials. Therefore, reliable CFM measurements require functionalizing tips with organic monolayers terminated by specific functional groups (e.g., OH or CH_3). A common method is based on the formation of self-assembled monolayers (SAMs) of alkanethiols on gold surfaces [9]. This procedure involves coating microfabricated cantilevers with a thin adhesive layer (Cr or Ti) via thermal evaporation, followed by a 15- to 100-nm-thick Au layer. The coated cantilevers are immersed in dilute (0.1–1 mM) ethanol solutions of the selected alkanethiol, followed by a rinse with ethanol, and dried using a gentle nitrogen flow. Although the protocol is fairly simple, it is important to validate the quality of the surface modification, which can be accomplished by treating model supports (glass, mica) in the same way as the tips and by characterizing them by means of surface analysis techniques (e.g., contact-angle measurements or X-ray photoelectron spectroscopy). To minimize surface contamination and alteration it is important to use the functionalized tips immediately after they are prepared.

A CFM experiment typically involves measuring the adhesion strength between chemical groups via force spectroscopy. The cantilever deflection is recorded as a function of the vertical displacement of the piezoelectric scanner, yielding a raw "voltage–displacement" curve that can be converted into a "force–distance" curve. Using the slope of the curves in the region where tip and sample are in contact, one can convert the voltage into a cantilever deflection. The cantilever deflection is then converted into a force F using Hooke's law: $F = -k \times d$, where k is the cantilever spring constant. The curve can be corrected by plotting F as a function of $(z - d)$. The zero separation distance is then determined as the position of the vertical linear parts of the curve in the contact region. The hysteresis or "pull-off" force observed during retraction is used to estimate the adhesion (unbinding) force between tip and sample. Adhesion force maps can be obtained by recording spatially resolved force–distance curves in the (x, y) plane.

Although hydrophobic forces have been known for 70 years and are of prime importance in biology (e.g., in protein folding and aggregation), their detailed mechanisms remain poorly understood. Recently, CFM was shown to be a powerful tool for measuring hydrophobic forces at biological interfaces and for resolving their nanoscale distribution. Figure 18.1 shows the proof of such CFM measurements on model surfaces [10]. SAMs of CH_3- and OH-terminated alkanethiols mixed in different proportions were probed using water contact-angle measurements and CFM with hydrophobic, CH_3-terminated tips. Consistent with expectations (work of adhesion, Young equation), the contact angle and adhesion force values measured on mixed SAMs increased gradually with the molar fraction of CH_3-terminated alkanethiols (Figure 18.1A, B), yielding a linear relationship between the adhesion force and the cosine of the contact angle (Figure 18.1C). This excellent agreement demonstrates that the measured adhesion forces reflect surface hydrophobicity. We point out that, interpretation of the data in terms of interfacial thermodynamics indicated that the measured adhesion forces do not originate from true, direct tip–sample interactions but rather reflect entropy changes associated with the restructuration of water near hydrophobic surfaces. This work is also of practical interest as it shows that CFM can be used for quality control of chemically modified tips, which is not possible when using common analytical techniques. As we discuss later, further interpretation of cellular data requires expressing the measured adhesion

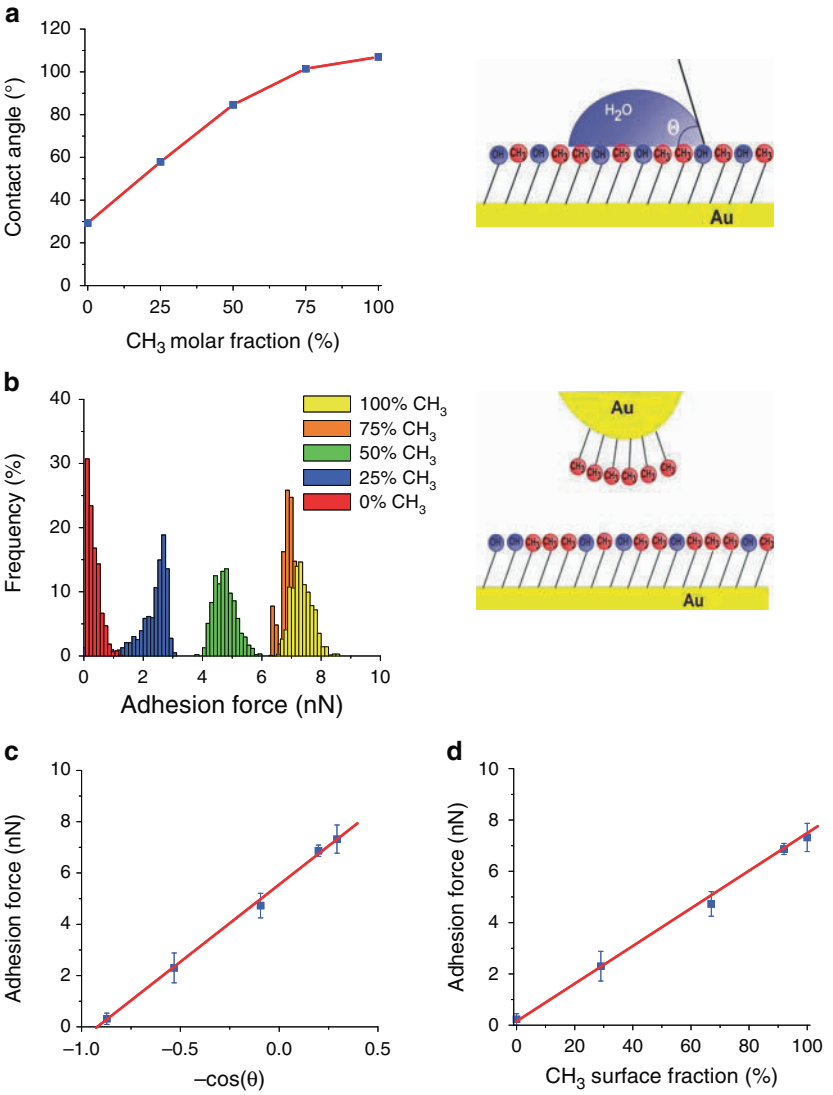


Figure 18.1. Measuring hydrophobic forces using chemical force microscopy. **A.** Water contact angle (θ) values measured for self-assembled mixed monolayers of CH_3- and OH-terminated alkanethiols as a function of the molar fraction of CH_3-terminated alkanethiols. **B.** Histograms of adhesion forces measured on the mixed self-assembled monolayers (SAMs) using hydrophobic CH_3 tips. **C.** Variation of adhesion forces as a function of the cosine of the water contact angle. **D.** Adhesion force as a function of the surface fraction of CH_3-terminated alkanethiols computed using Cassie's law. (Reprinted with permission from Alsteens et al. [10].)

force as a function of the surface fraction of CH_3-terminated alkanethiols, determined using Cassie's law and contact-angle values (Figure 18.1D). CFM with tips bearing charged groups can likewise be used to probe surface charge and electrostatic forces in aqueous solutions [11,12].

18.2.2. Chemical Imaging of Live Cells

CFM was recently used in imaging hydrophobic groups on the surface of the human opportunistic pathogen *Aspergillus fumigatus* (Figure 18.2) [9, 13]. Topographic images revealed the presence of regularly arranged rodlets on dormant spores. These structures are composed of hydrophobins, a family of small, moderately hydrophobic proteins that favour spore dispersion by air currents and mediate adherence to host cells. Force curves recorded across these surfaces with a hydrophobic tip showed large adhesion forces of ~3,000 pN magnitude. Comparing the data obtained on reference surfaces (Figure 18.1D) indicated that the conidial surface has a marked hydrophobic character, corresponding to a surface composed of ~10 CH_3 and ~15 OH groups, which is fully consistent with the presence of an outermost surface layer of hydrophobins and provides direct indications as to their putative functions as dispersion and adherence structures. In agreement with the uniform surface structure, adhesion maps were rather homogeneous, supporting the notion that the conidial surface is homogeneously hydrophobic.

Dramatic changes of hydrophobic properties can be tracked in real time [14]. A temperature-controlled AFM was used to obtain high-resolution images of the same *A. fumigatus* spore during the course of germination (Figure 18.2). Significant structural alterations were observed, and the rodlet layer changed into a layer of amorphous material, presumably reflecting the underlying polysaccharides. In addition, adhesion maps with hydrophobic tips revealed a dramatic loss of hydrophobicity over time. After 2-hr germination, heterogeneous hydrophobic contrast was observed, reflecting the coexistence of hydrophobic rodlets and hydrophilic polysaccharides. The observed changes are likely to be function related. Whereas the hydrophobic rodlets promote spore dispersion and adhesion to surfaces, the very hydrophilic nature of the germ tube cell wall is likely to favour hyphal growth in aqueous environment.

The impact of drug treatments on the chemical properties of microbes can also be studied using this technique. For instance, the surface of mycobacteria was shown to be uniformly hydrophobic, reflecting the presence of an outermost layer of hydrophobic mycolic acids [10]. This finding is of biomedical relevance because these hydrophobic constituents are thought to represent an important permeation barrier to common antibacterial agents. By contrast, treatment of the cells with drugs targeting specific cell wall constituents lead to a dramatic decrease of cell surface hydrophobicity, attributed to the removal of the mycolic acid layer [15].

In addition to hydrophobicity, surface charge also plays an important role in controlling antigen–antibody, cell adhesion, cell–virus, cell–drug, and cell–ion interactions. This trait can be probed on the nanoscale by using tips functionalized with ionizable carboxyl groups. For the yeast, *Saccharomyces cerevisiae*, force–distance curves were strongly influenced by pH [12]: Whereas no adhesion was measured at neutral/alkaline pH, reflecting the electrostatic repulsion between the negatively charged surfaces, multiple adhesion forces were recorded at low pH values, which are attributed to hydrogen bonding between the protonated tip surface and cell surface macromolecules. These changes were shown to be related to differences in the ionization state of the cell-surface functional groups: The adhesion force versus pH curve was correlated with microelectrophoresis data, with the pH of the largest adhesion force corresponding to the cell isoelectric point.

Hence, CFM is a powerful tool for imaging and quantifying the local hydrophobic and electrostatic properties of cells, thereby complementing the range of methods available for

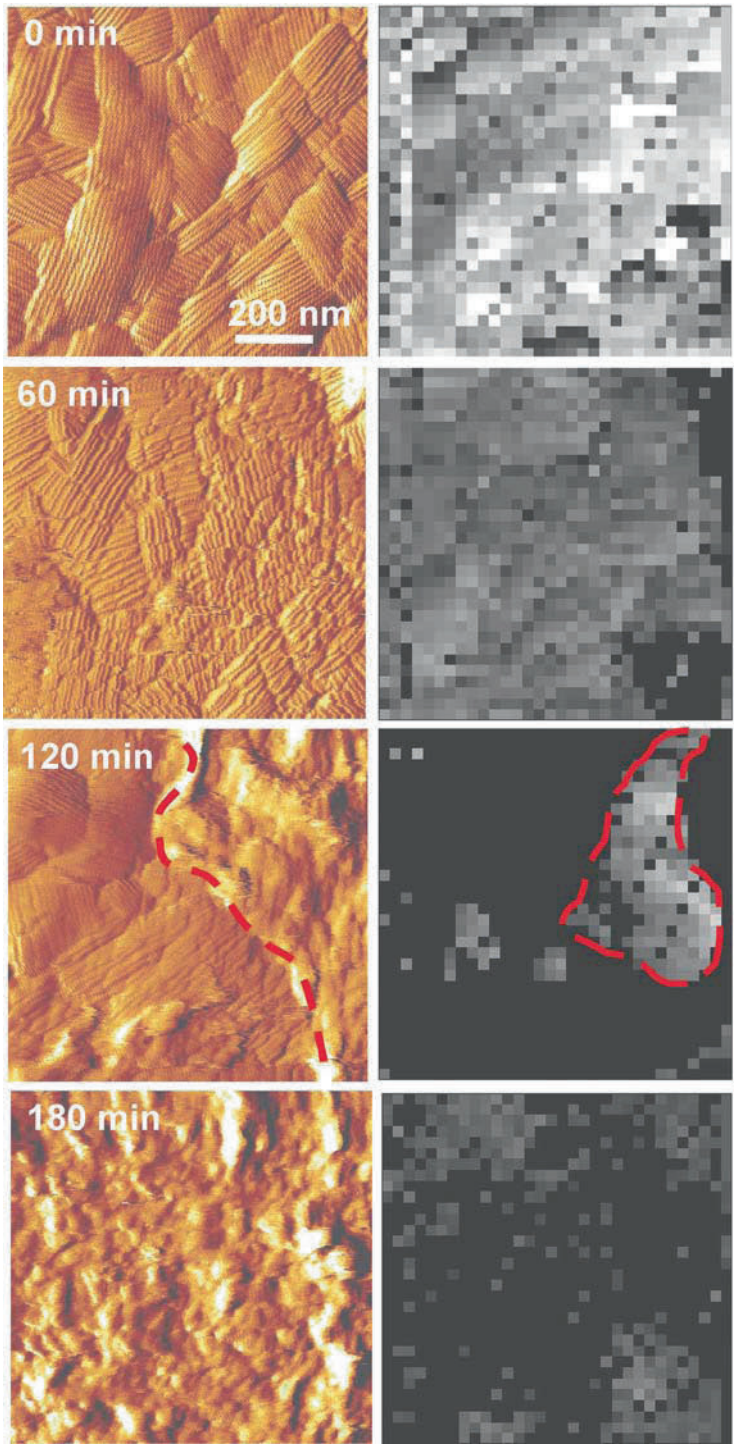


Figure 18.2. Chemical force microscopy of single live cells. Series of high-resolution deflection images **(left)** and adhesion force maps **(right)** recorded with a hydrophobic tip on a single *Aspergillus fumigatus* spore during

assessing cell surface properties. A unique feature of CFM is its ability to resolve nanoscale chemical heterogeneities on single live cells.

18.3. Recognition Imaging Using Force Spectroscopy

Single-molecule force spectroscopy (SMFS) provides unique possibilities for localizing individual receptors on biosurfaces. In addition, SMFS can measure molecular recognition interactions at the level of single molecules, providing valuable information on molecular dynamics within complexes.

18.3.1. Methods

Molecular recognition studies imply functionalizing the tips (and supports) with relevant biomolecules. Several issues should be considered. The forces that immobilize the molecules have to be stronger than the intermolecular force being studied. The attached biomolecules should have enough mobility so that they can freely interact with complementary molecules. The contribution of nonspecific adhesion to the measured forces should be minimized. Attaching biomolecules at a low surface density is recommended to ensure single-molecule detection. Site-directed coupling may be desired to orientate all the interacting molecules in the same way. Immobilization strategies commonly used for making such biological tips are illustrated in Figure 18.3. Using the nonspecific adsorption of biotinylated bovine serum albumin (BBSA), it is possible to measure the interaction between biotin and avidin (or streptavidin) down to the single-molecule level [16,17]. The adsorbed BBSA layer can actually further react with avidin/streptavidin to attach biotinylated molecules [18]. A second approach relies on the strong binding of thiols on gold-coated tips. Whereby proteins, oligonucleotides, and carbohydrates that bear thiol groups can directly be attached on the gold surfaces [19,20], they can also be covalently attached onto SAMs of functionalized alkanethiols on gold by using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) [21]. In this context, an interesting approach is to mix long-chain alkanethiols with COOH-terminal functions in a matrix of shorter OH-terminated alkanethiols, which ensures a certain mobility of the attached biomolecules and minimizes nonspecific adsorption [22]. We would like to point out that, it is possible to orientate all biomolecules in the same way by attaching recombinant histidine-tagged proteins onto an AFM tip coated with nitrilotriacetate (NTA)-terminated alkanethiols [23]. This coupling approach offers the advantage of allowing optimal exposure of the C-terminal or N-terminal domains, but it is limited by the rather low binding strength of the NTA–His bond. A third approach is to covalently anchor biomolecules on silicon tips by using various amine-functionalization procedures [24,25]. The amino-terminated surfaces are reacted with a cross-linker that provides the ligands with motional freedom and prevents their denaturation. Cross-linkers typically

Figure 18.2. (continued) germination. Within less than 3 hr, the crystalline rodlet layer changed into a layer of amorphous material, presumably reflecting inner-cell-wall polysaccharides. After 2 hr, both rodlet and amorphous regions were found to coexist (separated by dashed line). Consistent with this structural dynamics, substantial reduction of adhesion contrast was noted with time (right images), reflecting a dramatic decrease of hydrophobicity. After 2 hr, heterogeneous contrast was observed in the form of hydrophobic patches (*dashed line*), surrounded by a hydrophilic sea. (Reprinted with permission from Dague et al. [14].)

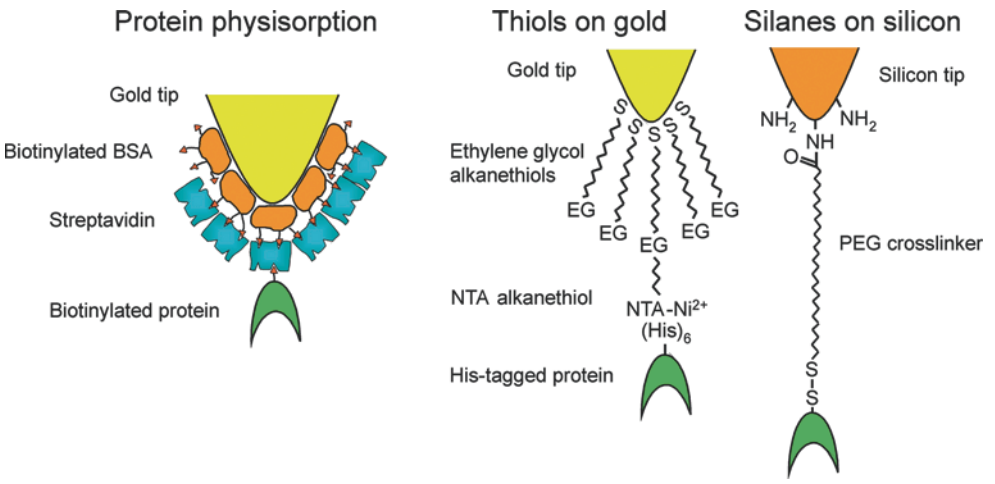


Figure 18.3. Strategies commonly used for modifying atomic force microscope tips with biomolecules. **Left to right.** Physisorption of proteins such as biotinylated bovine serum albumin (BSA), chemisorption of alkanethiols on gold, and covalent coupling of silanes on silicon oxide. NTA, nitrilotriacetate; PEG, poly(ethylene glycol). (Adapted with permission from Hinterdorfer and Dufrene [4].)

carry two different functional ends. This can be, for instance, an amine-reactive NHS group on one end for the coupling to tip surfaces and a 2-pyridyldithiopropionyl (PDP) or vinyl sulfone group on the other end, which can be covalently bound to thiol groups contained in ligands.

For cell adhesion studies, animal and microbial cells can be attached onto AFM cantilevers, allowing microscopists to probe various types of cellular interactions. Microbial cells have been immobilized by using glutaraldehyde treatments to create covalent cross-linking between cells and tips [26] or by attaching single cells with a small amount of glue [27]. However, these treatments are likely to affect the structure and properties of the cell surfaces. Alternatively, bacteria can be physically adsorbed onto poly-L-lysine–coated glass beads followed by attachment of the cell-coated beads to a cantilever by using a small amount of epoxy resin [28]. Another elegant approach is to attach individual cells to an AFM cantilever via lectins such as wheat germ agglutinin or concanavalin A, which then allows the measurement of the specific adhesion forces between two adjacent cells [29] or between cells and immobilized ligands [30].

18.3.2. Measuring Molecular Recognition Forces

Essentially, molecular recognition using SMFS implies recording force curves between the modified tip and sample, assessing the unbinding force between complementary receptor and ligand molecules from the adhesion "pull-off" force observed upon retraction, and displaying the values either as a histogram or an adhesion map. A typical force curve obtained by using a biotin-modified AFM tip and surface-attached avidin molecules is shown in Figure 18.4A [31]. If the ligand on the tip does not form a specific bond with the receptor on the cell surface, the recognition event is missing and the retrace (Figure 18.4B, blue line) looks like the trace (Figure 18.4B, red line). In addition, the specificity of ligand–receptor binding is usually demonstrated by blocking experiments with free ligands, which are injected into

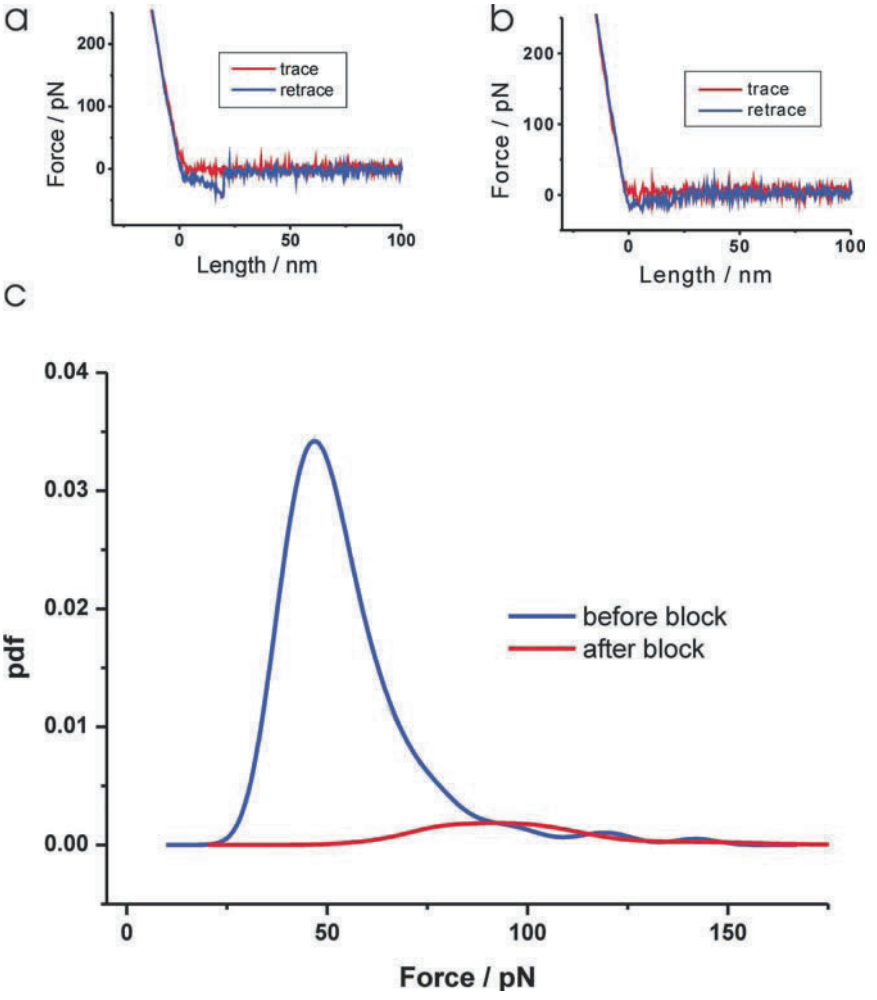


Figure 18.4. Typical molecular recognition force spectroscopy data using a biotinylated tip and an avidin-modified surface. **A.** A force–distance circle exhibiting an unbinding event in the retraction period (*blue line*). **B.** This unbinding event disappears when the tip is blocked by adding free streptavidin. **C.** The probability density function (pdf) of the unbinding forces (*blue line*) shows a clear maximum and a high binding probability. In contrast, the streptavidin-blocked tip has a significant lower and broader peak due to the hindered avidin–biotin complex formation. (Adapted from Ebner et al. [31].)

the solution to block the receptor sites on the surface. As a consequence, almost all specific recognition signals completely disappear and only occasional adhesion events are observed.

After acquiring hundreds of force curves, one can construct empirical probability density functions (pdf's) from the detected unbinding forces (Figure 18.4C). The maximum of the distribution reflects the most probable force on which a single ligand–receptor bond dissociates under the force ramp used. An overall binding probability, which is the probability to record an unbinding event, of 19% $\pm$ 6% (Figure 18.4C, blue line) was obtained. Blocking experiments performed by injecting free streptavidin molecules in solution reduced this probability to 2% $\pm$ 1% (Figure 18.4C, red line). No binding was found at all when a bare tip

was used instead of a ligand-coated tip. These results strongly support the specificity of the binding events observed.

Based on such force spectroscopy experiments, a variety of ligand–receptor forces have been measured at the single-molecule level, including those associated with avidin/streptavidin [16,17] antibodies [24,32], DNA [19,33,34], cadherins [35], and bacterial adhesins [23]. Knowledge of these forces contributes to the refinement of our understanding of the molecular basis of molecular recognition events such as those mediating cell adhesion and immunological processes and offers promising prospects in nanobiotechnology for the rapid detection of bioanalytes. Several reports showed that unbinding forces between receptors and ligands depend on the loading rate, that is, the rate at which the force is applied to the bond. We want to point out that, the interaction time may also shift the distribution of adhesion forces towards larger values. For VE-cadherins, this phenomenon was ascribed to the time-dependent association of the protein into complexes with higher-order adhesion strengths [35,36]. For the bacterial adhesin heparin-binding haemagglutinin adhesin (HBHA), the prolonged contact time required to establish strong interaction with heparin was suggested to reflect the time necessary for conformational changes within both molecules to allow an optimal fitting between charged groups [23]. Hence, it is essential in molecular recognition force measurements to vary the loading rate and interaction time to provide reliable data and assess kinetic parameters of the unbinding process.

18.3.3. Molecular Recognition Imaging Using SMFS

Molecular recognition imaging using SMFS is comprised of recording arrays of force curves in the x, y plane on an area of given size, assessing the unbinding force values for all curves, and displaying them as grey pixels [37]. This method has been applied to different cell types, including red blood cells [38], osteoclasts [39], and endothelial cells [40]. In microbiology, molecular recognition imaging may be used to map adhesion proteins on live bacteria (Figure 18.5) [23]. Bacterial infections are generally initiated by the interaction between adhesins, that is, cell adhesion proteins, on the bacterial pathogen and specific receptors on the host cell surface. A key example is *Mycobacterium tuberculosis*, which adheres to heparan sulphates on epithelial cells via the HBHA. To shed new light on the molecular basis of this interaction, *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) cells expressing HBHA were immobilized on a polycarbonate membrane, an approach that allows live cells to be imaged without using any drying or fixation step (Figure 18.5A). High-resolution images revealed a smooth and homogeneous surface (Figure 18.5B), consistent with earlier scanning electron microscopy observations. Adhesion force histograms (Figure 18.5C) and maps (Figure 18.5D) recorded on cells with a heparin-modified tip revealed adhesion events (bright pixels on the maps) in about half of the locations. The adhesion force magnitude was very close to the value expected for a single HBHA–heparin interaction, supporting the notion that single HBHAs were detected. This was confirmed by showing that a mutant strain lacking HBHA did not bind to the heparin tip. Of interest to, the HBHA distribution was not homogeneous but apparently concentrated into nanodomains, which may promote adhesion to target cells by inducing the recruitment of receptors within membrane rafts. In the future, these molecular recognition studies may help in the development of new drugs capable of blocking bacterial adhesion.

SMFS with antibiotic-modified tips may be used to gain insight into the architecture of bacterial cell walls, as well as into the action mode of antibiotics. In the first such study, AFM tips bearing vancomycin molecules were used to localize single D-Ala-D-Ala–binding

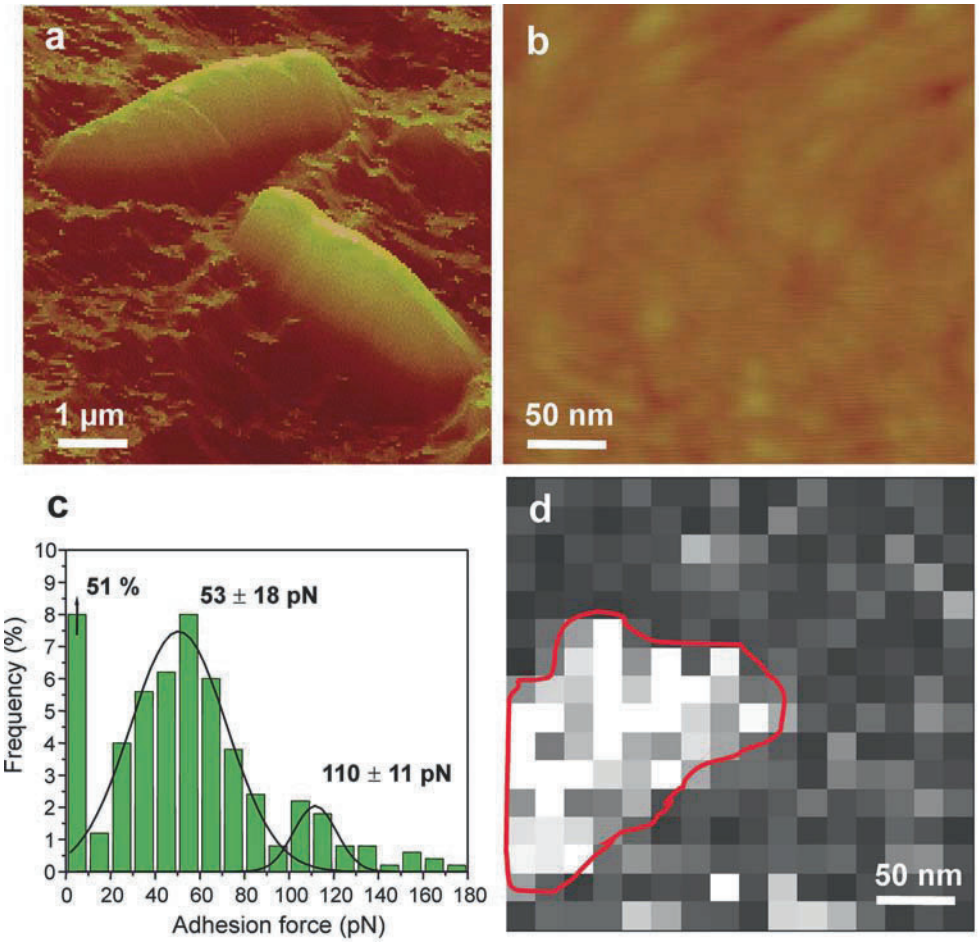


Figure 18.5. Imaging single-cell adhesion proteins on live bacteria using spatially resolved single-molecule force spectroscopy. **A.** Topographic image showing two living mycobacteria on a polymer support. **B.** Higher resolution of the cell surface revealing a smooth morphology. **C, D.** Adhesion force histogram and adhesion force map, respectively, recorded on a single cell with a heparin-modified tip. In localized regions, the map reveals adhesion events (clear pixels) due to the presence of adhesion proteins, which seem to be concentrated into nanodomains. (Reprinted with permission from Dupres et al. [23].)

sites up on dividing *Lactococcus lactis* cells [41]. Topographic images of the cells revealed a smooth and elongated cell morphology, as well as a well-defined division septum. Please note, that adhesion force maps recorded with vancomycin-terminated tips demonstrated that the vancomycin-binding sites were essentially located in the septum region, suggesting that newly formed peptidoglycan was inserted in these regions. This was in excellent agreement with fluorescence microscopy images obtained using a fluorescent vancomycin probe.

SMFS also enables researchers to map the distribution of polysaccharides on cell surfaces and to probe their adhesion, elasticity, and conformational properties [42]. For instance, lectin-modified tips were used to pick up and stretch single mannan polysaccharide chains on the two brewing yeast strains, *Saccharomyces carlsbergensis* and *S. cerevisiae* (Figure 18.6) [43]. Polysaccharides were clearly more extended on *S. cerevisiae*, suggesting that not only

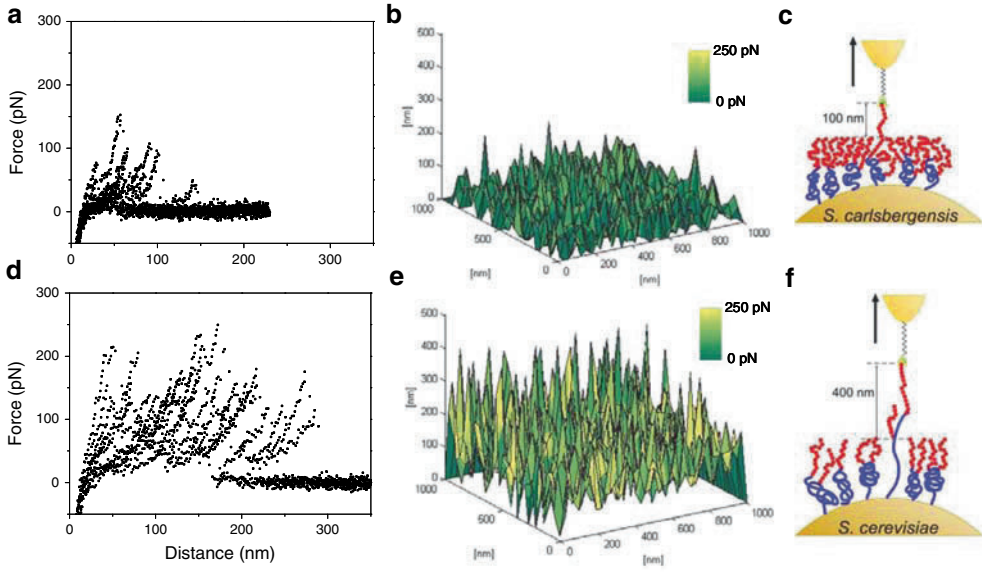


Figure 18.6. Localization and conformational analysis of single polysaccharide molecules on yeast using single-molecule force spectroscopy (SMFS). **A, D.** Superimpositions of retraction force curves recorded on *Saccharomyces carlsbergensis* **(A)** and *Saccharomyces cerevisiae* **(D)** yeast cells with an atomic force microscope tip functionalized with the lectin concanavalin A. **B, E**. Three-dimensional maps of polymer properties for *S. carlsbergensis* **(B)** and *S. cerevisiae* **(E)**, showing very different spatial distributions of adhesion forces (false colors; yellow means more adhesive) and of rupture distances (*z* levels). **C, F.** Schematics of the SMFS measurements: Whereas only the mannan oligosaccharides (*red*) are stretched on *S. carlsbergensis*, both the oligosaccharide (*red*) and polypeptide (*blue*) chains of mannoproteins are pulled on *S. cerevisiae*. This may explain why, unlike *S. carlsbergensis* (bottom-fermenting brewing strain), *S. cerevisiae* tends to associate with CO_2 bubbles (top-fermenting brewing strain). (Reprinted with permission from Alsteens et al. [43].)

oligosaccharides, but also polypeptide chains of the mannoproteins were stretched. These major differences in polymer properties could explain the very different aggregation properties of the two organisms. In another study, the method was used to reveal the coexistence of polysaccharide chains of different nature on surface of the clinically important probiotic bacterium *Lactobacillus rhamnosus* GG [44]. These SMFS measurements offer exciting prospects for understanding the molecular mechanisms of cell adhesion.

Molecular recognition imaging using SMFS is limited by its poor temporal resolution. The time required to record a molecular recognition image is about a 10 minute segment, depending on the acquisition parameters, which is much greater than the time scale at which dynamic processes usually occur in biology. In the future, progress in developing faster adhesion mapping modes, like the jumping mode [45], will help to circumvent this problem.

18.4. Topography and Recognition Imaging

In contrast to common recognition imaging using force spectroscopy, a recently developed imaging technique termed topography and recognition imaging (TREC) [46–52] overcomes some of the limitations regarding lateral resolution and imaging speed by using dynamic force microscopy with a functionalized sensor tip that is oscillated during scanning across the surface.

18.4.1. Methods

As a first step, the AFM tips must be upgraded into a molecular sensor by coupling a ligand molecule to the outer end of the scanning tip, preferably through a flexible tether such as poly(ethylene glycol) (PEG) [53]. Because it is an important parameter for adjusting the optimal oscillation amplitude, the length of the PEG tether plays a critical role here. It can range from 8 to 25 ethylene glycol units, corresponding to 3–10 nm of stretched tether length. The immobilization of the sensor molecule via the PEG tether gives the ligand the freedom to adopt the correct orientation so as to facilitate binding to the target molecule on the surface, while the sensor tips scan across the surface. In this approach, the so-called MacMode (Agilent, USA) is used. A magnetically coated cantilever is oscillated through an alternating magnetic field. In contrast to "normal" MacMode imaging, in TREC only the lower part of the oscillation is used to drive the feedback loop for obtaining the topography image. The upper part of the oscillation is used for the generation of the recognition image.

One-dimensional linear scans [52] allow a clear investigation of the nature of the topography and recognition signals. For this purpose, lysozyme molecules were immobilized onto a sample surface at a concentration at which they were singly distributed. The surface was scanned with a bare and with an antilysozyme (HyHel5) functionalized tip, respectively. Figure 18.7A shows the raw signal of a 7-kHz oscillation over a time range of 4 ms. In a presentation of significantly larger time spans of these traces (1 sec, corresponding to one full scan line of 500 nm) the oscillations appeared highly compressed, and only the envelope with the characteristic maxima and minima of each oscillation period remained visible (Figure 18.7B, C). Figure 18.7B shows the result of a scan with a bare tip, that is, a tip containing no antibody. The z positions of the minima varied significantly, and as a result singly distributed bulges with 10–15 nm width and about 1 nm in apparent height appeared along the scan axis. These bulges reflect single lysozyme molecules that resist the further downward movement of the tip towards the mica surface. Their widths and heights are a measure of the apparent molecule size (10–15 nm width, about 1 nm height), and their separation is a measure of the distance between the molecules on the surface (about 35 nm on average). In contrast, the positions of the oscillation maxima remained constant, aside from the minor, randomly occurring variations caused by the thermal noise of the cantilever.

Apparently, the information of the surface topography measured with a bare tip is solely contained in the minima of the cantilever oscillations, and cross-talk between minima and maxima does not exist at the conditions used (Q factor of the cantilever $\sim$1) [54]. Distinct minima were also clearly detected with an AFM tip carrying a specific antibody (Figure 18.7C), indicating that the topography information can also be obtained by using these chemically modified tips. In addition, however, the maxima were also significantly affected (Figure 18.7C). The antibody of the AFM tip binds to the specific antigenic sites on lysozyme during scanning, and the physical connection temporarily established between tip and substrate thereby reduces the upstroke of the cantilever oscillation. As a further consequence, recognition of lysozyme by the tip-conjugated antibody results in reduction of the oscillation maxima, which allows for detection of the lateral position of specific binding sites.

Topography and recognition images can be simultaneously obtained using a specially designed electronic circuit (PicoTREC, Agilent, Chandler, Arizona). The principle of TREC (Figure 18.8A) is hereafter explained. The time-resolved deflection signal of the oscillating cantilever is low-pass filtered to remove the thermal noise, DC-offset levelled, and amplified before it is split into the lower (U_{down}) and (U_{up}) upper parts of the oscillations. The signal passes a trigger threshold on each path, and the lower and upper peak of each oscillation

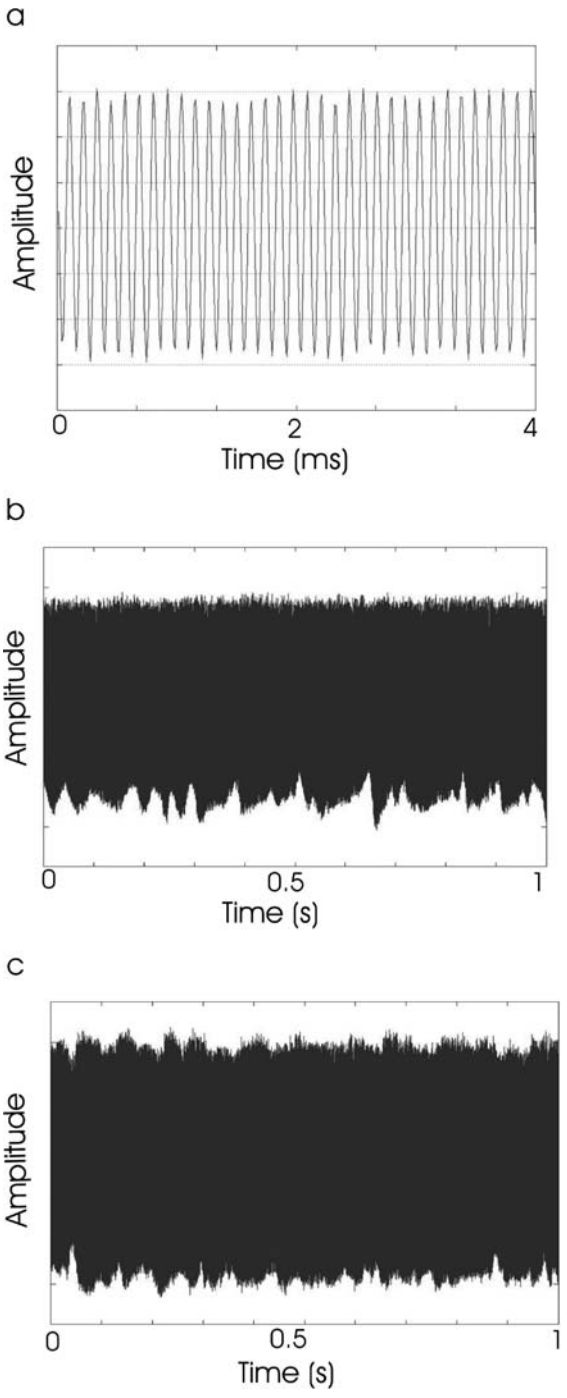


Figure 18.7. Deflection signals of magnetically oscillated cantilevers using disabled feedback. **A.** Oscillations signals over a time range of 4 sec with a peak-to-peak amplitude of ~5 nm. **B, C.** Full line traces over 500 nm (in 1 sec) with a bare tip and using a ligand-tethered tip, respectively. (Reprinted with permission from Stroh et al. [52].)

period is determined by means of sample and hold analysis. Succeeding peaks result in a staircase function, which is subsequently filtered and fed into the controller of the microscope, with U_{down} driving the feedback loop to record the height (i.e., topography) image and U_{up} providing the data for constructing the recognition image. If cantilevers with low Q factor ($\sim$1 in liquid) driven at frequencies below resonance are used, both types of information are independent. In this way, topography and recognition images can be recorded simultaneously and independently.

TREC imaging was applied to singly distributed avidin molecules by using a biotinylated tip (Figure 18.8B, C). The tip oscillation amplitude (8 nm) was chosen to be slightly

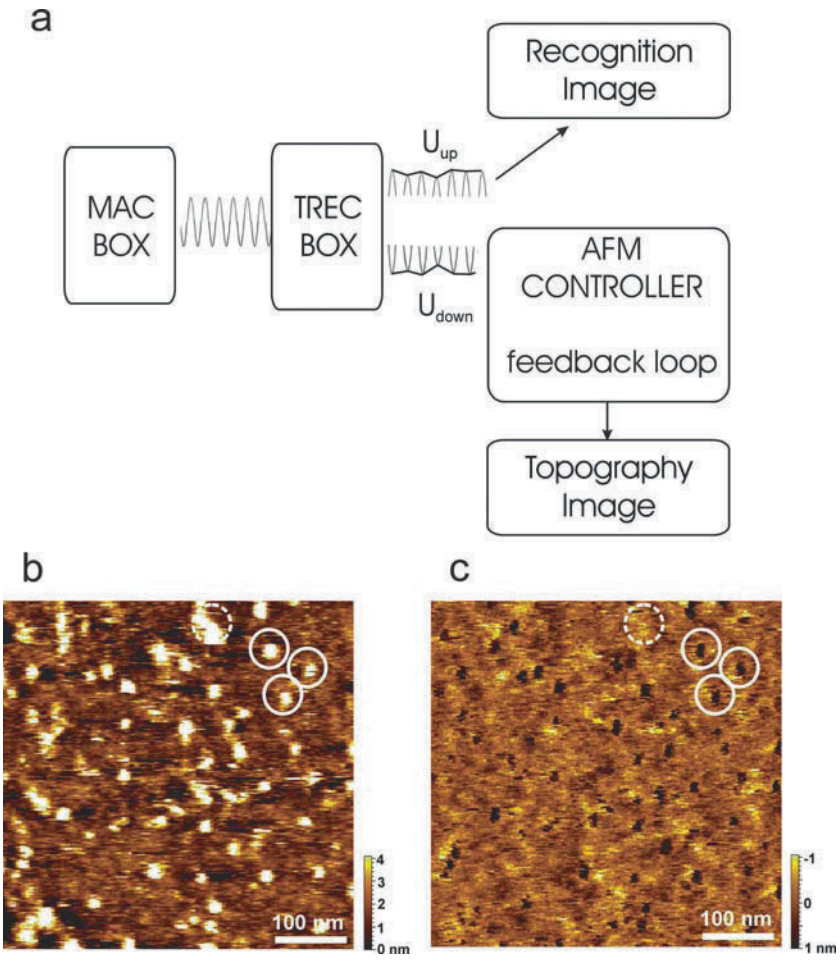


Figure 18.8. Topography and recognition imaging (TREC) signal processing. **A.** The raw cantilever deflection signal is fed into the TREC box, where the maxima (U_{up}) and the minima (U_{down}) of each oscillation period are depicted and used for the recognition and the topography image, respectively. AFM, atomic force microscope; MAC, magnetic alternating current mode. (Reprinted with permission from Stroh et al. [52].) **B, C.** Topographic and recognition images, respectively, using a biotin-functionalized tip. In the topographic image, single avidin molecules electrostatically adsorbed to mica are observable. In the simultaneously acquired recognition image, recognition events appear as black spots, showing a good correlation to the corresponding topography. (Reprinted with permission from Ebner et al. [66].)

smaller than the extended PEG cross-linker length (10 nm) through which the biotin was coupled to the AFM tip. Thus, both the biotin remained bound while passing a binding site and the reduction of the upwards deflection was of sufficient significance compared to the thermal noise. Because the spring constant of the polymeric cross-linker increases nonlinearly with the tip–surface distance, the binding force is only sensed close to full extension of the cross-linker (given at the maxima of the oscillation period). Therefore, the recognition signals are well separated from the topographic signals arising from the surface both in space ($\Delta z \sim$ 5 nm) and time (half-oscillation period $\sim$0.1 ms).

The visible bright dots with 2–3 nm in height and 15–20 nm in diameter in the topography image (Figure 18.8B) represent single avidin molecules stably adsorbed onto the flat mica surface. The recognition image shows black dots at positions of avidin molecules (Figure 18.8C) because the oscillation maxima are lowered due to the physical avidin–biotin connection established during recognition. The lateral positions of the avidin molecules obtained in the topography image are spatially correlated with the recognition signals of the recognition image. (circles in the images (Figure 18.8B, C)). Recognition between the biotin on the tip and the avidin on the surface took place for almost all avidin molecules (for exception, see dotted circle in Figure 18.8C). Thus, one would assume to have almost all binding epitopes oriented away from the mica surface and accessible to the biotin on the tip, resulting in a high binding efficiency.

In addition, the tether length (10 nm) via which the biotin was bound to the AFM tip was greater than the oscillation amplitude (8 nm) so that the biotin on the tip always had a chance to bind to the avidin when passing a binding site during lateral scans. Biotin–avidin recognition resulted in a reduction of the oscillation amplitude, and, indeed, due to the close proximity of the tip to the surface, the binding efficiency was high. It is important to note that topography and recognition images can be recorded at speeds typical for standard AFM imaging and are therefore considerably faster than conventional force mapping.

18.4.2. Applications of TREC Imaging

Chromatin

The fundamental particle of the chromatin structure—the nucleosome—is a self-assembled complex of basic histone proteins wrapped by approximately two turns of DNA. Using conventional AFM, it is challenging to identify and distinguish the single components because this technique only yields the shape and volume of the structures, which does not heavily depend on the molecular weight. In contrast, TREC allows recognition of a specific type of molecule in a complex structure while simultaneously yielding high-resolution topographic images. This strategy recently localized, histones H3 in remodelled chromatin structures [48]. For this procedure, nucleosomal arrays containing the mouse mammary tumor virus (MMTV) promoter region were salt-reconstituted and subsequently deposited on glutaraldehyde aminopropyltriethoxysilane (GD-APTES)–treated mica. Polyclonal anti–histone H3 antibodies were thiolated and attached to a PEG tether on the end of an AFM tip. Topographic and recognition images of nucleosomal arrays were simultaneously recorded using this tip. The dark recognition patches and the pattern in which they occur clearly coincided with the positions of the nucleosomes in the topographic image because histone H3 is one of the proteins that constitute the core of each nucleosome.

The specificity of the recognition process was tested by injecting exogenous proteins into the liquid cell in which the sample was imaged. When a BSA solution was added, the

recognition image remained unaffected. However, if a similar concentration of a peptide corresponding to amino acids 1–20 of the N-terminal tail of histone H3 was injected into this sample, the recognition signal was completely abolished. Thus, the ability of the tip-tethered antibody to produce a recognition signal was blocked by an excess of its specific antigen in solution, demonstrating that the recognition signals came from specific interactions between the antibody on the tip and its antigen on the surface. The major impact of TREC comes from its ability to monitor specific components in heterogeneous samples while they undergo biological processes. In such experiments, hSwi-Snf and chromatin were incubated together and then deposited onto mica. *In situ* images of the same nucleosomal arrays were taken before ATP was added (hSwi-Snf is inactive) and after ATP addition (hSwi-Snf is activated). The progress of remodelling was therefore tracked on individual molecules. Comparison of recognition images before and after ATP addition showed a variety of remodelling and compositional changes.

TREC on chromatin was also explored using another kind of recognition molecule, the DNA aptamer [55]. These are small, stem-loop, single-stranded DNA molecules generated via systematic evolution of ligands by exponential enrichment. Aptamers consist of a single strand of DNA and are therefore easy to synthesize and store. They are easily folded by thermal annealing in an appropriate buffer and can also be easily attached to an AFM tip by using commercially available DNA that is chemically modified at one end. Aptamers may be as specific as antibodies and have a high affinity for some small molecules. A proof-of-principle TREC study using aptamers as ligands with immunoglobulin E as receptor was performed [55] before histone H4 aptamers were investigated to elucidate their specificity to different histone classes [56].

Using an *in vitro* selected aptamer for histone H4, a series of recognition imaging experiments against recombinant histone H3, H2A, and H2B were performed. It was shown that the H4 aptamer recognized the H4, H3, H2A, and H2B proteins with recognition efficiencies of 80%, 29%, 3%, and 5%, respectively. The low level of recognition to the H2A and H2B proteins correlated with previous affinity measurements performed on the histone tail sequences. Comparison of different tail sequences revealed a GGX motif that was present, twice in H4 tails and once in H3 tails, but was absent from the H2A and H2B tail sequences. It thus appeared likely that this motif is important in aptamer binding. Because it was found that DNA aptamers represent a viable alternative to traditional antibodies in recognition imaging microscopy, it was suggested that this approach could be used to study key epigenetic modifications involved in chromatin remodelling [56].

Bacterial Surface Layers

Crystalline bacterial surface layers (S-layers) represent a unique self-assembly system optimized during billions of years of biological evolution [57–60]. The intrinsic ability of S-layers to self-assemble allows for *in vitro* formation of isoporous two-dimensional (2D) protein lattices in suspension, on lipid films, on liposomes, and on solid supports [57,61]. Strep-tagII, an eight–amino acid peptide and artificial ligand for Strep-Tactin (a genetically engineered streptavidin variant with higher affinity towards Strep-tagII than streptavidin [62]), was fused to a C-terminally truncated form of the S-layer protein SbpA of *Lysinibacillus sphaericus* CCM 2177. SbpA is one of the most extensively studied S-layer proteins, with the ability to self-assemble into a square (p4) lattice [57,61,63,64]. In previous studies, it was found that the fused Strep-tagII did not interfere with the self-assembly properties of S-layer

fusion protein SbpA–Strep-tagII and that the Strep-tagII was located on the exposed surface of the S-layer lattice [65].

We applied TREC for lateral imaging of the recognition sites on the nanoarray, to identify the location of Strep-tagII in the square S-layer lattice with nanoscale resolution. The AFM tip was covalently functionalized with Strep-Tactin via a short, elastic linker. The topography images of the nanoarrays with mixed S-layer proteins rSbpA–Strep-tagII and SbpA were acquired with MAC mode AFM. The 2D nanoarrays of mixed S-layer proteins with square lattice and with a center-to-center spacing of ∼14 nm, typical for the S-layer protein SbpA [60,65], could be observed in the topography images (Figure 18.9A). Figure 18.9B shows the simultaneously acquired recognition image generated by scanning with a Strep-Tactin–coupled AFM tip over the surface of mixed rSbpA–Strep-tagII/wild-type SbpA (1/7) lattice. When Strep-Tactin on the tip recognized Strep-tagII, the oscillating tip response was altered, and this change was converted into a recognition signal [48,66] that displayed the location of Strep-tagII in the S-layer nanoarrays. The dark spots in the recognition image (Figure 18.9B) indicate where recognition events between Strep-Tactin and Strep-tagII occurred and reflect the distribution of the fusion protein rSbpA–Strep-tagII in the cocrystallized S-layer lattice consisting of rSbpA–Strep-tagII/wild-type SbpA. Although the dark

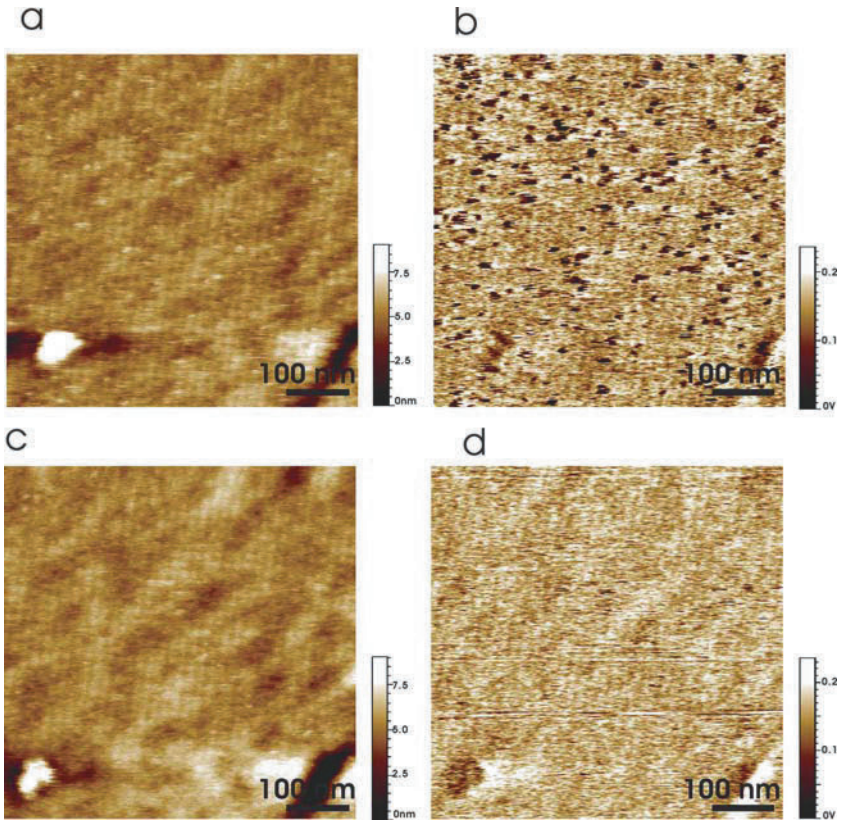


Figure 18.9. Topography and recognition images of surface-layer in the absence and presence of blocking agents. **A, B.** Topographic and recognition images, respectively, of the surface-layer proteins rSbpA–Strep-tagII and wild-type SbpA cocrystallized (in a molar ratio of 1:7) on a silicon chip. **C, D.** Topographic and recognition images acquired after Strep-Tactin on the atomic force microscope tip was blocked by adding free Strep-tagII.

spots were located all scattered over the nanoarray surface of the cocrystallized S-layer proteins, they were randomly located. The specificity of the recognition signals was proven by acquiring topography and recognition images at the same place after blocking Strep-Tactin on the tip (Figure 18.9C, D). The dark spots in the recognition image were completely abolished after injection of free Strep-tagII into the AFM fluid cell during imaging (Figure 18.9D). In this case, the ability of the tip-tethered Strep-Tactin to acquire a recognition signal was blocked by the excess of Strep-tagII in solution. This proved that the recognition signals shown in Figure 18.9B arose from specific interaction between Strep-Tactin coupled to the AFM tip and Strep-tagII on the surface of the cocrystallized S-layer lattice.

Subsequently, a smaller scan area on the surface of the cocrystallization of rSbpA–Strep-tagII and SbpA was imaged to increase the lateral resolution. Figure 18.10A, B show the simultaneously acquired topography and recognition images, with a scan size of 216 nm. To precisely localize the position of Stre*p*-tagII, a "center-of-mass" approach was used. In this case, the center-of-mass approach determined the position of the symmetry point, which was identified as the recognition site. Even though the diameter of the recognition spot was on the order of two times the length of PEG and protein (∼7 nm), the position of the recognition spots was determined with accuracy of less than 1 nm. By overlaying the position of Strep-tagII in recognition spots and the topography image (Figure 18.10C), 88% of the position of Strep-tagII in recognition spots were found on the corner of the square lattice, which indicated the location of Strep-tagII in the lattice of S-layer. Figure 18.10D shows a schematic drawing of the cocrystallization of rSbpA–Strep-tagII and wild-type SbpA and

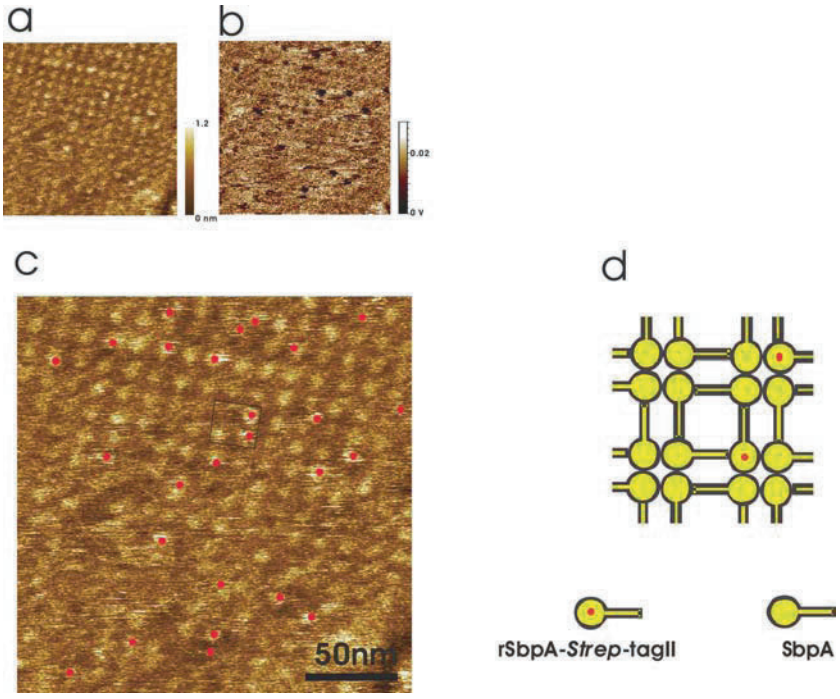


Figure 18.10. A, B. Topographic and recognition images, of the surface-layer proteins rSbpA–Strep-tagII and SbpA (in a molar ratio of 1:7) cocrystallized on a silicon chip. **C.** Superimposition of a recognition map of Strep-tagII onto the corresponding topographic image. **D.** Schematic drawing of cocrystallized rSbpA–Strep-tagII and wild-type SbpA.

represents the square area in Figure 18.10C. The red spot indicates where the Strep-tagII was located in correspondence to the dark spots in the recognition image. These results demonstrate that Strep-tagII was exposed on the surface of the S-layer lattice and located in the corner of the square S-layer lattice.

Membranes

In a recent study, the distribution of cystic fibrosis transmembrane regulator (CFTR) molecules in red blood cell membranes was investigated. CFTR is a chloride channel and a regulator protein for several membrane conductases [67]. It maintains the salt and water balance on the epithelium and regulates processes such as cell volume regulation [68]. CFTR dysfunction results in a severe disease, namely cystic fibrosis (CF). It is the most common genetic disease in people and is characterized by impaired epithelial transport mainly in the respiratory system, pancreas, and liver. The most typical mutation is the deletion of the amino acid phenylalanine at position 508, resulting in the misfolded protein (F508del CFTR). This leads to impaired trafficking to the plasma membrane. As a result, the number of CFTR molecules in the plasma membrane is strongly reduced [69,70]. There is strong evidence that CFTR is not only expressed in epithelia but also in human erythrocytes [71–76]. In this context, TREC imaging was used to accurately map CFTR sites directly on erythrocyte membranes.

Before starting with the TREC experiments the erythrocyte membrane preparation was investigated by dynamic force mode imaging by using a bare cantilever. The membranes appear as round, flat structures with a diameter of $\sim$10 μm and up to 25 nm in height. For TREC imaging, an anti-CFTR antibody was covalently coupled to an AFM tip via a lysine residue. The TREC measurements were performed on erythrocyte membranes of healthy (non-CF) and CF-positive patients, respectively. Topographic images of both non-CF (Figure 18.11A) and CF (Figure 18.11D) erythrocyte membranes revealed similar structures protruding out of the membranes with 10–12 nm in height, representing membrane proteins. These structures were comparable to the topographic measurements using an unmodified AFM tip with similar specifications. Visualization at single-molecule level was achieved without compromising its topographic imaging performance, despite the fact that the tip was carrying a tethered antibody. The simultaneously acquired recognition images (Figure 18.11E) showed dark spots, corresponding to interactions between the specific anti-CFTR antibody on the tip and the CFTR membrane proteins. These binding sites can be assigned to particular topographic structures, allowing the identification of CFTR among the abundance of different proteins present in the membrane.

The size of recognition spots for CF and non-CF membranes was 32 $\pm$ 5 and 33.5 $\pm$ 8 nm, respectively. The lateral resolution of TREC is limited by PEG-antibody-ligand system, which causes a lateral mobility of the antibody restricted by the linker length. The configuration used allows an antigen–antibody binding in a distance of up to $\sim$12 nm (linker length plus antibody) to the tip apex. Because the recognition signals were recorded relative to the tip position, a single antigen could generate a recognition spot with a diameter of up to 24 nm. The tip radius also contributes to the signal broadening, so that the overall diameter of a recognition spot could be up to 30–35 nm. Taking this into account, we see that the observed recognition spots with diameters of about 33 nm revealed the presence of at least one CFTR molecule. Although the cluster formation plays an important role, the exact number of CFTR molecules in one recognition spot cannot be given. Nevertheless, it could clearly be shown

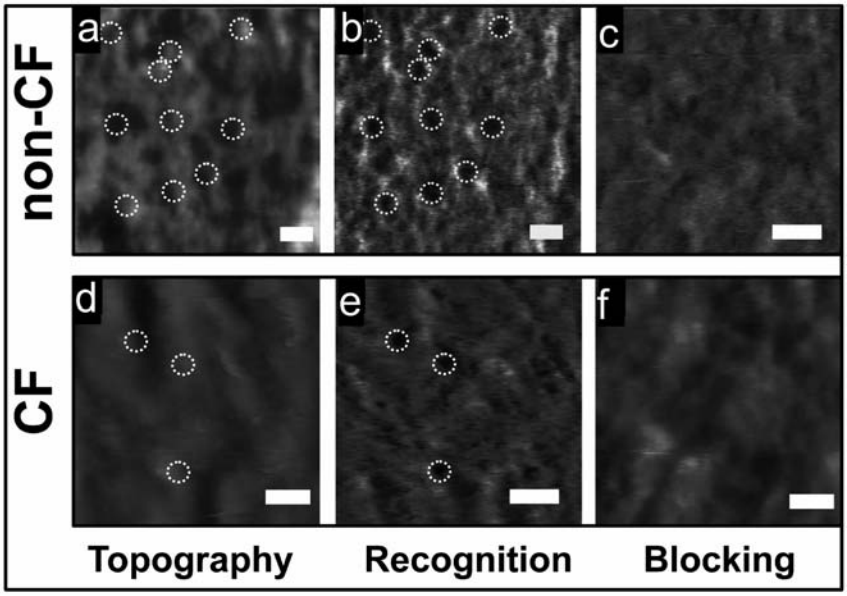


Figure 18.11. Topographic and recognition images of isolated erythrocyte membranes. Non–cystic fibrosis (CF) **(A)** and CF-positive **(D)** erythrocyte membranes show molecular recognition events **(B, E)** when scanned with an anti–cystic fibrosis transmembrane regulator (CFTR)–functionalized sensor tip. Specificity proof experiments using free anti-CFTR antibodies in solution show no recognition events **(C, F)**. Scale bar × nm.

that CF samples (Figure 18.11E) clearly revealed fewer recognition spots compared to non-CF ones (Figure 18.11B). Quantification of the recognition events revealed values of 6.2 ± 1.29 ($n = 4$) and 1.9 ± 0.90 ($n = 12$) CFTR molecules per μm^2 for non-CF and CF erythrocytes, respectively (mean ± SD). By extrapolation to the total erythrocyte surface area of 130 μm^2, one can estimate ~800 CFTR molecules/erythrocyte for non-CF samples and about 250 CFTR molecules/erythrocyte for CF samples (if multimeric clustering in one recognition spot is neglected).

To test the specificity of recognition, an excess of free anti-CFTR antibodies was injected into the measurement fluid chamber. This resulted in a hindering of tip–antibody surface–antigen complex by forming a CFTR–anti-CFTR antibody complex. Thus the free CFTR antibodies were bound to the membrane-embedded CFTRs, which hindered tip–antibody surface–antigen complex formation. As a result, in the TREC measurements, the recognition spots disappeared in the recognition image with an efficiency of more than 90% due to abolishment of the reduction of the upper oscillation amplitude (Figures 18.11C, F), whereas the sample topography was unaffected.

TREC recognition imaging allowed an accurate identification of CFTR within the plasma membrane at the single-molecule level. This provides considerable advantages because it removes the data-averaging drawback inherent in the conventional techniques that record measurements over large ensembles of molecules. With further advances in the TREC technique (increasing spatial resolution, improving sample preparation) this method provides more detailed information on protein clustering in native membranes. Moreover, elucidating the distribution of single proteins and their organization within the plasma membrane becomes feasible.

Cells

Real-time visualization and quantification of receptor-binding sites on cell surfaces is a fundamentally challenging task in molecular cell biology. The first TREC studies on cells were performed on microvascular endothelial cells of mouse myocardium (MyEnd) to locally identify vascular endothelial (VE)-cadherin binding sites and colocalize the receptor position with membrane topographic features [50] (Figure 18.12A). White VE-cadherin belongs to the widespread family of cadherins, transmembrane glycoproteins are known to play an important role in calcium-dependent homophilic cell-to-cell adhesion. VE-cadherin is located at intercellular junctions of essentially all types of endothelium, where VE-cadherin molecules are clustered and linked through their cytoplasmic domain to the actin cytoskeleton. The cadherin *cis*-dimer, which is formed by the association of two extracellular domains in physiological Ca^{2+} concentration (1.8 mM), represents a basic structural functional unit to promote a homophilic bond between cells [77].

To avoid lateral diffusion of VE-cadherin and increase the stability of the cell membrane, cells were gently fixed with glutaraldehyde, which is not only suitable for preventing the lateral mobility of receptors on the cell surface, but also maintains cell volume and generally preserves the filamentous structure on the cell cortex. TREC measurements were conducted with magnetically coated AFM tips (MAC tips), which were functionalized with a recombinant VE-cadherin–Fc *cis*-dimer via a soft and long (∼8 nm) PEG linker. Recognition signals correspond to the amplitude reduction due to a binding between VE-cadherin molecules on the AFM tip and the cell surface when specific *trans*-interaction occurs (seen as dark spots in recognition image). These dark spots are distributed nonuniformly and represent microdomains with dimensions from ∼10 to ∼100 nm (Figure 18.12B). Figure 18.12B illustrates high recognition efficiency, which remains practically unalterable on subsequent rescans. The addition of 5 mM ethylenediaminetetraacetic acid (Ca^{2+}-free conditions) leads to the disappearance of almost all binding events in the recognition image, whereas no change in the topography image is observed. A closer look at some recognition spots reveals that they consist of one to two large domains (50–80 nm) surrounded by smaller domains (10–20 nm) or even single-molecule spots (typically 1–4 pixels long; 1 pixel ∼4 nm) (Figure 18.12B). Given the size of the VE-cadherin *cis*-dimer (diameter 3 nm) and the free orientation of PEG linkers leading to specific binding even before/after (8 nm) the binding-site position, the dimensions of the single sites reflect the expectation. More than 600 single events were recognized and ∼6,000 active *cis*-dimers were estimated over the scanned area (4 μm^2). The shapes and the positions of VE-cadherin domains were correlated with topographic features of MyEnd cell surfaces. The topography of a scanned MyEnd cell surface area represents a complex picture of linear and branched filamentous structures, likely representing filaments of the peripheral actin belt and some globular features as well. We point out that, a few VE-cadherin domains were found directly on top of filaments. Nevertheless, most domains were located near and between filaments, indicating that at this stage of cell maturation (day 1 or 2 after seeding), clustering of VE-cadherin was incomplete.

In another study, TREC was used to visualize vascular endothelial growth factor receptors (VEGFRs) [78] on vascular endothelial cells. VEGFR is a transmembrane receptor tyrosine kinase, which is primarily expressed in vascular endothelial cells and plays an important role in mechanosensory functions such as focal adhesion turnover, actin cytoskeletal remodelling, and angiogenesis. Intermittent blocking of VEGFR2 in vascular endothelial cells promotes rapid blood vessel regression in animal models, but new strategies to inhibit/promote this signalling require enhanced understanding of VEGFR2 distribution and binding kinet-

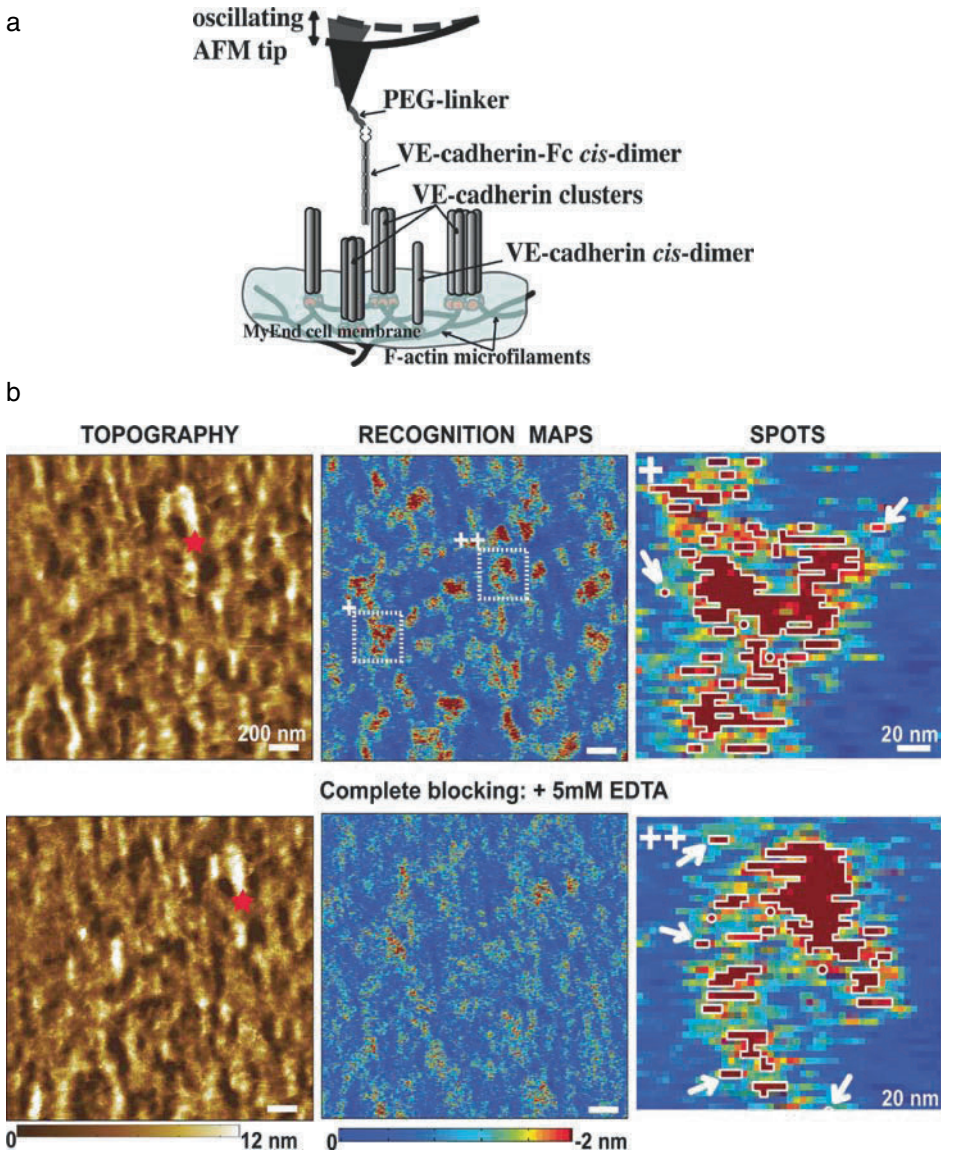


Figure 18.12. A. Scheme of dynamic recognition imaging for visualizing VE-cadherin–binding sites (here single *cis*-dimers and/or clusters) on a gently fixed MyEnd cell surface. **B.** Simultaneously recorded topography and recognition (TREC) images of a MyEnd cell surface obtained with VE-cadherin-Fc–functionalized tip. The addition of 5 mM ethylenediaminetetraacetic acid led to the disappearance of recognition clusters as the active VE-cadherin-Fc *cis*-dimers on the atomic force microscope (AFM) tip dissociated into inactive monomers, thereby abolishing specific VE-cadherin *trans*-interaction. After blocking experiments, topography remains unchanged, indicating that blocking does not affect membrane topography. Red stars in the topographic images show the AFM scanner lateral drift of ~5 nm/min. +, ++, Examples of recognition spots magnified from recognition image. Recognition areas are depicted by threshold analysis (threshold = – 1.7 nm) and bordered by white lines. Single VE-cadherin *cis*-dimers can be clearly detected (*arrows*). PEG, poly(ethylene glycol).

ics with biological or synthetic molecules. To identify VEGFR2 receptors, TREC imaging was performed on chemically fixed human umbilical vein endothelial cells (HUVEC) that endogenously express VEGFR2. By scanning the cell surface with a MAC tip tethered with monoclonal anti-VEGFR2 antibodies, dark regions with a diameter of 45.9 ± 8.9 nm were observed. These domains were assumed to be putative receptor locations that were consequently confirmed through the demonstration of binding specificity. The introduction of soluble anti-VEGFR2 to the imaging solution lead to a significant decrease in the number of recognitions sites over 60 min postblocking.

In addition, this nanoscale imaging showed that available VEGRFR2 receptors are nonuniformly distributed over 2-μm^2 regions, with a close spatial association to the underlying cortical cytoskeleton. This finding supports the current hypothesis that VEGRFR2 function is closely related to that of transmembrane integrin complexes transmitting force from the extracellular matrix to the actin cytoskeleton. Analysis of recognition images (1.5×1.5 μm^2) indicated about 9.8 sites per μm^2 and hence about 1.5×10^5 VEGRFR2 receptors per cell. This determination of receptors among individual cells is in reasonable agreement with cell population–averaged measurements through radiolabelled ligands.

TREC images were also acquired with MAC tips coated with antibody anti-Kv11.1 (against epitope tags present on the human ether-á-go-go-related gene [hERG] subunits) via PEG linker (Figure 18.13A). The recognition map represents an amplitude reduction due to specific binding between anti-Kv11.1 on the tip and epitope tags on the cell surface (dark spots) (Figure 18.13B). These dark spots (amplitude reduction up to 2 nm) were distributed nonuniformly and reflect microdomains with dimensions from ~30 up to ~350 nm, with a mean ± SD of 99 ± 81 nm ($n = 25$) for the long domain axis. During several subsequent rescans recognition maps of hERG channels remained unchanged. Hereafter, ErgTx1 was very slowly (50 μL/min) injected in the fluid cell while scanning the same sample. After the first and second injection of ErgTx1 (concentration of ~400 nM), no visual change in the recognition maps was observed. However, the recognition clusters partly disappeared after the third injection of ErgTx1 (~1 μM), whereas no change in the topography image was observed. The specific binding between anti-Kv11.1 and cellular surface was abolished when free ErgTx1 molecules bound to the hERG channels and thus blocked the antibody access to interact with epitope tags on hERG subunits. The topography of a scanned cell surface areas shows a complex picture of linear and branched filamentous structures with some globular features. Most domains are found to be located near and between filaments. TREC results suggest that ErgTx1 not only interacts with the extracellular surface of the pore domain (S5–S6), but also might interact with the voltage-sensing domains (S1–S4) of the hERG K^+ channel. These findings indicate a possibly new binding site of ErgTx1 in the voltage-sensing domain of the hERG K^+ channel. TREC is a suitable method for obtaining information about the structure and function of hERG K^+ channels on the cellar surface. The outcome of this study reveals that voltage-sensing domains (S1–S4) of hERG K^+ channel might be one of the binding sites of ErgTx1.

18.5. Immunogold Imaging

Immunogold labels, well known in electron and optical microscopy, may also be useful for discriminating specific molecules on a biological sample. The idea is to first incubate the sample with monoclonal antibodies directed against specific constituents, then further incubate it with the corresponding gold-conjugated secondary antibodies, and finally image the

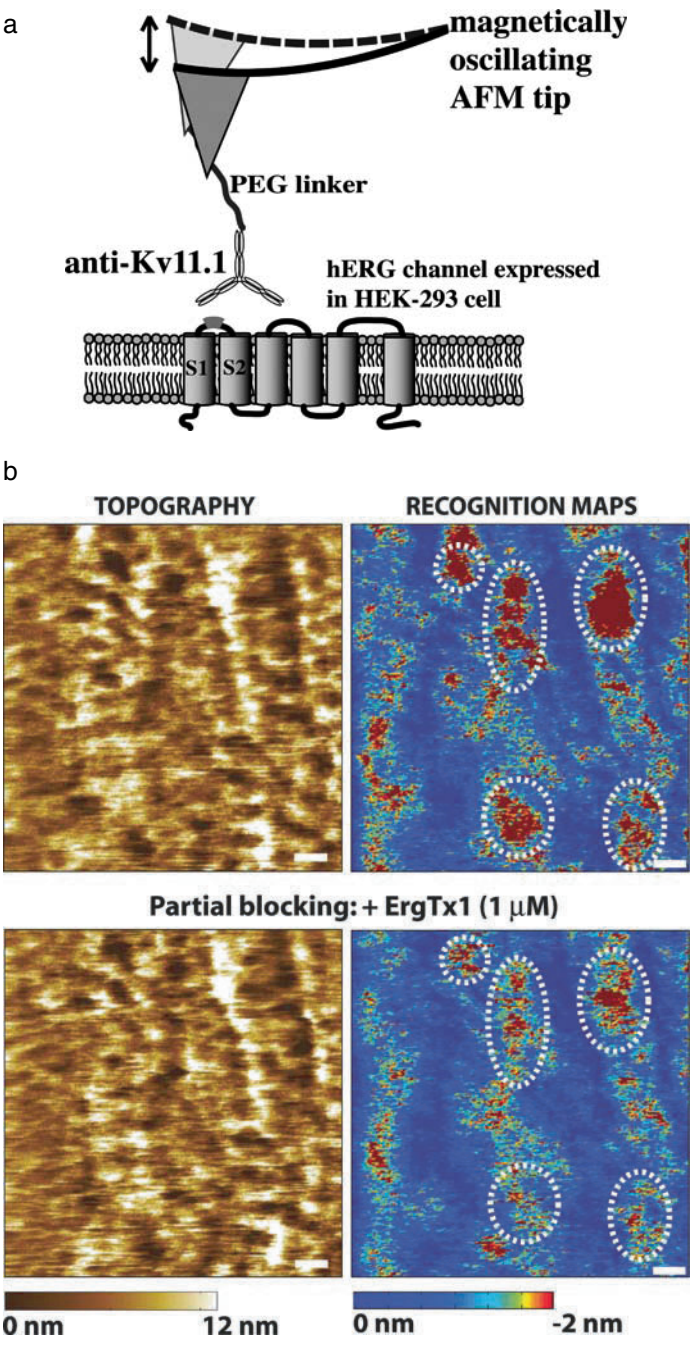


Figure 18.13. A. Schematic representation of recognition imaging to visualize human ether-á-go-go-related gene (hERG) K^+ channels (here binding sites at extracellular epitope between the S1 and S2 domains of hERG subunit [*light grey*]) on a gently fixed hERG HEK-293 cell surface. **B.** Topography and recognition imaging maps obtained on a hERG HEK-293 cell surface with anti-Kv11.1–functionalized tip. The most-pronounced recognition clusters are indicated by white lines. The presence of 400 nM ErgTx1 had practically no effect on the recognition map (data not shown), whereas the recognition clusters partly disappeared after the third injection of ErgTx1 (~1 μM). Blocking experiments do not affect membrane topography. Scale bars on all images are 170 nm. AFM, atomic force microscope; PEG, poly(ethylene glycol).

gold particles using AFM topographic imaging. The feasibility of using immunogold labels as cell-surface markers in AFM studies was first demonstrated on human lymphocytes [79]. Individual immunogold particles were clearly made visible on the cell surface, thus determining the location of antigens. More recently, a similar AFM-based immunogold technique revealed types I and II collagen fibers on rat fibroblasts and human chondrosarcoma cells [80]. Although well designed, these studies were performed in the dried state, which shows no clear benefit in using AFM over conventional immunogold electron microscopy.

Recently, this method was applied to hydrated bacteria, with the aim of localizing specific cell wall constituents (Figure 18.14) [15]. *M. bovis* BCG cells were treated with drugs to expose lipoarabinomannan (LAM) on the cell surface, then incubated with monoclonal anti-LAM antibodies, and further incubated with the corresponding gold-conjugated secondary antibodies. Images were obtained for native and treated cells, comparing contact and tapping (phase) modes. Gold particles were never seen in contact mode, emphasizing the need to use tapping (phase) mode for such *in situ* immunogold studies. With tapping mode, the surface

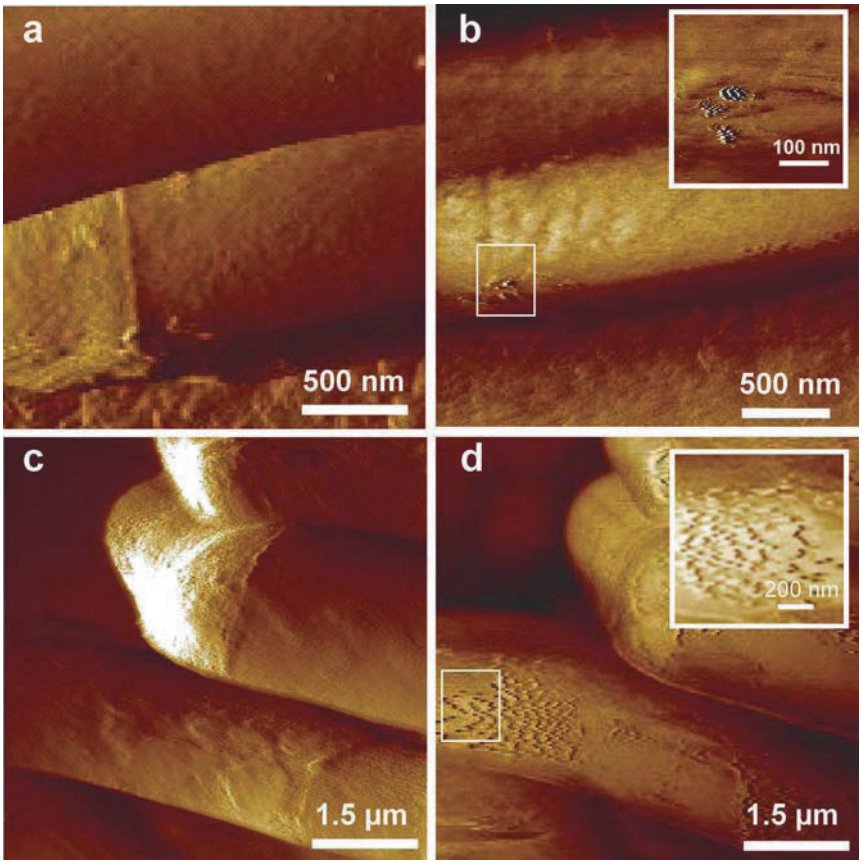


Figure 18.14. Localizing specific cell surface constituents by using immunogold atomic force microscopy. Contact mode (deflection) images **(A, C)** and tapping mode (phase) images **(B, D)** of immunogold-labelled *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) cells. **A, B.** Native cells. **C, D.** Cells treated for 24 hr with isoniazid. All cells were incubated with monoclonal anti-lipoarabinomannan (LAM) antibodies, followed by another incubation with the corresponding gold-conjugated secondary antibodies. (Reprinted with permission from Alsteens et al. [15].)

of native cells showed essentially no labelling, indicating that LAM was not exposed at the surface. This finding was consistent with the uniform distribution of hydrophobic mycolic acids measured on native cells. By contrast, drug-treated cells revealed a substantial coverage of gold particles, indicating that LAM was exposed. Particles were poorly resolved and rather fuzzy in tapping mode, presumably reflecting energy dissipation of the oscillating tip. These observations, which correlate with topographic and CFM data, demonstrate that antimycobacterial drugs lead to the massive exposure of LAM at the cell surface, thereby contributing to our understanding of their action modes. Hence, tapping mode imaging of immunogold labels is a valuable approach for localizing specific antigens on biosurfaces.

Acknowledgments

This work was supported by the National Foundation for Scientific Research (FNRS), the Université Catholique de Louvain (Fonds Spéciaux de Recherche), the Région Wallonne, the Federal Office for Scientific, Technical and Cultural Affairs (Interuniversity Poles of Attraction Programme), and the Research Department of the Communauté Française de Belgique (Concerted Research Action). YD and DA are Research Associate and Research Fellow of the FRS-FNRS, respectively. This work was further supported by the FFG program of the European Union projects Tips4Cells (LSHG-CT2005-512101), NASSAP (EC-STREP-13532), BioLightTouch (028181), and IMMUNANOMAP (MRTN-CT-2006-035946). AE and LC are Research Associates of the BioLigthTouch, and JT is Research Associate of the NASSAP project.

References

1. Dufrêne, Y. F. 2008. Towards nanomicrobiology using atomic force microscopy. Nat. Rev. Microbiol. 6:674–680.
2. Engel, A., and D. J. Muller. 2000. Observing single biomolecules at work with the atomic force microscope. Nat. Struct. Biol. 7:715–718.
3. Gerber, C., and H. P. Lang. 2006. How the doors to the nanoworld were opened. Nat. Nanotechnol. 1:3–5.
4. Hinterdorfer, P., and Y. F. Dufrene. 2006. Detection and localization of single molecular recognition events using atomic force microscopy. Nat. Methods 3:347–355.
5. Müller, D. J., K. T. Sapra, S. Scheuring, A. Kedrov, P. L. Frederix, D. Fotiadis, and A. Engel. 2006. Single-molecule studies of membrane proteins. Curr. Opin. Struct. Biol. 16:489–495.
6. Burnham, N. A., X. Chen, C. S. Hodges, G. A. Matei, E. J. Thoreson, C. J. Roberts, M. C. Davies, and S. J. B. Tendler. 2003. Comparison of calibration methods for atomic-force microscopy cantilevers. Nanotechnology 14:1–6.
7. Frisbie, C. D., L. F. Rozsnyai, A. Noy, M. S. Wrighton, and C. M. Lieber. 1994. Functional-group imaging by chemical force microscopy. Science 265:2071–2074.
8. Noy, A. 2006. Chemical force microscopy of chemical and biological interactions. Surf. Interface Anal. 38:1429–1441.
9. Dufrêne, Y. F. 2008. Atomic force microscopy and chemical force microscopy of microbial cells. Nat. Protocols 3:1132–1138.
10. Alsteens, D., E. Dague, P. G. Rouxhet, A. R. Baulard, and Y. F. Dufrêne. 2007. Direct measurement of hydrophobic forces on cell surfaces using AFM. Langmuir 23:11977–11979.
11. Vezenov, D. V., A. Noy, L. F. Rozsnyai, and C. M. Lieber. 1997. Force titrations and ionization state sensitive imaging of functional groups in aqueous solutions by chemical force microscopy. J. Am. Chem. Soc. 119: 2006–2015.
12. Ahimou, F., F. A. Denis, A. Touhami, and Y. F. Dufrene. 2002. Probing microbial cell surface charges by atomic force microscopy. Langmuir 18:9937–9941.

13. Dague, E., D. Alsteens, J. P. Latge, C. Verbelen, D. Raze, A. R. Baulard, and Y. F. Dufrêne. 2007. Chemical force microscopy of single live cells. Nano Lett. 7:3026–3030.
14. Dague, E., D. Alsteens, J. P. Latgé, and Y. F. Dufrêne. 2008. High-resolution cell surface dynamics of germinating *Aspergillus fumigatus* conidia. Biophys. J. 94:656–660.
15. Alsteens, D., C. Verbelen, E. Dague, D. Raze, A. R. Baulard, and Y. F. Dufrêne. 2008. Organization of the mycobacterial cell wall: A nanoscale view. Eur. J. Physiol. 456:117–125.
16. Florin, E. L., V. T. Moy, and H. E. Gaub. 1994. Adhesion forces between individual ligand receptor pairs. Science 264:415–417.
17. Lee, G. U., D. A. Kidwell, and R. J. Colton. 1994. Sensing discrete streptavidin biotin interactions with atomic force microscopy. Langmuir 10:354–357.
18. Fritz, J., A. G. Katopodis, F. Kolbinger, and D. Anselmetti. 1998. Force-mediated kinetics of single P-selectin ligand complexes observed by atomic force microscopy. Proc. Natl. Acad. Sci. USA. 95:12283–12288.
19. Lee, G. U., L. A. Chrisey, and R. J. Colton. 1994. Direct measurement of the forces between complementary strands of DNA. Science 266:771–773.
20. Touhami, A., B. Hoffmann, A. Vasella, F. A. Denis, and Y. F. Dufrene. 2003. Probing specific lectin–carbohydrate interactions using atomic force microscopy imaging and force measurements. Langmuir 19:1745–1751.
21. Dammer, U., O. Popescu, P. Wagner, D. Anselmetti, H. J. Guntherodt, and G. N. Misevic. 1995. Binding strength between cell-adhesion proteoglycans measured by atomic force microscopy. Science 267:1173–1175.
22. Berquand, A., N. Xia, D. G. Castner, B. H. Clare, N. L. Abbott, V. Dupres, Y. Adriaensen, and Y. F. Dufrêne. 2005. Antigen binding forces of single antilysozyme Fv fragments explored by atomic force microscopy. Langmuir 21:5517–5523.
23. Dupres, V., F. D. Menozzi, C. Locht, B. H. Clare, N. L. Abbott, S. Cuenot, C. Bompard, D. Raze, and Y. F. Dufrêne. 2005. Nanoscale mapping and functional analysis of individual adhesins on living bacteria. Nat. Methods 2:515–520.
24. Hinterdorfer, P., W. Baumgartner, H. J. Gruber, K. Schilcher, and H. Schindler. 1996. Detection and localization of individual antibody–antigen recognition events by atomic force microscopy. Proc. Natl. Acad. Sci. USA 93:3477–3481.
25. Allen, S., X. Y. Chen, J. Davies, M. C. Davies, A. C. Dawkes, J. C. Edwards, C. J. Roberts, J. Sefton, S. J. B. Tendler, and P. M. Williams. 1997. Detection of antigen–antibody binding events with the atomic force microscope. Biochemistry 36:7457–7463.
26. Ong, Y. L., A. Razatos, G. Georgiou, and M. M. Sharma. 1999. Adhesion forces between *E-coli* bacteria and biomaterial surfaces. Langmuir 15:2719–2725.
27. Bowen, W. R., N. Hilal, R. W. Lovitt, and C. J. Wright. 1998. Direct measurement of the force of adhesion of a single biological cell using an atomic force microscope. Colloids Surf. A Physicochem. Eng. Asp. 136:231–234.
28. Lower, S. K., M. F. Hochella, and T. J. Beveridge. 2001. Bacterial recognition of mineral surfaces: Nanoscale interactions between *Schewanella* and alpha-FeOOH. Science 292:1360–1363.
29. Benoit, M., D. Gabriel, G. Gerisch, and H. E. Gaub. 2000. Discrete interactions in cell adhesion measured by single-molecule force spectroscopy. Nat. Cell Biol. 2:313–317.
30. Zhang, X. H., E. Wojcikiewicz, and V. T. Moy. 2002. Force spectroscopy of the leukocyte function-associated antigen-1/intercellular adhesion molecule-1 interaction. Biophys. J. 83:2270–2279.
31. Ebner, A., P. Hinterdorfer, and H. J. Gruber. 2007. Comparison of different aminofunctionalization strategies for attachment of single antibodies to AFM cantilevers. Ultramicroscopy 107:922–927.
32. Ros, R., F. Schwesinger, D. Anselmetti, M. Kubon, R. Schafer, A. Pluckthun, and L. Tiefenauer. 1998. Antigen binding forces of individually addressed single-chain Fv antibody molecules. Proc. Natl. Acad. Sci. USA 95:7402–7405.
33. Strunz, T., K. Oroszlan, R. Schafer, and H. J. Guntherodt. 1999. Dynamic force spectroscopy of single DNA molecules. Proc. Natl. Acad. Sci. USA 96:11277–11282.
34. Schumakovitch, I., W. Grange, T. Strunz, P. Bertoncini, H. J. Guntherodt, and M. Hegner. 2002. Temperature dependence of unbinding forces between complementary DNA strands. Biophys. J. 82:517–521.
35. Baumgartner, W., P. Hinterdorfer, W. Ness, A. Raab, D. Vestweber, H. Schindler, and D. Drenckhahn. 2000. Cadherin interaction probed by atomic force microscopy. Proc. Natl. Acad. Sci. USA 97:4005–4010.
36. Baumgartner, W., H. J. Gruber, P. Hinterdorfer, and D. Drenckhahn. 2000. Affinity of *trans*-interacting VE-cadherin determined by atomic force microscopy. Single Mol. 1:119–122.
37. Ludwig, M., W. Dettmann, and H. E. Gaub. 1997. Atomic force microscope imaging contrast based on molecular recognition. Biophys. J. 72:445–448.

38. Grandbois, M., W. Dettmann, M. Benoit, and H. E. Gaub. 2000. Affinity imaging of red blood cells using an atomic force microscope. J. Histochem. Cytochem. 48:719–724.
39. Lehenkari, P. P., G. T. Charras, A. Nykänen, and M. A. Horton. 2000. Adapting atomic force microscopy for cell biology. Ultramicroscopy 82:289–295.
40. Almqvist, N., R. Bhatia, G. Primbs, N. Desai, S. Banerjee, and R. Lal. 2004. Elasticity and adhesion force mapping reveals real-time clustering of growth factor receptors and associated changes in local cellular rheological properties. Biophys. J. 86:1753–1762.
41. Gilbert, Y., M. Deghorain, L. Wang, B. Xu, P. D. Pollheimer, H. J. Gruber, J. Errington, B. Hallet, X. Haulot, C. Verbelen, P. Hols, and Y. F. Dufrene. 2007. Single-molecule force spectroscopy and imaging of the vancomycin/D-Ala-D-Ala interaction. Nano Lett. 7:796–801.
42. Camesano, T. A., Y. Liu, and M. Datta. 2007. Measuring bacterial adhesion at environmental interfaces with single-cell and single-molecule techniques. Adv. Water Resour. 30:1470–1491.
43. Alsteens, D., V. Dupres, K. Mc Evoy, L. Wildling, H. J. Gruber, and Y. F. Dufrêne. 2008. Structure, cell wall elasticity and polysaccharide properties of living yeast cells, as probed by AFM. Nanotechnology, 19:384005(9pp).
44. Francius, G., S. Lebeer, D. Alsteens, L. Wildling, H. J. Gruber, P. Hols, S. De Keersmaecker, J. Vanderleyden, and Y. F. Dufrêne. 2008. Detection, localization and conformational analysis of single polysaccharide molecules on live bacteria. ACS Nano, 2:1921–1929.
45. Sotres, J., A. Lostao, L. Wildling, A. Ebner, C. Gomez-Moreno, H. J. Gruber, P. Hinterdorfer, and A. M. Baro. 2008. Unbinding molecular recognition force maps of localized single receptor molecules by atomic force microscopy. Chemphyschem 9:590–599.
46. Ebner, A., F. Kienberger, G. Kada, C. M. Stroh, M. Geretschlager, A. S. M. Kamruzzahan, L. Wildling, W. T. Johnson, B. Ashcroft, J. Nelson, S. M. Lindsay, H. J. Gruber, and P. Hinterdorfer. 2005. Localization of single avidin–biotin interactions using simultaneous topography and molecular recognition imaging. Chemphyschem 6:897–900.
47. Kienberger, F., A. Ebner, H. J. Gruber, and P. Hinterdorfer. 2006. Molecular recognition imaging and force spectroscopy of single biomolecules. Acc. Chem. Res. 39:29–36.
48. Stroh, C., H. Wang, R. Bash, B. Ashcroft, J. Nelson, H. Gruber, D. Lohr, S. M. Lindsay, and P. Hinterdorfer. 2004. Single-molecule recognition imaging microscopy. Proc. Natl. Acad. Sci. USA 101:12503–12507.
49. Chtcheglova, L. A., F. Atalar, U. Ozbek, L. Wildling, A. Ebner, and P. Hinterdorfer. 2008. Localization of the ergtoxin-1 receptors on the voltage sensing domain of hERG K+ channel by AFM recognition imaging. Pflugers Arch. Eur. J. Physiol. 456:247–254.
50. Chtcheglova, L. A., J. Waschke, L. Wildling, D. Drenckhahn, and P. Hinterdorfer. 2007. Nano-scale dynamic recognition imaging on vascular endothelial cells. Biophys. J. 93:L11–13.
51. Ebner, A., D. Nikova, T. Lange, J. Haeberle, S. Falk, A. Duebbers, R. Bruns, P. Hinterdorfer, H. Oberleithner, and H. Schillers 2008. Determination of CFTR densities in erythrocyte plasma membranes using recognition imaging. Nanotechnology19:384017.
52. Stroh, C. M., A. Ebner, M. Geretschlager, G. Freudenthaler, F. Kienberger, A. S. M. Kamruzzahan, S. J. Smith-Gill, H. J. Gruber, and P. Hinterdorfer. 2004. Simultaneous topography and recognition imaging using force microscopy. Biophys. J. 87:1981–1990.
53. Ebner, A., L. Wildling, R. Zhu, C. Rankl, T. Haselgrubler, P. Hinterdorfer, and H. J. Gruber. 2007. Funtionalization of probe tips and supports for single-molecule recognition force microscopy. Top Curr. Chem 285:29–76.
54. Liu, Y. Z., S. H. Leuba, and S. M. Lindsay. 1999. Relationship between stiffness and force in single molecule pulling experiments. Langmuir 15:8547–8548.
55. Lin, L., H. Wang, Y. Liu, H. Yan, and S. Lindsay. 2006. Recognition imaging with a DNA aptamer. Biophys. J. 90:4236–4238.
56. Lin, L., D. Hom, S. M. Lindsay, and J. C. Chaput. 2007. *In vitro* selection of histone h4 aptamers for recognition imaging microscopy. J. Am. Chem. Soc. 129:14568–14569.
57. Sleytr, U. B., C. Huber, N. Ilk, D. Pum, B. Schuster, and E. M. Egelseer. 2007. S-layers as a tool kit for nanobiotechnological applications. FEMS Microbiol. Lett. 267:131–144.
58. Sleytr, U. B., M. Sara, D. Pum, and B. Schuster. 2001. Characterization and use of crystalline bacterial cell surface layers. Prog. Surf. Sci. 68:231–278.
59. Sleytr, U. B., and T. J. Beveridge. 1999. Bacterial S-layers. Trends Microbiol. 7:253–260.
60. Sleytr, U. B., P. Messner, D. Pum, and M. Sara. 1999. Crystalline bacterial cell surface layers (S layers): From supramolecular cell structure to biomimetics and nanotechnology. Angew. Chem. Int. Ed. 38:1035–1054.
61. Sleytr, U. B., E. M. Egelseer, N. Ilk, D. Pum, and B. Schuster. 2007. S-Layers as a basic building block in a molecular construction kit. FEBS J. 274:323–334.

62. Voss, S., and A. Skerra. 1997. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. Protein Eng. 10:975–982.
63. Ilk, N., P. Kosma, M. Puchberger, E. M. Egelseer, H. F. Mayer, U. B. Sleytr, and M. Sára. 1999. Structural and functional analyses of the secondary cell wall polymer of *Bacillus sphaericus* CCM 2177 that serves as an S-layer specific anchor. J. Bacteriol. 181:7643–7646.
64. Pleschberger, M., A. Neubauer, E. M. Egelseer, S. Weigert, B. Lindner, U. B. Sleytr, S. Muyldermans, and M. Sara. 2003. Generation of a functional monomolecular protein lattice consisting of an S-layer fusion protein comprising the variable domain of a camel heavy chain antibody Bioconjug. Chem. 14:440–448.
65. Tang, J., A. Ebner, N. Ilk, H. Lichtblau, C. Huber, R. Zhu, D. Pum, M. Leitner, V. Pastushenko, H. J. Gruber, U. B. Sleytr, and P. Hinterdorfer. 2007. High-affinity tags fused to S-layer proteins probed by atomic force microscopy. Langmuir 24:1324–1329.
66. Ebner, A., F. Kienberger, G. Kada, C. M. Stroh, M. Geretschlager, A. S. Kamruzzahan, L. Wildling, W. T. Johnson, B. Ashcroft, J. Nelson, S. M. Lindsay, H. J. Gruber, and P. Hinterdorfer. 2005. Localization of single avidin–biotin interactions using simultaneous topography and molecular recognition imaging. Chemphyschem 6:897–900.
67. Schwiebert, E. M., D. J. Benos, and C. M. Fuller. 1998. Cystic fibrosis: A multiple exocrinopathy caused by dysfunctions in a multifunctional transport protein. Am. J. Med. 104:576–590.
68. Braunstein, G. M., R. M. Roman, J. P. Clancy, B. A. Kudlow, A. L. Taylor, V. G. Shylonsky, B. Jovov, K. Peter, T. Jilling, Ismailov, II, D. J. Benos, L. M. Schwiebert, J. G. Fitz, and E. M. Schwiebert. 2001. Cystic fibrosis transmembrane conductance regulator facilitates ATP release by stimulating a separate ATP release channel for autocrine control of cell volume regulation. J. Biol. Chem. 276:6621–6630.
69. Dupuit, F., N. Kalin, S. Brezillon, J. Hinnrasky, B. Tummler, and E. Puchelle. 1995. CFTR and differentiation markers expression in non-CF and delta-F-508 homozygous CF nasal epithelium. J. Clin. Invest. 96:1601–1611.
70. Kalin, N., A. Claass, M. Sommer, E. Puchelle, and B. Tummler. 1999. Delta F508 CFTR protein expression in tissues from patients with cystic fibrosis. J. Clin. Invest. 103:1379–1389.
71. Lange, T., P. Jungmann, J. Haberle, S. Falk, A. Duebbers, R. Bruns, A. Ebner, P. Hinterdorfer, H. Oberleithner, and H. Schillers. 2006. Reduced number of CFTR molecules in erythrocyte plasma membrane of cystic fibrosis patients. Mol. Membr. Biol. 23:317–323.
72. Sprague, R. S., M. L. Ellsworth, A. H. Stephenson, M. E. Kleinhenz, and A. J. Lonigro. 1998. Deformation-induced ATP release from red blood cells requires CFTR activity. Am. J. Physiol. Heart Circ. Physiol. 275:H1726–H1732.
73. Sterling, K. M., S. Shah, R. J. Kim, N. I. F. Johnston, A. Y. Salikhova, and E. H. Abraham. 2004. Cystic fibrosis transmembrane conductance regulator in human and mouse red blood cell membranes and its interaction with ecto-apyrase. J. Cell. Biochem. 91:1174–1182.
74. Stumpf, A., J. Almaca, K. Kunzelmann, K. Wenners-Epping, S. M. Huber, J. Haberle, S. Falk, A. Duebbers, M. Walte, H. Oberleithner, and H. Schillers. 2006. IADS, a decomposition product of DIDS activates a cation conductance in *Xenopus* oocytes and human erythrocytes: New compound for the diagnosis of cystic fibrosis. Cell. Physiol. Biochem. 18:243–252.
75. Stumpf, A., K. Wenners-Epping, M. Walte, T. Lange, H. G. Koch, J. Haberle, A. Dubbers, S. Falk, L. Kiesel, D. Nikova, R. Bruns, H. Bertram, H. Oberleithner, and H. Schillers. 2006. Physiological concept for a blood based CFTR test. Cell. Physiol. Biochem. 17:29–36.
76. Verloo, P., C. H. M. Kocken, A. Van der Wel, B. C. Tilly, B. M. Hogema, M. Sinaasappel, A. W. Thomas, and H. R. De Jonge. 2004. *Plasmodium falciparum*-activated chloride channels are defective in erythrocytes from cystic fibrosis patients. J. Biol. Chem. 279:10316–10322.
77. Vincent, P. A., K. Xiao, K. M. Buckley, and A. P. Kowalczyk. 2004. VE-cadherin: Adhesion at arm's length. AJP Cell Physiol. 286:C987–997.
78. Lee, S., J. Mandic, and K. J. Van Vliet. 2007. Chemomechanical mapping of ligand–receptor binding kinetics on cells. Proc. Natl. Acad. Sci. USA 104:9609–9614.
79. Putman, C. A. J., B. G. Degrooth, P. K. Hansma, N. F. Vanhulst, and J. Greve. 1993. Immunogold labels: Cell-surface markers in atomic force microscopy. Ultramicroscopy 48:177–182.
80. Arntz, Y., L. Jourdainne, G. Greiner-Wacker, S. Rinckenbach, J. Ogier, J. C. Voegel, P. Lavalle, and D. Vautier. 2006. Immunogold detection of types I and II chondrocyte collagen fibrils: An *in situ* atomic force microscopic investigation. Microsc. Res. Tech. 69:283–290.

19

Atomic Force Microscopy of Protein–Protein Interactions

Xiaohui Zhang, Felix Rico, Amy J. Xu, and Vincent T. Moy

Abstract Since its invention in 1986, the atomic force microscope (AFM) has emerged as a flexible and powerful tool for exploring a variety of biological processes, including cell adhesion, protein folding, and protein–protein interactions. This review focuses on the application of the AFM to studies of protein–protein interactions. It describes the commonly used methodologies and reviews the theoretical framework used to analyze single-molecule protein–protein unbinding measurements. Finally, the chapter summarizes recent progress in the field and shows that the AFM provides an excellent tool for probing interactions on the cell surface and for understanding the energy landscapes that govern the dynamics of protein interactions.

19.1. Introduction

In recent years, new technologies have been developed to directly measure the dynamic processes of biological interactions at high spatial and temporal resolutions. In particular, single-molecule approaches such as atomic force microscopy and optical tweezers, with their ultrahigh resolutions down to subnanometer and submicrosecond levels, have opened new avenues to measuring individual molecular interactions in real time (1,2). This chapter focuses mainly on the application of the atomic force microscope (AFM) (3,4), a widely used single-molecule tool, to the study of protein interactions.

Proper protein–protein interactions are essential to normal physiological function in biological systems (5). Until recently, studies of protein–protein interactions usually involved biochemical methods that determine binding affinities or rate constants. Although these bulk measurements offer averaged estimates of interaction rates and free energies, they often miss

X. Zhang • Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China
F. Rico, V. T. Moy • Department of Physiology and Biophysics, University of Miami Miller School of Medicine, 1600 NW 10th Avenue, Miami, FL 33136, USA
A. J. Xu • MD/PhD program, HMS-MIT Division of Health Sciences and Technology, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_19,

important features of interaction dynamics that uncover various intermediate states or alternative reaction pathways. Single-molecule approaches, however, avoid the intrinsic ensemble average of bulk methods and make it possible to follow the trajectories of individual interacting molecules in real time (1,6). Single-molecule methods, therefore, have become better tools for understanding the structure of the energy landscape governing the association, dissociation, and intermediate states of protein interactions (2,7). Moreover, on the cell surface, external forces continuously stretch molecules. For example, traction forces induce the cell surface adhesion receptors to undergo cycles of adhesion and de-adhesion as the cells migrate (8). Conventional techniques cannot reveal the influence of these internal and external forces associated with protein–protein interactions (9). Now, however, by using such force-measuring single-molecule tools as the AFM, optical tweezers, and biomembrane force probe, one can directly measure the forces holding protein complexes together and can access the stresses and strains associated with these reactions (2). Although each method has its merits, this chapter only discusses the use of the AFM in measuring protein interactions.

19.2. AFM Experimentation

19.2.1. AFM Principles

The essential components of an AFM are the cantilever; the tip; the sample stage; the piezoelectric translator, which accurately displaces the sample stage or cantilever; and the optical deflection system, comprising a laser and a photodetector, which records the changes in cantilever deflection (Figure 19.1).

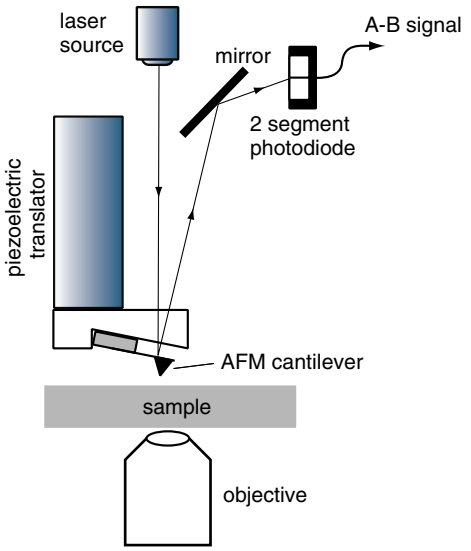


Figure 19.1. Schematic of the atomic force microscope (AFM). A beam of laser light is directed onto the cantilever and is reflected into a split photodiode. The difference (A-B) signal reports the forces exerted on the tip. Because the cantilever obeys Hooke's law for small displacements, the interaction force between the tip and the sample can be determined from the photodiode signal after a proper calibration of the spring constant of the cantilever.

Originally designed as an imaging tool (3), the AFM acquires topographic images by scanning the specimen with a flexible cantilever. In imaging mode, the AFM scans its tip, which is mounted to the end of a soft cantilever, over the specimen in the x, y plane (Figure 19.1). A piezoelectric element constantly corrects the position of the cantilever while scanning, to keep the deflection constant. Atomic-level resolution is acquired by translating the positioning of the cantilever into a topographic image of the sample surface.

Two formats are used to probe sample surfaces: a traditional design, which couples the sample stage to the movement of a three-dimensional piezoelectric translator, and a newer design, which instead scans the cantilever mounted on a piezo over a fixed sample (Figure 19.1). In the second approach, the AFM head is usually placed on the stage of an inverted optical microscope, allowing simultaneous acquisition of optical and AFM images. Therefore, the latter design is more feasible for working with living cells because it allows simultaneous visualization of cell shape and morphology.

The difference (A-B) signal from the photodetector reports the bending of the AFM cantilever due to forces exerted on the tip. The cantilever must be calibrated in order to translate its deflection into units of force. To this end, it is necessary to determine the spring constant of the cantilever, k_C. Different techniques are used for cantilever calibration (10) — measuring the cantilever's deflection (x) with the application of a constant known force (F) (11), or measuring of the cantilever's thermal oscillation resonant frequency (12). The most popular method for cantilever calibration, based on Hutter and Bechhoefer (12), treats the cantilever as a simple harmonic oscillator whose power spectrum of thermal fluctuations can be used to derive the spring constant. Since each vibration mode of the cantilever receives the thermal energy commensurate with one degree of freedom, $k_B T/2$, the measured variance of the cantilever's deflection $\langle x^2 \rangle$ can be used to calculate its spring constant, that is, $\frac{1}{2} k_B T = \frac{1}{2} k_C \langle x^2 \rangle$, where k_B and T are Boltzmann's constant and the absolute temperature, respectively.

19.2.2. AFM Measurement of Single-Molecule Interactions

The AFM can also be operated in force scan mode to measure interactions between two apposing surfaces at the single-molecule level. In this mode, a soft cantilever, usually with a spring constant of 10–100 mN/m and functionalized with a protein ligand, is positioned a few microns above a sample surface decorated with the molecule (such as a protein) of interest. The pyramid-shaped AFM tip, with a rounded apex of 10–50 nm in diameter, limits the contact between the two interacting surfaces so that it is possible to restrict the interaction between the surfaces to a single protein–protein bond (13). Figure 19.2 illustrates how unbinding experiments are performed.

When the cantilever is withdrawn from the surface, the unbinding forces of an individual protein–protein complex are determined experimentally from retraction curves by measuring the force jumps relative to the "zero-force" offset (Figure 19.2). In many experiments, force curves are obtained at cantilever velocities that can be as high as tens of microns per second. Due to the drag force exerted by the surrounding medium, the cantilever experiences a constant bending proportional to the applied velocity during its free approach and retraction. This hydrodynamic effect adds a systematic bias to the measured unbinding forces that has to be corrected when high velocities are applied. The drag force depends not only on the cantilever velocity and on the viscosity of the medium, which varies with temperature, but also on the geometry of the cantilever, the topography of the sample, and the relative separation between cantilever and sample (14,15). It is, thus, crucial to correct for this effect by systematically determining the viscous drag at each experimental condition. A straightforward

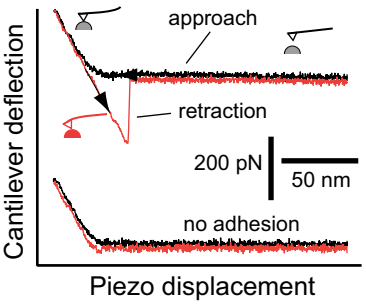


Figure 19.2. A representative unbinding trace. A representative atomic force microscope (AFM) force scan (i.e., optical signal vs. piezo displacement) under force scan mode. The two interacting surfaces are the AFM tip functionalized with streptavidin and an agarose bead with immobilized biotin. When the measurements started, the piezo first expanded and moved the agarose bead closer to the cantilever. Before the bead and cantilever touched, the deflection/force signal maintained a stable baseline level corresponding to "zero force." When the bead touched the cantilever, the cantilever bent upward, leading to a positive change of the deflection/force signal. The piezo ceased to expand when a preset compression force was reached (200 pN in this case). During the preset contact duration, the streptavidin and biotin had time to interact with one another. Adhesive contact created during this time was revealed by a downward deflection of the cantilever when the piezo started to contract and gradually moved the bead to its initial position. Further retraction movement of the bead resulted in a gradual increase of tension applied to the avidin–biotin bond(s) until the bond(s) were ruptured, indicated by a sharp vertical transition in the retraction trace, after which the cantilever returned to its original "zero-force" resting position.

approach to determining the viscous drag force is to measure the force shift just after the jump between the "zero force" values of the approach and retraction traces.

Using the force scan mode of an AFM, Lee et al. (16) and Florin et al. (13) directly measured the unbinding force of a single avidin–biotin interaction. Since the method was first reported in 1994, hundreds of different protein–protein interactions have been studied using this or similar methods. Table 19.1 provides a partial list of the unbinding forces and Bell–Evans model parameters of measured protein interaction pairs.

19.2.3. Tip and Sample Preparation

In studies of protein–protein interactions, the protein or ligand is immobilized on the surface of an AFM cantilever tip while its cognate binding partner is attached to a suitable substrate. After bringing the two surfaces into contact, the interaction force is acquired from the deflection of the cantilever during its withdrawal from the substrate. Therefore, proper tip and sample preparation is essential for the success of this type of study. The techniques commonly used to functionalize cantilevers involve either physisorption (13) or covalent coupling of the ligand to the tip via an extended linker (17). The linker between the tip and the ligand lends greater mobility and access to receptors on the surface being probed. Figure 19.3 outlines both methods for functionalizing tips.

Similar to tip functionalization, samples may be immobilized to surfaces by using physisorption or covalent coupling methods. Mica, polystyrene, and glass are widely used substrates for physisorption of proteins. Commonly used covalent coupling methods include thiol surface chemistry after gold deposition and cross-linking molecules of interest with a silanized surface.

Although the protocols for tip and surface functionalization are well established, one should bear in mind that coupling efficiency varies among different types of cantilevers or

Table 19.1. Summary of reported atomic force microscope unbinding studies

Ligand–Receptor Pair	Loading Rate (pN/sec)	Rupture forces (pN)	Barrier width (Å)	Ref.
$\alpha_5\beta_1$/fibronectin	50–50,000	40–170	0.9, 4	27
Aptamer/IgE	12,000–1,700,000	50–190	0.91, 2.54	53
Avidin-biotin	0.05–60,000	5–170	1.2, 3, 30	25
Azurin/cytochrome c551	7,000–150,000	50–150	1.4	54
Con A/D-mannose	400–5000	80–125	2.7	32
Digoxigenin/antibody	30–63,000	20–80	0.35, 1.15	55
GTPase Rap/impβ	300–80,000	40–90, 75–160[1]	N/A	30
LFA-1/ICAM-1	50–50,000	20–320	0.2, 2	26
LFA-1/ICAM-2	50–60,000	20–120	0.31–4.5	40
L-Selectin/PSGL-1[2]	10–100,000	20–160	0.6, 4	24
L-Selectin/sLeX[2]	10–100,000	10–180	0.6, 4	24
Mannuronan/AlgE 4	300–30,000	73–144	2.3	56
p53/azurin	3 nN/s	70	N/A	57
PDZ domain/peptide	3,800–140,000	40–220	0.4, 2.1	58
Plant lectin/asialofetuin	100–30,000	37–65	4–6	59
P-Selectin/PSGL-1	–[3]	110–170	2.5	60
P-selectin/PSGL-1	100–10,000	30–220	1.4	61
P-selectin/sLeX	70–100,000	20–220	0.8, 4.5	28
SfiI/DNA	2,100–630,000	25–100	1.8	62
Streptavidin–biotin	0.05–60,000	5–170	1.2, 5	25
TGF-β-1/receptor	100–13,000	40–200	0.73, 2.93	63
Transferrin(holo)/receptor	400–70,000	40–140	1.5, 9.3	29
Transferrin(apo)/receptor	500–40,000	25–40	8.1	29
VE-cadherin pair	–[3]	30–50	5.9	64
VLA-4/VCAM-1 (WT)	30–200,000	15–130	1, 5.5	65
VLA-4/VCAM-1 (D40A)	300–100,000	25–70	5.9	65
VLA-4/VCAM-1 (Q38G)	200–100,000	25–100	1.7, 5.8	65
VLA-4/VCAM-1 (L43K)	200–100,000	20–100	1.5, 5.7	65
VLA-4/VCAM-1 (D143A)	300–200,000	25–140	0.95, 5.8	65
Anti-γ-GT/γ-GT	N/A	131 ± 44	N/A	66
Antiferritin/ferritin	N/A	49 ± 10	N/A	67
Anti-HSA/HSA	N/A	240 ± 48	N/A	68
Anti-βhCG/βhCG	N/A	100 ± 47	N/A	69
Glycoprotein csA pair	N/A	23 ± 8	N/A	37
GroEL pair	N/A	420 ± 100	N/A	70
ICAM-1/anti-ICAM-1	N/A	100 ± 50	N/A	34
Insulin pair	N/A	1,340–1,350	N/A	71
Meromyosin/actin	N/A	15–25	N/A	72
Ocular mucin pair	N/A	100–4,000	N/A	73

Table 19.1. (continued)

Ligand–Receptor Pair	Loading Rate (pN/sec)	Rupture forces (pN)	Barrier width (Å)	Ref.
Proteoglycan pair	N/A	40 ± 15	N/A	74
$\alpha_5\beta_1$/GRGDSP peptide	N/A	32 ± 2	N/A	31
$\alpha_v\beta_3$/Osteopontin	N/A	50 ± 2	N/A	31

Con A, concanavalin A; βhCG, β subunit of human chorionic gonadotropin; GRGDSP, Gly-Arg-Gly-Asp-Ser-Pro; γ-GT, γ-glutamyl-transpeptidase; GTP, guanosine triphosphate; HAS, heat-stable antigen; ICAM-1, -2, intercellular adhesion molecules 1 and 2; IgE, immunoglobulin E; impβ, nuclear import receptor importin β1; LFA-1, leukocyte function-associated antigen-1; N/A, not available; PSGL-1, P-selectin glycoprotein ligand-1; TGF-β-1, transforming growth factor beta 1; VCAM-1, vascular cell adhesion molecule-1; VLA-4, very late antigen-4.

[1]The authors reported two populations of unbinding forces, reflecting the existence of two conformation states in the Rap/impβ complexes.

[2]Studies using biomembrane force probe.

[3]The authors reported unbinding forces versus different pulling velocities; we were unable to convert the velocity to loading rates.

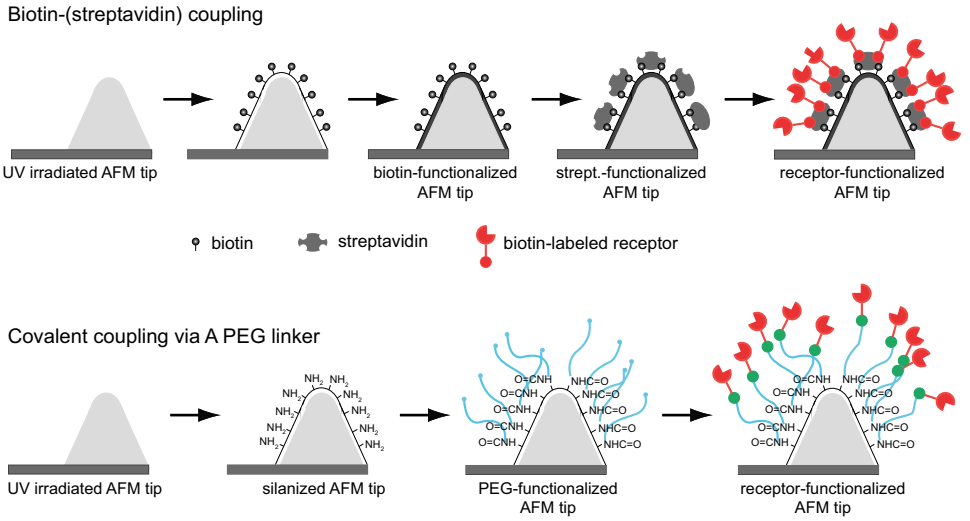


Figure 19.3. Schematics of surface chemistries commonly used for modifying atomic force microscope (AFM) tips. **A.** Functionalization of AFM tips with streptavidin using physisorption. **B.** Covalent coupling of proteins via a heterobifunctional poly(ethylene glycol) (PEG) cross-linker: The amine-reactive *N*-hydroxysuccinimide (NHS) end of the cross-linker reacts with amines on the silicon tip, yielding a stable amide bond, and the reactive 2-[pyridyldithio]propionate (PDP) group forms a bond with free thiols presented by cysteines in the protein, resulting in a stable disulfide bond. UV, ultraviolet.

surfaces, as well as among different molecules of interest. Thus, before performing force measurements, a detailed validation of the quality of functionalized tips and coupled samples is strongly recommended.

19.3. Determination of the Energy Landscape from the AFM Measurements

An AFM unbinding measurement (Figure 19.2) reveals the force required to break a ligand–receptor bond under a constant retraction velocity of the piezoelectric translator. For most of the measurement, the tension applied to break the bond increases linearly with time. The rate of change in tension, $rf = df/dt$, is known as the "loading rate" and can be experimentally controlled by varying the cantilever retraction velocity and/or the spring constant of the cantilever. For many ligand–receptor interactions, the unbinding force has been shown to change as the loading rate varies. In these studies, dependence of the unbinding forces on the measured loading rate can be used to reveal the energy landscape of the ligand–receptor complex.

The theoretical framework for understanding how a pulling force affects the dissociation of a protein–protein complex was first formulated by Bell in 1978 (18) and later expanded on by Evans and other researchers (19). Although recent refinements and generalizations have been developed (20–22), the Bell–Evans model can be seen as a first approximation to determining the energy landscape properties. The Bell–Evans model describes the influence of an external force on the rate of bond dissociation. This model is based on the conventional transition-state theory, in which a molecular complex needs to overcome an activation energy barrier before final separation. If only a single barrier dominates the dissociation process, the dissociation rate of this interaction is given by

$$k_{\text{off}} = \alpha \frac{k_{\text{B}}T}{h} \exp\left(\frac{-\Delta G^0}{k_{\text{B}}T}\right) \tag{19.1}$$

where ΔG^0 is the activation energy, T is the absolute temperature, k_{B} is Boltzmann's constant, h is Planck's constant, and α is a prefactor that characterizes the energy potential well. When the complex is exposed to a pulling force, the applied force adds a $-fx$ term to the potential of the system. If the potential barrier is steep, adding this term to the free energy reduces the activation barrier by approximately $f\gamma$, where γ is the width between the bound state and the transition state along the reaction coordinate. Thus, the force-dependent dissociation rate of the complex is given by

$$k_{\text{off}}(f) = \alpha \frac{k_{\text{B}}T}{h} \exp\left(\frac{-(\Delta G^0 - f\gamma)}{k_{\text{B}}T}\right) = k^o \exp\left(\frac{f\gamma}{k_{\text{B}}T}\right) \tag{19.2}$$

where k^o is the unstressed dissociation rate. Hence, the model predicts that the dissociation rate of the complex increases exponentially with a pulling force. The two parameters k^o and γ are often referred to as the Bell model parameters. These two parameters characterize the energy potential of the protein–protein complex. The depth of the potential is characterized by k^o, and γ characterizes the width of the potential and dictates the force resistance of a molecular complex. If a complex has a small γ (i.e., the activation potential is narrow), then an external force will have less effect on its force-dependent dissociation rate $k_{\text{off}}(f)$. On the other hand, if the activation potential is wide, the complex will be more sensitive to an external force because $f\gamma$ adds a larger term to the intermolecular potential.

Equation (19.2) describes how bond dissociation is changed by a constant pulling force. However, a constant pulling force is difficult to maintain in an AFM experiment. Instead, a dynamic force approach is generally used to characterize the forced dissociation

of ligand–receptor complexes (19,23). Under conditions of constant loading rate r_f, the probability density function for the forced unbinding of an adhesion complex is given by

$$P(f) = k^o \exp\left(\frac{\gamma f}{k_B T}\right) \exp\left\{\frac{k^o k_B T}{\gamma r_f}\left[1 - \exp\left(\frac{\gamma f}{k_B T}\right)\right]\right\} \tag{19.3}$$

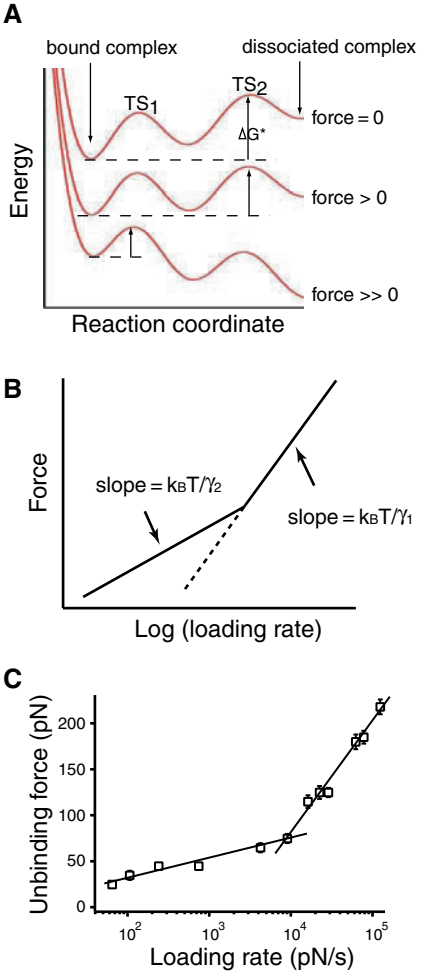


Figure 19.4. Dynamic force spectroscopy. **A.** Effects of an applied force on a protein–protein interaction potential consisting of two transition states, TS1 and TS2. In the absence of an applied force (*top*), the dissociation kinetics of the complex is determined by the outer energy barrier (i.e., TS2). An external force tilts the energy potential and suppresses the outer barrier (*middle*). Further increase in force results in a potential that is governed by the properties of the inner energy barrier (i.e., TS1) (*bottom*). **B.** Two linear regimes are predicted for a cascade of two sharp energy barriers. The increase of slope indicates that the outer barrier has been suppressed and the inner barrier has become the dominant kinetic impedance to detachment. **C**. Dynamic force spectra of the P-selectin/sLeX interactions (28). The best-fit curves (*solid lines*) were obtained using Eq. (19.4) applied to each of the two loading regimes. Error bars are the standard error of the mean (SEM). Some error bars are smaller than the symbols.

From (19.4), the most probable unbinding force f^* [i.e., the maximum of the distribution $\partial P(f)/\partial f = 0$] is

$$f^* = \frac{k_B T}{\gamma} \ln\left(\frac{\gamma}{k^o k_B T}\right) + \frac{k_B T}{\gamma} \ln(r_f) \tag{19.4}$$

Equation (19.4) shows that f^* is a linear function of the logarithm of the loading rate. The Bell model parameters are obtained from the plot of f^* versus $\ln(r_f)$, the dynamic force spectrum (DFS) of the complex (19).

The energy landscape of a complex may consist of multiple sharp activation barriers (Figure 19.4A) (19,23). In this case, the DFS is predicted to have multiple linear regimes with ascending slopes, as shown in Figure 19.2B. The increase in slope from one regime to the next indicates that an outer barrier has been suppressed by force and that an inner barrier dominates the dynamic response of the complex (Figure 19.3A). This theory is supported by recent experiments from different groups (24–27). Multiple activation barriers were found in a number of protein–protein systems, including the (strep)avidin/biotin complexes and all tested integrin/ligand complexes. A partial list of these studies is tabulated in Table 19.1.

19.4. Recent Applications

The following sections present some noteworthy and recent progress in using the AFM for protein–protein interaction studies.

19.4.1. Molecular Basis for Multiple Energy Barriers along with Protein–Protein Dissociation

Hundreds of protein pairs have been studied using the AFM. With the DFS approach, the dissociation of many protein complexes has been found to involve one or multiple potential energy barriers (Table 19.1). Recent experiments using site-directed mutagenesis and/or different physiological conditions have yielded important insight into the molecular and structural components that make up these energy barriers. This section highlights three studies investigating the molecular basis of protein interactions, with applications ranging from leukocyte extravasation, to iron transport, to receptor-mediated transport between the nucleus and cytoplasm.

The interaction between $\alpha_4\beta_1$ and vascular cell adhesion molecule-1 (VCAM-1) critically mediates leukocyte adhesion onto the vascular endothelium, and this interaction exhibits remarkable mechanical strength in resisting the large shear forces imposed by the bloodstream. To understand the molecular basis by which the $\alpha_4\beta_1$/VCAM-1 complex resists a pulling force, we employed single-molecule DFS to reveal that the dissociation of the $\alpha_4\beta_1$/VCAM-1 complex involves at least two activation potential barriers: a steep inner barrier granting the complex tensile strength to withstand large pulling forces (>50 pN) and a more elevated outer barrier that is stabilized by integrin activation (Table 19.1) (28). This special kinetic profile may reflect a biophysical basis permitting a dual physiological function (i.e., cell rolling and firm adhesion) of the $\alpha_4\beta_1$/VCAM-1 interaction. To correlate such features in the dissociation potential with molecular determinants, site-directed mutagenesis was applied to VCAM-1. Both using Mg^{2+} ion chelation and mutating Asp40 to the neutral residue alanine (D40A) suppressed the unbinding forces of the $\alpha_4\beta_1$/VCAM-1 complex, supporting the hypothesis that the inner barrier in the $\alpha_4\beta_1$/VCAM-1interaction is largely due

to the interaction between the chelated Mg^{2+} ion of the β_1 domain and the Asp40 residue of VCAM-1, an electrostatic interaction implicated by previous biochemical studies as most crucial to the stabilization of the complex. Mutagenesis of other residues in the C-D loop of D1 of VCAM-1 yielded smaller unbinding forces in both the slow- and fast-loading regimes, suggesting that variation in the dynamic strength of C-D loop mutants is due to differing widths of the inner barrier. In contrast, mutations in the D2 of VCAM-1 suppressed the unbinding force in the slow-loading regime but had no effect on forces in the fast-loading regime, suggesting that D2 of VCAM-1 helps to stabilize the activation energy of the outer barrier of the $\alpha_4\beta_1$/VCAM-1 complex (Table 19.1) (28). These DFS studies of D1 and D2 mutants provide a molecular explanation for the functionally relevant kinetic properties of the $\alpha_4\beta_1$/VCAM-1 interaction.

Multiple energy barriers have also been implicated in interactions between the iron transporter protein transferrin (Tf) and its receptor (TfR). In most mammalian cells, Tf R binds iron-loaded Tf (holo-TF) and transports it to endosomes, where acidic pH favors iron release. After returning to the cell surface, iron-free Tf (apo-TF) dissociates from TfR. Using an AFM tip functionalized with holo-Tf or apo-TF to probe TfR on both mica and cell surfaces, Yersin et al. (29) revealed striking differences between holo-Tf–TFR and apo-Tf–TfR interactions. Consistent with the accepted model of TfR cycling, the forces necessary to unbind holo-TF and TfR were always stronger than the unbinding forces required for the apo-Tf–TfR interaction, and the apo-Tf–TfR interaction was found to be pH dependent. Moreover, DFS measurements indicated that the dissociation of holo-Tf–TfR complexes involves overcoming two energy barriers, whereas apo-Tf–TfR complex dissociation involves overcoming only one. These distinct energy landscapes support a model of different binding points for holo-Tf–TfR and apo-Tf–TfR interactions.

Single-molecule studies have also yielded insight into the functional control of receptor-mediated transport of macromolecules between the nucleus and cytoplasm. The small GTPase Ran regulates the assembly and disassembly of receptor–cargo complexes by binding to the nuclear import receptor importin β1 (impβ) in its guanosine triphosphate (GTP)–bound form. Release of Ran from impβ is induced by effector proteins; however, it is unknown whether this change in stability is effected through an induced fit model whereby ligand binding triggers structural alterations or through dynamic population shifts in which changes in protein function and properties result from redistributions of preexisting conformational substates in response to binding. Using DFS to measure unbinding forces between single Ran–impβ pairs, Nevo et al. (30) found that Ran-GDP forms a single, weak complex with impβ, whereas Ran associated with the nonhydrolyzable GTP analog GppNHp leads to two distinct bound states whose fractional occupancy can be altered by applying mechanical force. The RanQ69L mutant, known for its markedly lowered GTPase activity, yielded a similar force spectrum to the wild-type analog when loaded with GDP. However, the mutant gave rise to a unimodal distribution of rupture forces in its GTP-bound form, suggesting that only the lower-strength state is accessible to GTP-bond RanQ69L. These data support a model of allosteric interactions regulating dynamic shifts between preexisting conformational isomers, extending the concept of multiple conformational states to macromolecular complexes.

19.4.2. AFM Measurements on Living Cells

Because the properties of purified molecules may not be the same as when they are in physiological conditions, it is best to study protein–protein interactions on living cells. However, the surface of a living cell is far more complicated than that of a bead or of mica coated

with protein. Furthermore, the low density of receptors on the cell surface and its nonspecific interactions introduce additional challenges. Lehenkari and Horton (31) were able to measure RGD (arginine-glycine-aspartic acid)-integrin binding on the surface of osteoblasts and osteoclasts, and we measured the binding force between concanavalin A and its receptor on the surface of NIH-3T3 fibroblasts (32).

Recent experiments in affinity imaging point to promising directions for studying protein–protein interactions with the AFM. Affinity imaging combines the force measurements of ligand–receptor interaction with the imaging function of the AFM. When the cantilever tip is coated with a specific protein, the AFM can provide an adhesion map detailing the density of the binding partner on a surface (33). Affinity images of antigens immobilized on a substrate were obtained using antibody-functionalized cantilevers (34). This technique has been extended to cell surfaces. Based on the specific interaction between *Helix pomatia* lectin and group A red blood cell plasma membrane proteins, Grandbois et al. (35) were able to discriminate group A from group O red blood cells. Similarly, the Dufrene group mapped numerous cell surface molecules on living bacteria (36).

The AFM has also been adapted for cell–cell adhesion studies (37,38). These experiments differed from those discussed earlier, in that a cell was attached to the end of the cantilever and used as a probe in cell adhesion studies. This approach allowed the study of both the ligand and the receptor under conditions close to their native environment. Potential applications of this approach include the study of modulated adhesion following cell activation. An example of such experiments is shown in Figure 19.5, in which a single leukocyte was attached to a concanavalin-A–functionalized cantilever. Using this protocol, we were able to measure the unbinding force of individual integrin leukocyte function–associated antigen-1 (LFA-1)/ICAM-1 (28,39,40) and integrin α5β1/fibronectin (27) bonds. Several other groups have also reported studies on different types of cell–cell interactions (41,42).

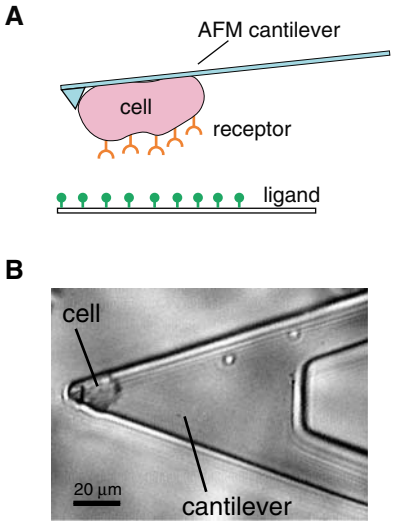


Figure 19.5. Study of protein–protein interactions in living cells. **A.** A live cell coupled to the atomic force microscope (AFM) cantilever. The cell expresses cell surface receptors, which will bind to the ligands immobilized on the tissue culture dish. **B.** Micrograph of a live leukocyte attached to the AFM cantilever. The bar is 20 µm.

19.4.3. Energy Landscape Roughness of Protein–Ligand Interaction

In its simplest representation, the energy landscape of molecular interactions between receptors and ligands is described by a barrier of certain height and width that determines the dissociation rate of the complex, as well as its dynamic strength. Some interactions, however, require a more complex landscape with additional barriers. Moreover, the surfaces of the energy landscape describing molecular interactions have been shown to be generally not smooth but to have varying heights that render them highly corrugated or rugged (43,44). This roughness slows down the dissociation kinetics of the interaction and contributes to its dynamic strength (45). Recently, it was shown that an estimate of the energy landscape roughness of protein–ligand interactions can be obtained from single-molecule dynamic force spectroscopy experiments conducted at different temperatures (46). These studies predicted that surface roughness increases the unbinding force of a protein–ligand complex at a given loading rate because it retards the unbinding process. As noted, the effect of surface roughness is temperature dependent, thus making it possible to estimate the energy roughness from AFM single-molecule measurements acquired at two temperatures T_1 and T_2. Specifically, the surface roughness ϵ of a protein–ligand interaction can be computed using Nevo's modification of the Hyeon and Thirumalai approximation,

$$\varepsilon^2 \approx \frac{x_\beta(T_1)k_\mathrm{B}T_1 x_\beta(T_2)k_\mathrm{B}T_2}{x_\beta(T_2)k_\mathrm{B}T_2 - x_\beta(T_1)k_\mathrm{B}T_1}\left[\Delta G^0\left(\frac{1}{x_\beta(T_1)}-\frac{1}{x_\beta(T_2)}\right)+\frac{k_\mathrm{B}T_1}{x_\beta(T_1)}\ln\frac{r_\mathrm{f}(T_1)x_\beta(T_1)}{k^o(T_1)k_\mathrm{B}T_1}-\frac{k_\mathrm{B}T_2}{x_\beta(T_2)}\ln\frac{r_\mathrm{f}(T_2)x_\beta(T_2)}{k^o(T_2)k_\mathrm{B}T_2}\right]$$

which takes into account the possible temperature dependence of the Bell model parameters $x_\beta(T)$ and $k^o(T)$ (45,46). Here, ΔG^0 is the height of the potential and $r_\mathrm{f}(T_1)$ and $r_\mathrm{f}(T_2)$ are the respective loading rates at the two different temperatures that give rise to the same unbinding force.

Using this approach, we recently obtained energy roughness values of $\sim 5.6 k_\mathrm{B}T$ and $7.5 k_\mathrm{B}T$ for the inner and outer barriers of the streptavidin–biotin complex, respectively (47). Similar values have been obtained for other systems, such as the unbinding of GTPase Ran from its receptor importin-β ($5.7 k_\mathrm{B}T$), the unbinding of complementary DNA strands ($10 k_\mathrm{B}T$), and the unfolding of the transmembrane helices of bacteriorhodopsin ($4 k_\mathrm{B}T$ to $6 k_\mathrm{B}T$) (46,48,49). In addition, when a force clamp technique was used to measure the unfolding rate of ubiquitin, Brujic et al. obtained a power-law distribution of unfolding rates that could be explained by assuming an exponential distribution of energy barriers with mean value of $6.6 k_\mathrm{B}T$ (50). The similarity in the results on such dissimilar systems suggests the possibility of a common origin for the roughness, perhaps due to the oversimplification involved in reducing the tridimensional dissociation pathway into a single reaction coordinate.

19.5. Concluding Remarks

Although the AFM is under improvement to enhance sensitivity (4), the major challenge for AFM research may stem from its availability, which is restricted to a limited number of laboratories. With the completion of various genome projects (including mapping the human genome, as well as the genomes of other animals and of plants), many gene sequences have been revealed. However, the proteins that genes encode do most of the work, such as building cells and running living organisms (51). Understanding how proteins interact with each other is a new frontier that may provide important insights into a variety of disease processes (52).

In the future, an AFM-based force spectroscopy approach will become increasingly important to elucidating the biophysical properties of many more protein–protein complexes. This methodology will also be increasingly adapted by biologists from different fields.

Acknowledgments

We are grateful for support by the National Institutes of Health (GM55611 to VTM), the American Heart Association (0215139B to XZ), the Fulbright-Generalitat de Catalunya postdoctoral fellowship (to FR) and the Harvard Medical School Medical Scientist Training Program fellowship (to AJX)

References

1. Wennmalm, S., and S. M. Simon. 2007. Studying individual events in biology. Annual Review of Biochemistry 76:419–446.
2. Greenleaf, W. J., M. T. Woodside, and S. M. Block. 2007. High-resolution, single-molecule measurements of biomolecular motion. Annual Review of Biophysics and Biomolecular Structure 36:171–190.
3. Binnig, G., C. F. Quate, and C. Gerber. 1986. Atomic force microscope. Physical Review Letters 56:930–933.
4. Horber, J. K., and M. J. Miles. 2003. Scanning probe evolution in biology. Science 302:1002–1005.
5. Lodish, H. F., A. Berk, P. Matsudaira, C. Kaiser, M. Krieger, M. Scott, L. Zipursky, and J. E. Darnell. 2004. Molecular Cell Biology, 5th ed. Scientific American Press, N.Y.
6. Barkai, E., Y. J. Jung, and R. Silbey. 2004. Theory of single-molecule spectroscopy: Beyond the ensemble average. Annual Review of Physical Chemistry 55:457–507.
7. Bustamante, C., J. C. Macosko, and G. J. L. Wuite. 2000. Grabbing the cat by the tail: Manipulating molecules one by one. Nature Reviews Molecular Cell Biology 1:130–136.
8. Orsello, C. E., D. A. Lauffenburger, and D. A. Hammer. 2001. Molecular properties in cell adhesion: A physical and engineering perspective. Trends in Biotechnology 19:310–316.
9. Bustamante, C., Y. R. Chemla, N. R. Forde, and D. Izhaky. 2004. Mechanical processes in biochemistry. Annual Review of Biochemistry 73:705–748.
10. Sader, J. E., I. Larson, P. Mulvaney, and L. R. White. 1995. Method for the calibration of atomic-force microscope cantilevers. Review of Scientific Instruments 66:3789–3798.
11. Senden, T. J., and W. A. Ducker. 1994. Experimental-determination of spring constants in atomic-force microscopy. Langmuir 10:1003–1004.
12. Hutter, J. L., and J. Bechhoefer. 1993. Calibration of atomic-force microscope tips. Review of Scientific Instruments 64:1868–1873.
13. Florin, E. L., V. T. Moy, and H. E. Gaub. 1994. Adhesion forces between individual ligand-receptor pairs. Science 264:415–417.
14. Alcaraz, J., L. Buscemi, M. Puig-de-Morales, J. Colchero, A. Baro, and D. Navajas. 2002. Correction of microrheological measurements of soft samples with atomic force microscopy for the hydrodynamic drag on the cantilever. Langmuir 18:716–721.
15. Janovjak, H., J. Struckmeier, and D. J. Muller. 2005. Hydrodynamic effects in fast AFM single-molecule force measurements. European Biophysics Journal with Biophysics Letters 34:91–96.
16. Lee, G. U., D. A. Kidwell, and R. J. Colton. 1994. Sensing discrete streptavidin–biotin interactions with AFM. Langmuir 10:354–361.
17. Hinterdorfer, P. 2002. Molecular recognition studies using the atomic force microscope. Methods in Cell Biology 68:115–139.
18. Bell, G. I. 1978. Models for the specific adhesion of cells to cells. Science 200:618–627.
19. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. Biophysical Journal 72:1541–1555.
20. Dudko, O. K., G. Hummer, and A. Szabo. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. Physical Review Letters 96:108101.
21. Friddle, R. W. 2008. Unified model of dynamic forced barrier crossing in single molecules. Physical Review Letters 100:138302.

22. Husson, J., and F. Pincet. 2008. Analyzing single-bond experiments: Influence of the shape of the energy landscape and universal law between the width, depth, and force spectrum of the bond. Physical Review E 77:026108.
23. Evans, E. A., and D. A. Calderwood. 2007. Forces and bond dynamics in cell adhesion. Science 316:1148–1153.
24. Evans, E., A. Leung, D. Hammer, and S. Simon. 2001. Chemically distinct transition states govern rapid dissociation of single L-selectin bonds under force. Proceedings of the National Academy of Sciences of the United States of America 98:3784–3789.
25. Merkel, R., P. Nassoy, A. Leung, K. Ritchie, and E. Evans. 1999. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. Nature 397:50–53.
26. Zhang, X., E. Wojcikiewicz, and V. T. Moy. 2002. Force spectroscopy of the leukocyte function-associated antigen-1/intercellular adhesion molecule-1 interaction. Biophysical Journal 83:2270–2279.
27. Li, F., S. D. Redick, H. P. Erickson, and V. T. Moy. 2003. Force measurements of the alpha(5)beta(1) integrin–fibronectin interaction. Biophysical Journal 84:1252–1262.
28. Zhang, X., D. F. Bogorin, and V. T. Moy. 2004. Molecular basis of the dynamic strength of the sialyl Lewis X–selectin interaction. Chemphyschem 5:175–182.
29. Yersin, A., T. Osada, and A. Ikai. 2008. Exploring transferrin-receptor interactions at the single-molecule level. Biophysical Journal 94:230–240.
30. Nevo, R., C. Stroh, F. Kienberger, D. Kaftan, V. Brumfeld, M. Elbaum, Z. Reich, and P. Hinterdorfer. 2003. A molecular switch between alternative conformational states in the complex of Ran and importin beta1. Nature Structural Biology 10:553–557.
31. Lehenkari, P. P., and M. A. Horton. 1999. Single integrin molecule adhesion forces in intact cells measured by atomic force microscopy. Biochemical and Biophysical Research Communications 259:645–650.
32. Chen, A., and V. T. Moy. 2000. Cross-linking of cell surface receptors enhances cooperativity of molecular adhesion. Biophysical Journal 78:2814–2820.
33. Kienberger, F., A. Ebner, H. J. Gruber, and P. Hinterdorfer. 2006. Molecular recognition imaging and force spectroscopy of single biomolecules. Accounts of Chemical Research 39:29–36.
34. Willemsen, O. H., M. M. Snel, K. O. van der Werf, B. G. de Grooth, J. Greve, P. Hinterdorfer, H. J. Gruber, H. Schindler, Y. van Kooyk, and C. G. Figdor. 1998. Simultaneous height and adhesion imaging of antibody-antigen interactions by atomic force microscopy. Biophysical Journal 75:2220–2228.
35. Grandbois, M., W. Dettmann, M. Benoit, and H. E. Gaub. 2000. Affinity imaging of red blood cells using an atomic force microscope. Journal of Histochemistry and Cytochemistry 48:719–724.
36. Dufrene, Y. F. 2008. AFM for nanoscale microbe analysis. Analyst 133:297–301.
37. Benoit, M., D. Gabriel, G. Gerisch, and H. E. Gaub. 2000. Discrete interactions in cell adhesion measured by single-molecule force spectroscopy. Nature Cell Biology 2:313–317.
38. Razatos, A., Y. L. Ong, M. M. Sharma, and G. Georgiou. 1998. Molecular determinants of bacterial adhesion monitored by atomic force microscopy. Proceedings of the National Academy of Sciences of the United States of America 95:11059–11064.
39. Wojcikiewicz, E. P., X. Zhang, A. Chen, and V. T. Moy. 2003. Contributions of molecular binding events and cellular compliance to the modulation of leukocyte adhesion. Journal of Cell Science 116:2531–2539.
40. Wojcikiewicz, E. P., M. H. Abdulreda, X. H. Zhang, and V. T. Moy. 2006. Force spectroscopy of LFA-1 and its ligands, ICAM-1 and ICAM-2. Biomacromolecules 7:3188–3195.
41. Krieg, M., Y. Arboleda-Estudillo, P. H. Puech, J. Kafer, F. Graner, D. J. Muller, and C. P. Heisenberg. 2008. Tensile forces govern germ-layer organization in zebrafish. Nature Cell Biology 10:429–U122.
42. Puech, P. H., K. Poole, D. Knebel, and D. J. Muller. 2006. A new technical approach to quantify cell-cell adhesion forces by AFM. Ultramicroscopy 106:637–644.
43. Ansari, A., J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young. 1985. Protein states and protein quakes. Proceedings of the National Academy of Sciences of the United States of America 82:5000–5004.
44. Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. Science 254:1598–1603.
45. Hyeon, C., and D. Thirumalai. 2007. Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. Journal of Physics—Condensed Matter 19:113101.
46. Nevo, R., V. Brumfeld, R. Kapon, P. Hinterdorfer, and Z. Reich. 2005. Direct measurement of protein energy landscape roughness. EMBO Reports 6:482–486.
47. Rico, F., and V. T. Moy. 2007. Energy landscape roughness of the streptavidin-biotin interaction. Journal of Molecular Recognition 20:495–501.
48. Janovjak, H., H. Knaus, and D. J. Muller. 2007. Transmembrane helices have rough energy surfaces. Journal of the American Chemical Society 129:246–247.

49. Schumakovitch, I., W. Grange, T. Strunz, P. Bertoncini, H. J. Guntherodt, and M. Hegner. 2002. Temperature dependence of unbinding forces between complementary DNA strands. Biophysical Journal 82:517–521.
50. Brujic, J., R. I. Hermans, K. A. Walther, and J. M. Fernandez. 2006. Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. Nature Physics 2:282–286.
51. Walhout, A. J. M., and M. Vidal. 2001. Protein interaction maps for model organisms. Nature Reviews Molecular Cell Biology 2:55–62.
52. Wells, J. A., and C. L. McClendon. 2007. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450:1001–1009.
53. Yu, J. P., Y. X. Jiang, X. Y. Ma, Y. Lin, and X. H. Fang. 2007. Energy landscape of aptamer/protein complexes studied by single-molecule force spectroscopy. Chemistry—An Asian Journal 2:284–289.
54. Bonanni, B., A. S. M. Kamruzzahan, A. R. Bizzarri, C. Rankl, H. J. Gruber, P. Hinterdorfer, and S. Cannistraro. 2005. Single molecule recognition between cytochrome C 551 and gold-immobilized azurin by force spectroscopy. Biophysical Journal 89:2783–2791.
55. Neuert, G., C. Albrecht, E. Pamir, and H. E. Gaub. 2006. Dynamic force spectroscopy of the digoxigenin–antibody complex. EBS Letters 580:505–509.
56. Sletmoen, M., G. Skjak-Braek, and B. T. Stokke. 2004. Single-molecular pair unbinding studies of mannuronan C-5 epimerase AlgE4 and its polymer substrate. Biomacromolecules 5:1288–1295.
57. Taranta, M., A. R. Bizzarri, and S. Cannistraro. 2008. Probing the interaction between p53 and the bacterial protein azurin by single molecule force spectroscopy. Journal of Molecular Recognition 21:63–70.
58. Maki, T., S. Kidoaki, K. Usui, H. Suzuki, M. Ito, F. Ito, Y. Hayashizaki, and T. Matsuda. 2007. Dynamic force spectroscopy of the specific interaction between the PDZ domain and its recognition peptides. Langmuir 23:2668–2673.
59. Dettmann, W., M. Grandbois, S. Andre, M. Benoit, A. K. Wehle, H. Kaltner, H. J. Gabius, and H. E. Gaub. 2000. Differences in zero-force and force-driven kinetics of ligand dissociation from beta-galactoside-specific proteins (plant and animal lectins, immunoglobulin G) monitored by plasmon resonance and dynamic single molecule force microscopy. Archives of Biochemistry and Biophysics 383:157–170.
60. Fritz, J., A. G. Katopodis, F. Kolbinger, and D. Anselmetti. 1998. Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy. Proceedings of the National Academy of Sciences of the United States of America 95:12283–12288.
61. Hanley, W., O. McCarty, S. Jadhav, Y. Tseng, D. Wirtz, and K. Konstantopoulos. 2003. Single molecule characterization of P-selectin/ligand binding. Journal of Biological Chemistry 278:10556–10561.
62. Krasnoslobodtsev, A. V., L. S. Shlyakhtenko, and Y. L. Lyubchenko. 2007. Probing interactions within the synaptic DNA-Sfil complex by AFM force spectroscopy. Journal of Molecular Biology 365:1407–1416.
63. Yu, J. P., Q. Wang, X. L. Shi, X. Y. Ma, H. Y. Yang, Y. G. Chen, and X. H. Fang. 2007. Single-molecule force spectroscopy study of interaction between transforming growth factor beta 1 and its receptor in living cells. Journal of Physical Chemistry B 111:13619–13625.
64. Baumgartner, W., P. Hinterdorfer, W. Ness, A. Raab, D. Vestweber, H. Schindler, and D. Drenckhahn. 2000. Cadherin interaction probed by atomic force microscopy. Proceedings of the National Academy of Sciences of the United States of America 97:4005–4010.
65. Zhang, X., S. E. Craig, H. Kirby, M. J. Humphries, and V. T. Moy. 2004. Molecular basis for the dynamic strength of the integrin alpha4beta1/VCAM-1 interaction. Biophysical Journal 87:3470–3478.
66. Wielert-Badt, S., P. Hinterdorfer, H. J. Gruber, J. T. Lin, D. Badt, B. Wimmer, H. Schindler, and R. K. Kinne. 2002. Single molecule recognition of protein binding epitopes in brush border membranes by force microscopy. Biophysical Journal 82:2767–2774.
67. Allen, S., X. Chen, J. Davies, M. C. Davies, A. C. Dawkes, J. C. Edwards, C. J. Roberts, J. Sefton, S. J. Tendler, and P. M. Williams. 1997. Detection of antigen–antibody binding events with the atomic force microscope. Biochemistry 36:7457–7463.
68. Hinterdorfer, P., W. Baumgartner, H. J. Gruber, K. Schilcher, and H. Schindler. 1996. Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. Proceedings of the National Academy of Sciences of the United States of America 93:3477–3481.
69. Allen, S., J. Davies, M. C. Davies, A. C. Dawkes, C. J. Roberts, S. J. Tendler, and P. M. Williams. 1999. The influence of epitope availability on atomic-force microscope studies of antigen-antibody interactions. Biochemical Journal 341:173–178.
70. Vinckier, A., P. Gervasoni, F. Zaugg, U. Ziegler, P. Lindner, P. Groscurth, A. Pluckthun, and G. Semenza. 1998. Atomic force microscopy detects changes in the interaction forces between GroEL and substrate proteins. Biophysical Journal 74:3256–3263.
71. Yip, C. M., C. C. Yip, and M. D. Ward. 1998. Direct force measurements of insulin monomer-monomer interactions. Biochemistry 37:5439–5449.

72. Nakajima, H., Y. Kunioka, K. Nakano, K. Shimizu, M. Seto, and T. Ando. 1997. Scanning force microscopy of the interaction events between a single molecule of heavy meromyosin and actin. Biochemical and Biophysical Research Communications 234:178–182.
73. Berry, M., T. J. McMaster, A. P. Corfield, and M. J. Miles. 2001. Exploring the molecular adhesion of ocular mucins. Biomacromolecules 2:498–503.
74. Dammer, U., O. Popescu, P. Wagner, D. Anselmetti, H. J. Guntherodt, and G. N. Misevic. 1995. Binding strength between cell adhesion proteoglycans measured by atomic force microscopy. Science 267:1173–1175.

20

A New Approach to Analysis of Single-Molecule Force Measurements

Evan Evans, Ken Halvorsen, Koji Kinoshita, and Wesley P. Wong

Abstract A common aim in probing single molecular bonds or the structural stability of proteins is to measure the kinetic rates at which a bond dissociates or a protein changes conformation under conditions of changing force. Using sample data taken from tests of ligand–receptor unbinding and protein unfolding/refolding, we show that populations of "single molecule" events, arranged into statistical arrays expressing the numbers of bonds or initial conformers remaining as a function of time and cumulated into histograms of transitions over fixed time increments, provide the bases for a model-independent assay of the kinetic rates of transition throughout the course of an experiment. Most important, this assay for kinetic rates can be employed with any deterministic mode of force spectroscopy, whether the pulling force increases *or* decreases with time.

20.1. Introduction

In tests of molecular complexes or structures with force probes, the common objective is to measure the kinetic rate $k_{\rightarrow}$ at which a single complex (bond) dissociates or a single molecule changes conformation when subjected to a pulling force. As such, success in molecular force measurements requires an experimental approach that will produce a high probability of single-molecule events plus reliable criteria for excluding contaminating events that arise from irrelevant (nonspecific) and multiple molecule attachments. After briefly discussing ways to best satisfy these requirements, we show that arranging the refined population of putative "single-molecule" events into (1) an array $N(t_i)$ expressing the numbers of bonds or initial conformers remaining as a function of time and (2) its histogram $\Delta N_k(t_k)$ of transition events cumulated over fixed time increments Δt together establish a direct (model-independent) assay for the kinetic rates of transition, $k_{\rightarrow}(t_k) \approx (\Delta N_k/\Delta t)/N(t_k)$ defined at the mean sample times t_k. Although formulated here in terms of the ratios of probability density/probability for a two-state molecular transition [i.e., $p(t_k)/S(t_k) = k_{\rightarrow}(t_k)$],

E. Evans, K. Kinoshita • Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA, and Physics and Pathology, University of British Columbia, Vancouver, BC Canada V6T 2A6
K. Halvorsen, W.P. Wong • The Rowland Institute at Harvard, Harvard University Cambridge MA 02142, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_20,

the statistics-based approach can be readily extended to encompass more hierarchal and complex transitions if backed by appropriate evidence. Of fundamental significance, the experimental assay for kinetic rates provides a general method to analyze results obtained with any deterministic mode of force spectroscopy, $f(t)$, whether the pulling force increases *or* decreases with time. Using examples of ligand–receptor unbinding and protein unfolding/refolding, we demonstrate that plotting the experimental assay for kinetic rates on a logarithmic scale versus pulling force $f(t_k)$, and invoking Arrhenius phenomenology, $k_{\rightarrow}(f) = k_0 \exp[-\Delta G(f)/k_B T]$,[1] reveals both the impact of force on the chemical energy barrier governing the kinetics of single-molecule transitions, $\Delta G(f)/k_B T$, and the rate prefactor, $k_0 = 1/t_0$, characterizing persistence of the force-free state. As a preface to these examples, we first outline principles of instrument design and procedures for data processing that lead to reliable assays of molecular transitions—the obvious challenge being to acquire "single-molecule events".

20.2. Force Probe Design and the Quest for Single-Molecule Statistics

A molecular force probe is an ultrasensitive "spring" with a reactive tip that initiates the attachments to be tested. Thus, touched repeatedly to a target surface, the probe tip is prepared with a very low concentration of the initiator molecule to ensure rare discrete attachments. Since the reacting molecules on a tip and target surface are usually limited in range of mobility, the likelihood of forming attachments on touch depends primarily on the area of contact and to a much lesser extent on contact time. Consequently, the probe instrument should be designed to touch the target surface with a small (less than a few piconewtons) compressive force regulated by feedback. With both the impingement force and duration of touch controlled, the conditions leading to attachments will be nearly identical for every touch to a target, allowing the rare events to be viewed as a stationary-random process (amenable to characterization by Poisson statistics as described later). Equally important for the assay presented here, a probe must have sufficient precision to report forces with ~1-pN accuracy over a wide range from minus a few piconewtons to +100 pN or more. Thus, to achieve this accuracy, the probe transducer should have a low spring constant (less than or equal to a piconewton per nanometer), as well as a nanometer-precision system for tracking its tip displacement.

As an example, a biomembrane force probe (BFP) [1,2] was used to test the ligand–receptor bonds described later, providing the force–time traces seen in Figure 20.1. With a tunable spring constant set typically between $0.3 < \kappa_f < 0.5$ pN/nm and with 3- to 4-nm precision in video tracking of its tip, the BFP can measure forces in the range from –30 to +300 pN with a standard deviation (SD) of ±1–2 pN. As other examples, two types of laser optical trap force probes (OTFP) were used to test the unfolding/refolding of polyprotein domains [3,4]. With spring constants set typically between $0.005 < \kappa_f < 0.1$ pN/nm, the fast video tracking systems in these OTFP instruments, one with 2- to 3-nm precision [3] and the other with <1-nm precision [4,5], provide much better force accuracy, SD < ±0.1 pN, as demonstrated in Figure 20.2 (also later in Figure 20.9), than a BFP.

[1] Thermal energy $k_B T$ is ~4.08×10^{-21} J at room temperature.

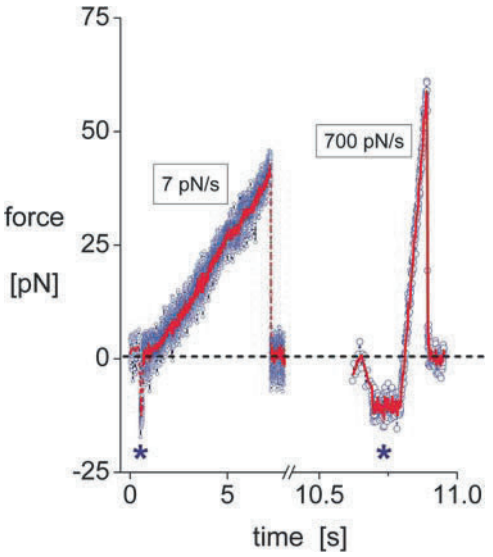


Figure 20.1. Force–time traces obtained from biomembrane force probe (BFP) in tests of ligand–receptor bonds: The left trace is at a slow (7 pN/s) ramp and the right trace is at a fast (700 pN/s) ramp of pulling force. A tenfold break in time axis separates the two traces to illustrate the major reduction in bond lifetime (i.e., from ~7 sec to ~0.1 sec) that accompanies the increase in its mechanical strength (i.e., from ~40 pN to ~60 pN) when ramps increase from slow to fast rates. Stars mark feedback control of the soft impingement (~–10 pN) and a brief, 0.1-sec period of contact ending with piezo retractions of the microsphere targets at fixed speeds. The red traces are five-point moving averages of the direct traces (blue) obtained at a sampling rate of 1,500/sec. Limited by uncertainty (~±4 nm standard deviation [SD]) in video tracking of the BFP tip deflection, the spread in force (~± 2-pN SD; *blue traces*) increases and decreases, in proportion to the spring constant k_f chosen for the experiments (here set to 0.5 pN/nm by micropipet suction).

However, smaller ranges of force are usually accessible to optical trap instruments, for example, –10 to +50 pN. Since rates of molecular transitions increase rapidly when subjected to increasing force, probe instruments have to provide high levels of precision at very fast sampling frequencies, which are limited to ~2–3 × 10^3/sec by viscous damping for the BFP and OTFP.

In most of the experimental examples described here, molecules were stressed by retracting a tip or target away from contact at fixed speed v_{pull}, resulting in a "force ramp." Although programmed to achieve a ramp of $\kappa_f v_{\text{pull}}$, the actual ramp produced by steady speed retraction depends on the effective stiffness κ_{eff} of the probe spring in series with the tip–target interfacial structure. For tests exhibiting linear force increases with time (cf. Figures 20.1 and 20.2), the tip–target interfaces respond like elastic springs, $\kappa_{t\sim t}$, and therefore, $\kappa_{\text{eff}} = \kappa_f \kappa_{t\sim t}/(\kappa_f + \kappa_{t\sim t})$. Thus, the ratio of observed ramp/nominal ramp, $c_{\text{rf}} = r_f/(\kappa_f v_{\text{pull}})$, provides a direct assay of the tip–target interfacial stiffness, $\kappa_{t\sim t} = \kappa_f\, c_{\text{rf}}/(1 - c_{\text{rf}})$. As an example, using amino- or mercapto-silane and hetero-bifunctional polymer linkers to immobilize proteins on glass microspheres, one finds the ratios c_{rf} to depend on the molecules chosen, typically exhibiting values between 0.7 and 0.9 with small SD (<±0.1). Close to 1, this range of values shows that the tip–target interface is much stiffer than the probe spring.

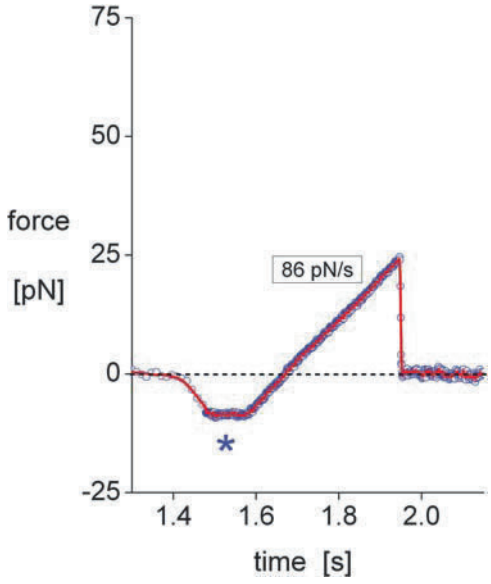


Figure 20.2. Force–time trace from a laser OTFP test. Replacing a biomembrane force probe (BFP) as the transducer in a nearly identical apparatus, the probe is a microsphere confined by a focused laser beam [3]. Controlled by feedback and moved by a piezo as in Figure 20.1, a target microsphere was driven to contact the probe with ∼–8-pN impingement (*star*) for 0.1 sec, then retracted at a fixed speed to stress an attachment with a ramp of 90 pN/sec until rupture. The red trace is a five-point moving average of the direct trace (*blue*) obtained at a sampling rate of 1,500/sec. Limited by the uncertainty (∼± 2-nm standard deviation [SD]) in video tracking of the probe deflection, the narrow spread in force (∼±0.1-pN SD, *blue trace*) demonstrates the better precision obtained with the softer OTFP spring constant (here, $k_f \sim 0.05$ pN/nm) set through control of the laser intensity.

20.3. Identifying Events Arising from Nonspecific and Multiple-Specific Attachments

Even using the type of probe instrument just described and controlling contact conditions, we still need to define procedures for dealing with the contaminating events that arise from nonspecific interactions and multiple-molecule attachments. Used to great advantage in testing unfolding/refolding of polyprotein domains with atomic force microscopy (AFM), one approach has been to accept only force–time traces showing "sawtooth" patterns with force increases and sharp drops equal to or less than the number of domains present in a single protein chain. Similarly, in testing weak molecular bonds, force traces with obvious multiple force drops are rejected. Yet, this method of inspection will not identify nonspecific attachments. Moreover, because of limits in time resolution, probe response, and fluctuations in force, a significant fraction of multiple attachments seem to escape detection[2] when compared to predictions from Poisson statistics (described later). Consequently,

[2] Subjected to large forces, preemptive failure of one component in a cluster of attachments will immediately transfer the load to the remaining components, resulting in an extremely rapid sequence of failures with little separation in time [6]. Because of probe response and fluctuations, only the final outcome of this sequence of events—a single drop to zero force—is likely be detected in most force–time traces. The only evidence of the multiple bonds is the large magnitude of breaking force reported by the probe.

we will demonstrate a simple conservative procedure to identify the major subset of "raw statistics" that most likely contains the "single-molecule" events. After describing the procedures, we will illustrate their impact on "raw statistics" with data from tests of molecular unbinding.

20.3.1. Dealing with Nonspecific Events

Performed under the same conditions of target impingement and retraction speed as in the tests of functional properties, we employ control experiments to identify which bins in histograms are likely to contain significant numbers of nonspecific events. Since nonspecific interactions in controls arise mainly at low forces, the simplest approach is to ignore this low range of forces in histograms when the numbers become significant. Although the approach is unsophisticated, we will show that ignoring these short-time statistics has no effect on the assay for transition rates at longer times and higher forces. The procedure merely sets a lower bound to the range of data analyzed. Most of the studies to be described here involve molecules immobilized on rigid surfaces at concentrations adjusted to yield <15%–20% specific attachments from thousands of repeated contacts. With these preparations, control experiments typically yield tenfold fewer attachments. As a demonstration relevant to later examples, Figure 20.3 shows force histograms obtained from tests in NaCl buffer plus 5 mM ethylenediaminetetraacetic acid (EDTA) between a microsphere tip linked covalently with intercellular adhesion molecule-1 (ICAM-1) and a microsphere target presenting recombinant β_2-integrin lymphocyte function-associated antigen-1 (LFA-1) held by a covalently linked activating antibody (240Q [7,8]). Touched ~1,400–1,800 times at a force of –10 pN (for 0.1 sec), less than 1%–2% of the tip–target contacts in EDTA indicated an attachment.

In marked contrast to microsphere targets, even when performed with the same soft touch and brief contact, controls for interactions at surfaces of biological cells yield much higher frequencies of nonspecific attachments and larger ranges of forces affecting statistics. Strongly dependent on the speed of target retraction, the increased number of nonspecific attachments and range of forces are primarily consequences of a much larger area of contact at a cell surface and the stronger hydrodynamic coupling ("suction") that accompanies fast retraction of the "soft" target [2]. Demonstrating the enhanced nonspecific interaction with cell surfaces, Figure 20.4 shows another set of controls relevant to LFA-1 interactions in which a probe linked with an irrelevant protein (bovine serum albumin) was touched against human neutrophilic leukocytes (PMNs), this time in buffer plus 2 mM Mn^{2+}. Although large numbers of nonspecific forces are seen at fast retraction speeds in Figure 20.4, well-defined populations of specific-integrin attachments can be observed beyond the range contaminated by nonspecific events when a probe decorated with low concentration of ICAM-1 is touched to PMNs in Mn^{2+}. As we show in a later example (Figure 20.8), simply ignoring this initial force range leaves the prominent peak of specific events amenable to analysis.

20.3.2. Dealing with Multiple-Specific Events

Unlike nonspecific interactions, the procedure for identifying and excluding hidden events that originate from multiple attachments is not benign and must be considered carefully. Since definitive observational criteria are usually not available for rejecting events,

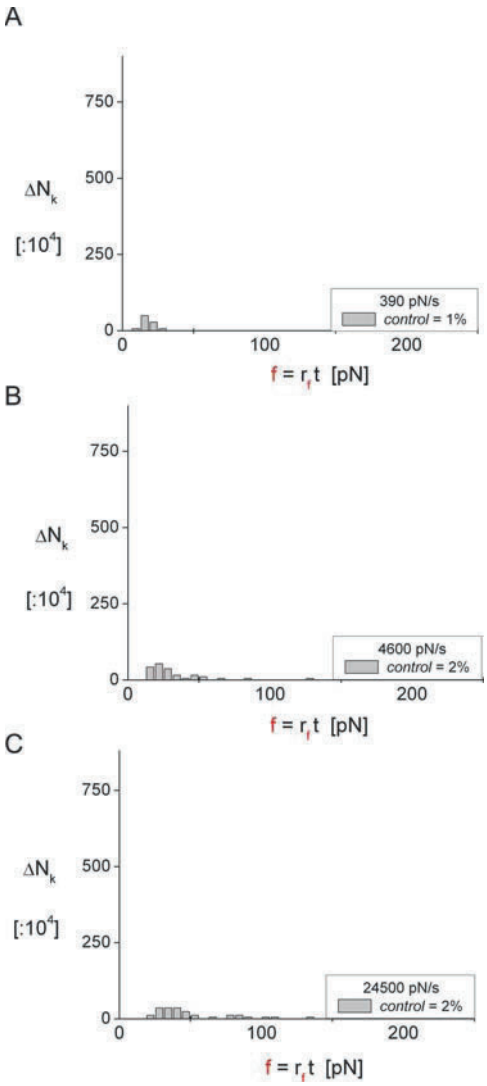


Figure 20.3. Controls for nonspecific interactions in tests of lymphocyte function-associated antigen-1 interactions that involve immobilized constructs on glass microspheres. Performed at three force rates in NaCl buffer plus 5 mM ethylenediaminetetraacetic acid, force histograms show the numbers of nonspecific attachments (scaled per 10^4 touches) obtained from 1,400–1,800 touches between a biomembrane force probe tip linked covalently with intercellular adhesion molecule-1 and the microsphere targets presenting recombinant $\alpha_L\beta_2$ integrin held by a covalently linked activating antibody (240Q [6]).

the primary objective in all experiments should be to establish a stationary-random process for formation of attachments, employing sensitive feedback to control the initial contact and reducing the concentration of molecular reactants to achieve *rare* events. Once accomplished, the attachment frequency A_ω can be used to define the probability $P_0 = (1 - A_\omega)$ for no attachment, from which Poisson estimates follow for the probabilities of a single attachment,

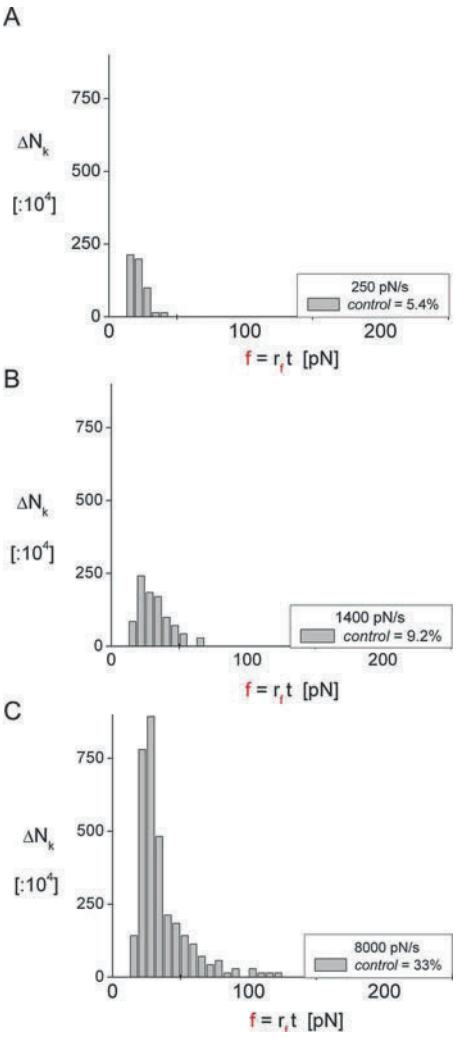


Figure 20.4. Controls for nonspecific interactions in tests of lymphocyte function-associated antigen-1 (LFA-1) interactions at the surfaces of human neutrophilic leukocytes (PMNs). Performed at three force rates in buffer plus 2 mM Mn^{2+}, force histograms show the numbers of nonspecific attachments (scaled per 10^4 touches) obtained from ~1,000 touches between a biomembrane force probe tip linked covalently with an irrelevant protein (bovine serum albumin) and PMNs expressing LFA-1.

$P_1 = (A_\omega - 1)\ln(1 - A_\omega)$, and all multiples, $P_{n\geq 2} = A_\omega + (1 - A_\omega)\ln(1 - A_\omega)$. Since multiple attachments share the applied force (albeit with unknown distribution), we expect their transitions (detachments or conformational changes) to occur at unusually long times and very large forces.Hence, we use the probability for multiples to conservatively identify a small fraction of events $P_{n\geq 2}/A_\omega$ (typically ~10%–15% or less) to be excluded at the longest transition times (largest forces). Although there are only a few events, we show next that truncating "raw statistics" by this precise fraction produces arrays of events $N(t_i)$ with appropriate convergence at long times.

20.4. Establishing Estimators for Initial State Probability and Distribution of Transitions

Using data from tests between a BFP linked with ICAM-1 and microsphere targets linked with recombinant LFA-1 ($\alpha_L\beta_2$-LZ [9]), we will first demonstrate how to obtain estimators for initial state probability and its probability distribution of transitions.[3] As noted in the introduction, these two observables form the basis for the direct assay of the kinetic rates in force probe experiments. Figures 20.5A, B show examples of the arrays $N(t_i)$ (open blue circles) defining unnormalized probabilities of bond survival, which were obtained by excluding a small number of "Poisson multiples" at the terminus (long-time tail) of each "raw data" set (open black circles). Equivalent to time (i.e., $f = r_f t$), the numbers of integrin attachments $N(t_i)$ that survived in 2 mM Ca^{2+} are plotted as a function of force in Figure 20.5A for a ramp of 770 pN/sec. Based on 417 attachments from 1,703 cycles of approach-touch-retraction, the Poisson fraction $P_{n\geq 2}/A_\omega \approx 0.134$ marked 56 events for exclusion from the "raw data" tail. Similarly, the numbers of integrin attachments that survived in 2 mM Mg^{2+} are plotted in Figure 20.5B as a function of force for a ramp of 89 pN/sec. In total, 375 attachments were detected from 1,804 cycles of approach-touch-retraction. Again from the attachment frequency, the Poisson fraction $P_{n\geq 2}/A_\omega \approx 0.112$ marked 42 events for exclusion from the "raw data" tail. Establishing unnormalized estimators for initial state probabilities $[S_1(t_i)]$, the arrays $N(t_i)$ (open blue circles) in Figures 20.5A, B are expected to be enriched significantly in "single-molecule" events, which is consistent with the rapid convergence in statistics observed at large forces (long times). Plotted directly below in Figures 20.5C, D, histograms of the "raw data" show the bins of events that were excluded at long times (cross-hatched), as well as the few nonspecific events expected from controls (gray bins) at low forces. As described later, the intervening (bright yellow) bins represent the numbers ΔN_k of unbinding transitions within fixed sampling periods Δf ($= r_f\ \Delta$t), providing the estimators for probability density p(t_k) and the rate of decrease in the initial state probability, $p(t_k) = -[dS_1/dt]_k$. Unless supported by additional evidence, more-aggressive truncations of the raw data to obtain these estimators would be unsubstantiated and likely to introduce artifacts in the assay of kinetic rates.

20.5. Two-State Transitions and the Direct Experimental Assay for Kinetic Rates

Applicable to many types of single-molecule tests, the estimators for initial state probability, i.e., $S_1(t) \approx N(t_i)/N(0)$, characterize the likelihood that single-molecular complexes or conformers remain in the initial state throughout the course of an experiment. Treated here as a two-state transition, a first-order Markov equation describes the temporal evolution of this initial state to a subsequent state $S_2(t)$ of molecular dissociation or new conformation, that is,

$$dS_1(t)/dt = -k_{\rightarrow}(t)S_1(t) + k_{\leftarrow}(t)[1 - S_1(t)]$$

$$S_2(t) \equiv [1 - S_1(t)]$$

[3] Used in this chapter to illustrate the analysis of force distributions in ligand–receptor dissociation under force ramps, the sample data for interactions between ICAM-1 and β_2-integrin (LFA-1 on leukocytes and recombinant $\alpha_L\beta_2$-LZ heterodimer) are taken from a two manuscripts describing many tests of LFA-1 integrin interactions, which have been submitted to the *Biophysical Journal* for publication.

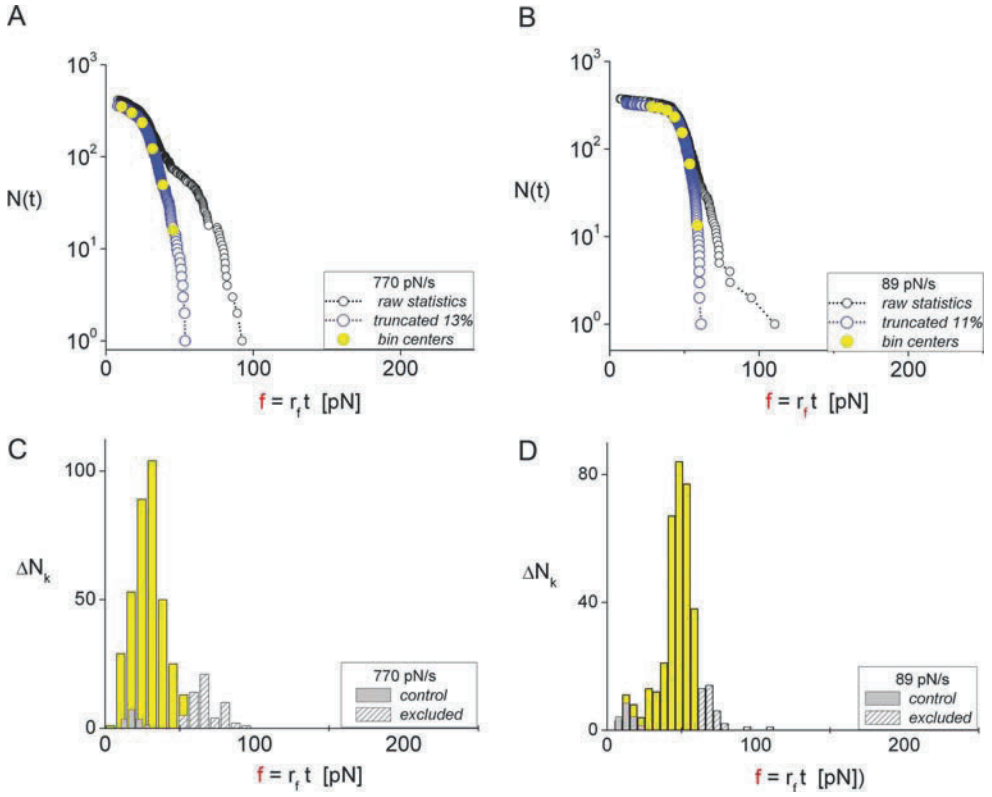


Figure 20.5. Unbinding intercellular adhesion molecule-1 attachments to $\alpha_L\beta_2$ immobilized on microspheres with ramps of force. **A.** Tests in 2 mM Ca^{2+} at a ramp rate of 770 pN/sec produced 417 attachments from 1,703 contacts. Thus, a Poisson fraction of 56 putative multiples was marked for exclusion from the "raw statistics" (*black open circles*), yielding the unnormalized estimator $N(f_k)$ for initial-state probability (*blue open circles*). **B.** Tests in 2 mM Mg^{2+} at a ramp rate of 89 pN/sec produced 375 attachments from 1,804 contacts. Thus, a Poisson fraction of 42 putative multiples was marked for exclusion from the "raw statistics," again yielding the unnormalized estimator $N(f_k)$ for initial-state probability. **C, D.** Histograms of the "raw data" appearing in panels A and B. The cross-hatched bins identify the events truncated from the "raw statistics" at long times to obtain the unnormalized estimator for initial state probability $N(t_i)$. The gray bins represent the nonspecific events expected from controls scaled to match the appropriate number of contacts. The bright-yellow bins containing the majority of events represent the numbers ΔN_k of transitions occurring within fixed sampling periods $\Delta f\,(= r_f\,\Delta t)$, yielding the unnormalized estimators for probability density, $p(t_k) = (r_f/\Delta f)\Delta N_k$, at the bin centers $f_k = r_f t_k$. (The closed yellow circles in panels A and B identify interpolation points.)

Governing the two-state dynamics are the instantaneous frequencies of forward $k_{\rightarrow}(t)$ and reverse $k_{\leftarrow}(t)$ transitions. Since application of pulling forces with a "soft" spring eliminates reverse transitions [10,11], we approximate $k_{\leftarrow}$ as $\sim$0 and reduce the two-state dynamics to a first-order decay process, $dS_1(t)/\mathrm{dt} \approx -k_{\rightarrow}(t)S_1(t)$, the solution of which reflects the cumulated frequency of forward transition, that is, $S_1(t) = \exp[-\int_{0\rightarrow t} k_{\rightarrow}(t')\,dt']$. As such, the probability density $p(t)$ now represents the "instantaneous" rate, $k_{\rightarrow}(t)S_1(t)$, at which forward transitions drain the initial state, and the ratio of probability density/probability becomes equivalent to the forward kinetic rate, $k_{\rightarrow}(t)$, at any time. Hence, equally important in our approach to data analysis, the numbers of events ΔN_k in histograms, cumulated over discrete

time differences Δt, provide key estimators for probability density, $p(t_k) \approx (1/\Delta t)[\Delta N_k/N(0)]$, at the bin center times t_k. Interpolating among the array $N(t_i)$ to obtain the values $N(t_k)$ (identified by solid yellow circles in Figures 20.5A–B), we establish the ingredients needed to specify ratios $p(t_k)/S(t_k)$ of probability density/probability and to directly assay the forward kinetic rates, $k_{\rightarrow}(t_k)$. Perhaps most significant, this generic assay for kinetic rates is valid even when transition rates vary over time, which allows the approach to be applied in all modes of force spectroscopy as long as the history of force follows a precise function in time $f(t)$. Since the estimator for probability density/probability, $(1/\Delta t)\ [\Delta N_k/N(t_k)]$, is independent of $N(0)$, the results do not depend on the initial force (i.e., time) where we commence the analysis, which allows us to simply ignore bins at low forces in histograms containing large numbers of nonspecific events.

In the sections that follow, we will demonstrate this method of analysis using data from a variety of experiments. Along with the requirement for a well-defined population of single-molecule transitions, the key assumption underlying validity of the two-state approach is that the internal dynamics of transitions (e.g., pausing at metastable states or involving multiple reaction pathways) must be orders of magnitude faster than the mean time needed to reach the end state. Even so, prominent consequences of internal transitions can be perceived through exploring a large range of force or force ramps, as recognized some time ago [12].

20.6. Experimental Example: Dissociating ICAM-1 from β_2-Integrin with Force Ramps

20.6.1. Microsphere Targets

In analyzing data from force ramp experiments, we employ an equivalent and very useful transformation that replaces the statistical estimators based on transition times with those based on forces, that is, $p(t_k)/S(t_k) \equiv r_f\, p(f_k)/S(f_k)$.[4] By changing from transition times t_k to force statistics $f_k = r_f t_k$, we will determine the ratios of probability density/probability, $p(t_k)/S(t_k) \approx (r_f\,/\Delta f)[\Delta N_k/N(f_k)]$, from the bin contents in force histograms ΔN_k and values for $N(f_k)$ obtained at the bin centers f_k by interpolation among the initial state statistics $N(f_i)$. Finding these ratios from data at several ramps like the examples shown in Figure 20.5, we plot the results on a logarithmic scale in Fig. 20.6A, B to expose the dramatic increases in off-rates that accompany increases in the forces. Moreover, starting from their intercepts at zero force, the off-rates of ICAM-1 from LFA-1 are seen to differ significantly in the two divalent cation environments. Extended to higher and lower forces with ratios from faster and slower ramps (solid green, magenta, red circles in Fig. 20.6A, B), the marked difference in kinetics becomes clear when comparing the linear regressions superposed on each data set. Converted to a linear scale, the fits yield exponentials: $k_{\rm off} \approx (2/\text{sec})\ \exp(f\,/9.6\ \text{pN})$ for activation in Ca^{2+} and $k_{\rm off} \approx (0.007/\text{sec})\ \exp(f\,/6.8\ \text{pN})$ for activation in Mg^{2+}. Differing more than 100-fold, the force-free dissociation rates obtained from the single-molecule assay are found to be consistent with results ($k_0 \sim 1$/sec in Ca^{2+} and ~ 0.01/sec in Mg^{2+}) known from previous solution-based assays [13].

[4] Note, after we submitted this chapter, a similar version of the approach described here based on analysis of force histograms was published independently by Dudko, Hummer, and Szabo (in *Proc. Natl. Acad. Sci.*105: 15755–15760, 2008). Included in their paper are excellent applications to data from tests of nanopore unzipping of DNA hairpins and tests of unfolding a protein attached by flexible linkers to an atomic force microscope.

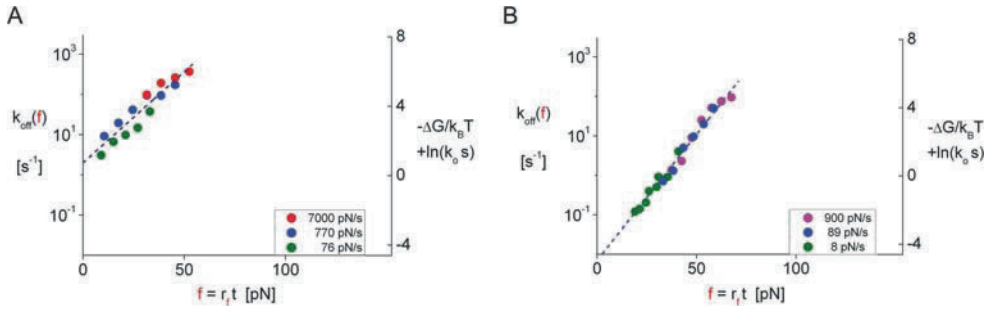


Figure 20.6. Rates of unbinding of intercellular adhesion molecule-1 (ICAM-1) attachments to $\alpha_L\beta_2$ immobilized on microspheres under slow to fast ramps of force. (Examples of the experimental estimators defining these ratios are shown in Figure 20.5 for one ramp rate.) Plotted on logarithmic scales versus unbinding force, the ratios of probability density/probability defining off rates **(A)** for tests in 2 mM Ca^{2+} at 76 and 770 pN/sec (plus those up to 55 pN at 7,000 pN/s) and **(B)** for tests in 2 mM Mg^{2+} at 7, 90, and 740 pN/sec. The regression lines superposed on the data reveal that the off-rates of single-molecular complexes in both cation environments increase exponentially over these ranges of pulling force (expressions for the off-rate dependences appear in the text). Scales at the right quantify the shifts in activation energy of ICAM-1 bonds to $\alpha_L\beta_2$ produced by force, decreasing by nearly 8 units of thermal energy (~20 kJ/M) relative to the force-free state.

Viewed in the context of Arrhenius phenomenology, that is, $\ln[k_{off}(f)] - \ln(k_0) = -\Delta G(f)/k_BT$, the single-molecule force assays in Figures 20.6A, B reveal that large reductions in the activation energies account for the enormous increases in off-rates (note the scales added at the right). The thermal force f_β for lowering an activation energy by one k_BT characterizes the length gained in a molecular transition[10,11], which, based on the regressions in Figures 20.6A, B, are ~0.4 and 0.6 nm, respectively. At the same time, the correlations to off-rates (like those plotted in Figures 20.6A, B) provide a direct check on the quality and self-consistency of each assay. Converted to analytical expressions for forward transition rates $k_{\rightarrow}(f)$, the correlations predict initial state probabilities, $S_1(f) = \exp[-(1/r_f) \int_{0\rightarrow f} k_{off}(f')\, df']$, that can be compared directly to the statistics from all experiments.

Even when one is matching data from tests over a large span in ramps, a subsequent prominent deviation from the expected initial state probability at faster ramps can reveal a significant change in the energy landscape or pathways governing the kinetic rate. This outcome is demonstrated in Figures 20.7A, B with results for the LFA-1 interactions obtained at very fast ramps, 7,000 pN/s in Ca^{2+} and 8,400 pN/s in Mg^{2+}, respectively. Coinciding with the departures of the experimental statistics $N(t)$ (open blue circles) from the statistics $N_0S_1(t)$ (black dashed curves) predicted by off-rates at ramps <1,000 pN/s, broad wings of forces in distributions are found to extend beyond >55 pN in tests at 7,000 pN/s with Ca^{2+} and beyond >75 pN in tests at 8,400 pN/s with Mg^{2+} (bin locations marked by red-rimmed white circles in Figures 20.7A, B). Including measurements of forces and times at the very fast ramps, we see in Figures 20.7C, D that a new branch of off-rates (open red circles) has emerged from the results at ramps <1,000 pN/sec. The off-rates now increase little with increase in force, suggesting a significant alteration in the kinetic pathway for dissociation. [Note that, as demonstrated by the sharp drops in $N(t)$ beyond the last bin analyzed (last red-rimmed white circle), the largest forces arising from the emergent pathway ultimately became obscured by the Poisson fraction of putative multiples marked for exclusion. Hence, to obtain an accurate picture of any emergent pathway, it is essential to have a reliable (and conservative) procedure for truncating the "raw statistics".]

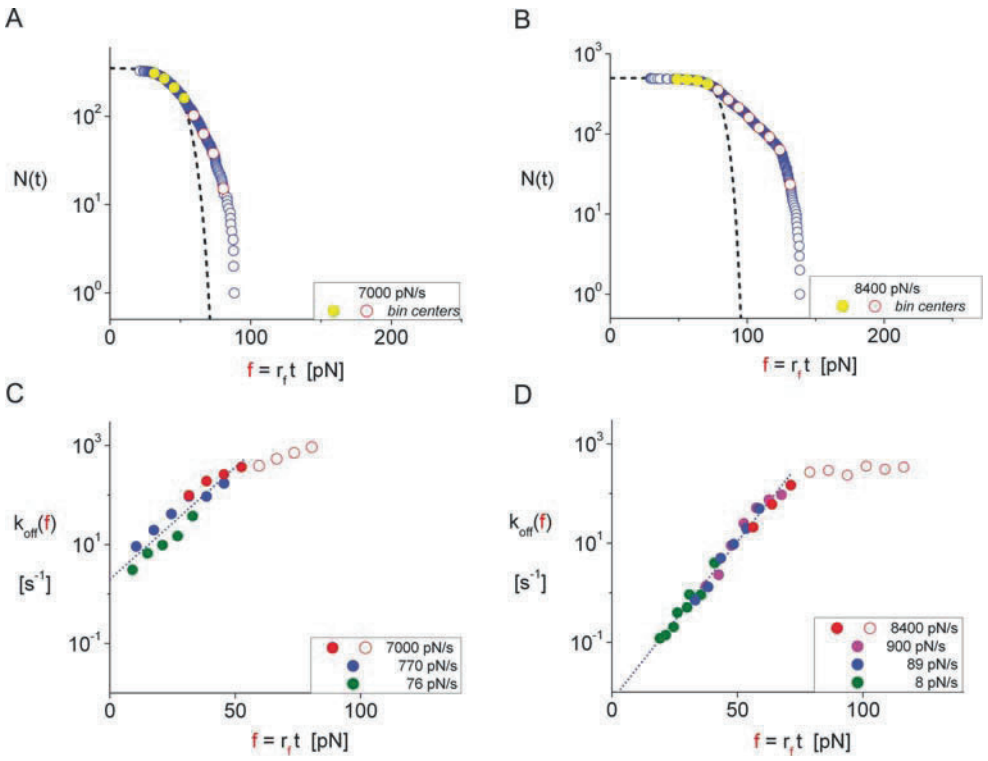


Figure 20.7. Rates of unbinding of intercellular adhesion molecule-1 attachments to $\alpha_L\beta_2$ immobilized on microspheres under extremely fast ramps of force. Plotted on a log scale versus force, experimental statistics for bond survival (*open blue circles*) for **(A)** tests in Ca^{2+} at 7,000 pN/sec and **(B)** for tests in Mg^{2+} at 8,400 pN/sec, both sets of data showing marked deviations from the statistics $N_0S_1(t)$ (*black dashed curves*) predicted by off-rate correlations at ramps <1,000 pN/sec. Coinciding with these deviations and identified by red-rimmed white circles for the bin centers, histograms of the data yield broad distributions of forces that extend well beyond the ranges expected from results in Figures 20.6A, B. By incorporating statistics from tests at these very fast ramps, the added ratios of probability density/probability (*red open circles*) are seen **(C, D)** to branch abruptly from the correlations at ramps <1,000 pN/sec, suggesting a significantly altered or reconfigured kinetic pathway for dissociation.

20.6.2. PMN Targets

For comparison to the experiments with recombinant β_2-integrin immobilized on microspheres, the following example demonstrates application of the approach to tests of LFA-1 interactions at the surfaces of human neutrophilic leukocytes (PMNs). Designed to quantify the impact on off-rates of a phenomenon called "inside-out" signaling in cells, the data presented in Figure 20.8 were taken from tests of an ICAM-1 probe against PMN surfaces following suspension in a buffer containing 2 mM Mg^{2+} plus soluble IL-8 (recombinant human chemokine, interleukin-8). With attachment frequency decreasing monotonically following exposure to IL-8, experiments had to be performed as quickly as possible (typically within a window of 5–10 min) after chamber preparation. By doing this, sufficient numbers of events were collected to clearly define narrow peaks like that seen in Figure 20.8A.

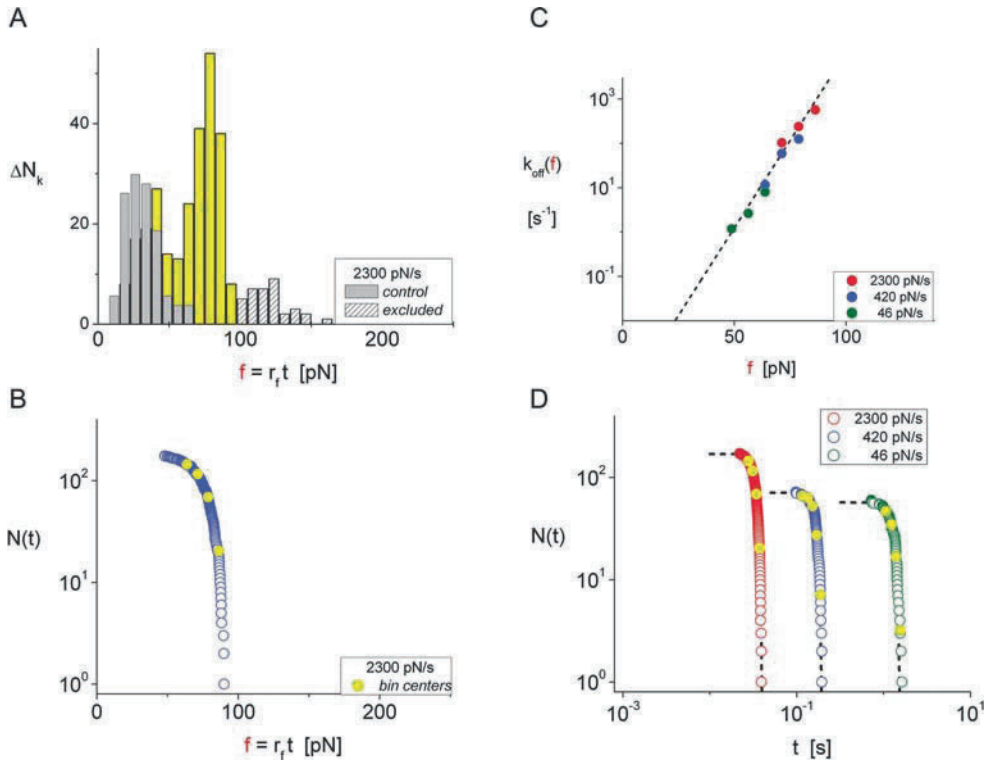


Figure 20.8. Rates of unbinding intercellular adhesion molecule-1 (ICAM-1) attachments to lymphocyte free antigen-1 (LFA-1) expressed on neutrophilic leukocytes (PMNs) under force ramps when stimulated by the chemokine interleukin-8 (IL-8). **A.** Obtained in 10 nM IL-8 plus Mg^{2+}, the histogram of "raw data" at 2,300 pN/sec illustrates the experimental estimator for probability density. **B.** Plotted directly below (on a log scale versus force are the corresponding statistics that characterize probability of bond survival (*open blue circles*). **C.** With ratios of probability density/probability from tests at three ramps in IL-8 plus Mg^{2+}, the off-rates reveal that unbinding of ICAM-1 from LFA-1 on PMNs stimulated by chemokines is dramatically slower than unbinding from LFA-1 on microspheres in Mg^{2+}. **D.** Providing a critical check on the assay for off-rates, the experimental statistics obtained at all the ramps are compared on a log-log scale versus time to the unnormalized probabilities $N_0 S_1(t)$ (*black dashed curves*) predicted by the off-rate correlation in panel C, i.e., $k_{\text{off}}(f) \approx (0.0001/\text{sec}) \exp(f/5.5\ \text{pN})$.

Narrow peaks in force ramp distributions reveal precipitous shortening in the lifetime of the bonds under the rising force, which appears as very rapid attenuation in long-time tails of the arrays $N(t)$ as shown in Figure 20.8B. Again with the ratios of probability density/probability provided by these experimental estimators, we plot the values on a logarithmic scale in Figure 20.8C to assay the off-rates of ICAM-1 from LFA-1 on chemokine-stimulated PMNs in Mg^{2+}. When compared to the results in Figure 20.6B for LFA-1 on microspheres in Mg^{2+}, the outcome demonstrates that chemokine stimulation and the "inside-out" signaling process dramatically increase the lifetime and strength of a single LFA-1 interaction. Again as a check on quality of the assay, the experimental statistics for different ramps are compared on a log-log scale versus time in Figure 20.8D to the unnormalized probabilities $N_0 S_1(t)$ (black dashed curves) predicted by the correlation in Figure 20.8C.

20.7. Experimental Example: Unfolding/Refolding of a Polyprotein with Force Ramps

Demonstrating the diversity of this approach, our final examples illustrate applications to unfolding and refolding of polyproteins. Created by Jane Clark at Cambridge University [14], the polyprotein in this example was an eight-mer sequence of four heterodimers: the R16 triple-helical repeat from chicken brain spectrin, followed by an I27 domain from cardiac muscle titin. Pulling on these polyproteins with an AFM, the creators showed that the spectrin triple-helical repeats unfold at forces between ∼30 and 50 pN under pulling speeds in a range from $\sim 1\text{–}4 \times 10^3$ nm/sec [14]. Yet, as expected from earlier studies (see references in Ref. 15), the I27 domains remain tightly folded until several-times larger forces are reached (∼150–180 pN). Less well established in probe tests of weak domains like R16, the objective here is to demonstrate a direct assay for rates of unfolding and refolding polyproteins using the experimental statistics and histograms of transitions obtained under ascending and descending ramps of force, respectively.

To achieve the precision in force and time needed to quantify these kinetic rates, the poly-R16/I27 construct was tested with two arrangements of a laser OTFP, one providing data at frequencies of 1,500/sec for fast-pulling experiments (see Figure 20.2 and Ref. 3) and the other providing data at slower frequencies of ∼200/sec but with subnanometer precision in chain length displacements (see Figure 20.9 and Refs. 4 and 5). Tethered to glass targets through linkages to a cysteine residue at the C-terminus, the polyprotein attachments in tests were obtained by capturing the polyhistidine residues at the N-terminus with an anti-His–linked OTFP microsphere. Illustrating the kinematics of unfolding and refolding four R16 domains, Figure 20.9 shows the changes in length (upper panel) of a single poly-R16/I27 protein produced by the cycle of positive–zero–negative force ramps seen in the lower panel.

20.7.1. Unfolding Kinetics

Examining unfolding first, we focus on forces obtained from traces showing one or more unfolding events at two ramp rates (10 and 473 pN/sec). Unfortunately, when tested at the fast ramp, preemptive failure of the noncovalent linkages left just over half of the traces showing only one unfolding event, with the next significant population showing two unfolding events; taken together, the two types of traces accounted for nearly 90% of the results. Based on statistics of the first unfolding events at 473 pN/sec, Figures 20.10A, B demonstrate the force histogram and the experimental statistics characterizing the fastest unfolding kinetics in the OTFP experiments. For more slowly unfolding kinetics, we use the forces for two unfolding events from experiments performed with the high-resolution instrument at the slow ramp 10 pN/sec, where most traces contained more than two unfolding events. Under slow loading, longer times are needed to unfold domains and the events occur at lower forces. Finding ratios of probability density/probability from both sets of data, we plot two assays for unfolding rates on a logarithmic scale versus force in Figure 20.10C. As shown by gray dashed lines in Figure 20.10C, both assays correlate well with a common exponential dependence on force [i.e., $\sim\exp(f/2.8\text{ pN})$], separated by approximately twofold difference in the preexponential factors describing "spontaneous" unfolding rates. This shift is qualitatively consistent with the different numbers of unfolding events treated in the two cases. However, not knowing the total number of folded domains in the attachments prior to test, we can only estimate the frequency for unfolding a single spectrin R-16 domain. Thus, assuming that four folded and

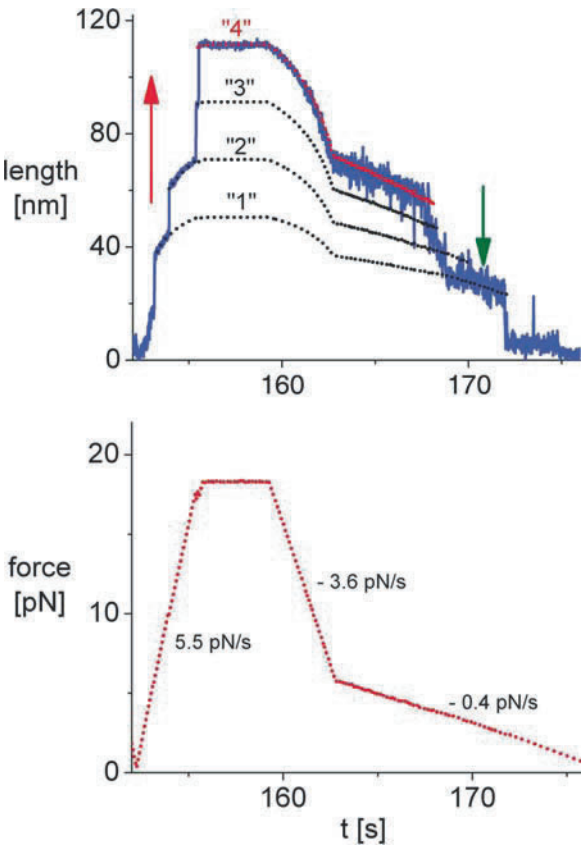


Figure 20.9. Changes of length (*blue traces, upper panel*) produced in a single poly-R16/I27 protein by a sequence of positive– zero–negative force ramps (*red traces, lower panel*). Performing repeated cycles of pulling and relaxing attached chains parallel to the optical axis of an inverted microscope, the OTFP apparatus employed reflection interference contrast optics and unique image analysis software to measure displacements of the probe microsphere with subnanometer precision [4,5]. Centered well above the coverglass floor of a microchamber, the strength of the trapping potential was varied by controlling the laser intensity. Because of the large distance to the trap center and the soft confinement along the optical axis, the force history applied to the microsphere was unaffected by the increases and decreases in length of the polyprotein, appearing only as extremely small blips in the red force trace. Correlated to the length increases in unfolding (*upward red arrow*) and decreases in refolding (*downward green arrow*), separations between the black dashed curves are predicted by adding a length $L_c \sim 37$ nm at each step to the thermally fluctuating contour length described by a "worm-like chain" with persistence length $b \sim 0.7$ nm while, at the same time, substracting 5 nm to account for the folded domain.

noninteracting domains were present in each attachment, the correlation for first unfolding events in the fast ramp test (red closed circles, Figure 20.10D) implies an unfolding rate of $0.009/\text{sec} \div 4 \approx 0.0023/\text{sec}$ for a single domain, which is close to results from solution-based assays [16]. Again checking the quality and self -consistency of the assays, experimental statistics from both tests are compared on a log-log scale versus time in Figure 20.10D to the unnormalized probabilities $N_0 S_1(t)$ (black dashed curves) predicted by the probability of folded domains, $S_1(t) = \exp[-(1/r_f)\int_{0\to f} k_{\text{unfld}}(f')\,df']$, and the correlations to unfolding rates at each ramp.

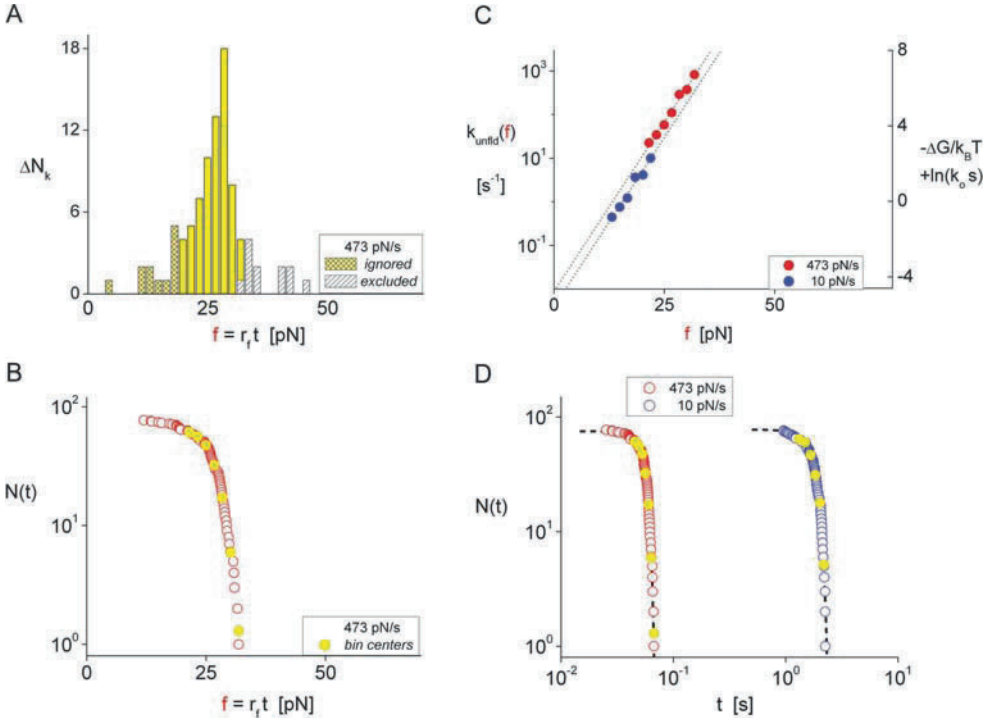


Figure 20.10. Rates of unfolding weak domains in poly-R16/I27 proteins under ramps of force. Used to assay the fastest unfolding rates at 473 pN/sec [3], the histogram of first unfolding events **(A)** provides the estimator for probability density; the corresponding array of statistics **(B)** provides the unnormalized estimator for the probability of remaining folded at different forces. [Again of no consequence in our analysis, we ignore the bins (marked by yellow meshes at low forces). The events in the cross-hatched bins (∼13% of total) have been truncated from the "raw statistics" to obtain the array $N(t)$. The Poisson fraction of putative multiples was determined from an attachment frequency of ∼26%.] Plotting the ratios of probability density/probability obtained at a fast and a slow [4] ramp on a logarithmic scale **(C)**, we see that the unfolding rates produced by the two ramps follow parallel lines (slope ∼2.8 pN) with a slight separation reflecting an approximate twofold difference in force-free unfolding rates (0.009/sec vs. 0.004/sec). **D.** Providing a critical check on quality of the assay for unfolding rates, experimental statistics for probabilities of remaining folded at the fast and slow ramps are compared on a log-log scale versus time to the unnormalized probabilities $N_0 S_1(t)$ (*black dashed curves*) predicted by the linear fits for unfolding rates (gray lines in panel C), that is, $N_0 \exp[-(1/r_f) \int_{0\to f} k_{\mathrm{unfld}}(f')\, \mathrm{d}f']$. The close match between the data and predictions reveals the major impact of ramp speed on unfolding kinetics.

20.7.2. Refolding Kinetics

Emphasizing the generality of our approach, we end by demonstrating that statistics of reverse transitions obtained with *negative* force ramps also provide a direct assay for refolding rates. Starting from an initial force f_0, each negative ramp applied to the unfolded chain in Figure 20.9 was produced by a fixed period of steady reduction in the trapping force at a rate $-r_f$, that is, $f(t) = f_0 - r_f\, t$. Accurately controlled in this way by laser power, the unloading process allowed the collapse and refolding of polyprotein domains previously unfolded by force. Usually, refolding is only likely to occur in a reasonable period of time when low levels of force are reached. Yet, at the same time, much larger forces are usually needed to initially unfold the domains. Thus, the sequence of positive– zero–*two* negative force ramps

was developed for the purpose of unfolding and refolding the R-16 domains, as illustrated in Figure 20.9.

In this example of refolding, we will use estimators for ratios of probability density/probability that characterize refolding from the level "2" to "1" identified by the labels above black dotted lines sketched in Figure 20.9. As seen throughout the slow negative ramp in Figure 20.9, transient up/down jumps appear along the length trace at all levels. Summarized by histograms in Figure 20.11A for unloading at –0.37 pN/s, the downward refolding events from a state of two to one unfolded domain ("2>1") appear as yellow bins, the intermittent-upward unfolding events returning to two unfolded domains ("2<1") as blue bins, and the downward refolding events from three to two unfolded domains ("3>2") as red bins. [The histrogram ("3<2") containing a few transient returns to the state of three unfolded domains is left out of Figure 20.11A for clarity.] Careful analysis (4) of the down/up jumps (e.g. "3<2" and "2<1") has revealed that refolding R-16 domains involves a long-lived metastable intermediate (mean lifetime ~0.3 s). Although complicated by the multi-level dynamics, we show that approximating the refolding process as a two-state transition provides a lower bound to the rates of refolding. As such, the statistical array N(f) for refolding from "2>1" plotted in Figure 20.11B becomes the un-normalized estimator for probability of two unfolded domains, which diminishes monotonically with increase in time or equivalently decrease in force. Using the arrays N(f) and histograms of transitions ("2>1") obtained with three negative ramps, we have plotted values (closed red, green, blue circles) of the approximation for ratios of probability density/probability on a logarithmic scale as a function of force squared (f^2/k_BT) and on a linear scale of rate versus force in Figure 20.11D. The estimates of refolding rates for the three negative ramps follow a negatively sloped line with a zero force intercept implying a spontaneous refolding rate of ~400/s from a state of two unfolded domains. Also plotted in Figure 20.11B is the un-normalized probability "N S2" derived from summation of all transitions to/from level "2" at –0.37 pN/s. Obtaining un-normalized probabilities in this way for each negative force ramp (details to appear elsewhere), and computing the true estimators for probability density/probability, the outcome (red dotted lines in Figures 20.11C, D) reveals that refolding actually occurs at much faster rates because of the transitions to/from the intervening metastable state. As indicated by the true estimators for probability density/probability (open red circles) from the fast unloading rate of –1.36 pN/s, the true refolding rates and the two-state approximation begin to converge when refolding rates greatly exceed the rate of transient unfolding from the metastable state. The reason for correlating the results to f^2/k_BT is that the unfolded domains are very floppy at low forces, and thus the energetics governing changes of end-to-end length is described by a harmonic potential [$\sim -f^2/(2\kappa_{chain})$]. Also following from polymer physics, the spring constant of the floppy chain κ_{chain} approximately equals the spring constant for a single unfolded domain κ_{unfld} divided by the number of unfolded domains, $\kappa_{chain} \approx \kappa_{unfld}/n_{unfld}$. Hence, the activation energy for refolding one domain is expected to be augmented by the difference in elastic energy required to shorten the chain, $\Delta[f^2/(2\kappa_{chain})]/\Delta n \approx f^2/(2\kappa_{unfld})$. Correlated to f^2/k_BT, the slopes obtained from linear regressions to the approximate and true refolding rates in Figure 20.11C yield apparent spring constants (0.22 and 0.28 pN/nm respectively) that characterize the apparent harmonic-energy barrier. These slopes bracket the harmonic stiffness, $3k_BT/(2bL_c) \approx 0.236$ pN/nm, based on persistence length ($b \sim 0.7$ nm) and contour length ($L_c \sim 37$ nm) used to match the force-distance response of the poly protein with all four R-16 domains unfolded. Converting the refolding kinetics in Figure 20.11C to a linear-linear scale, Figure 20.11D demonstrates the major suppression of refolding kinetics that accompanies pulling force.

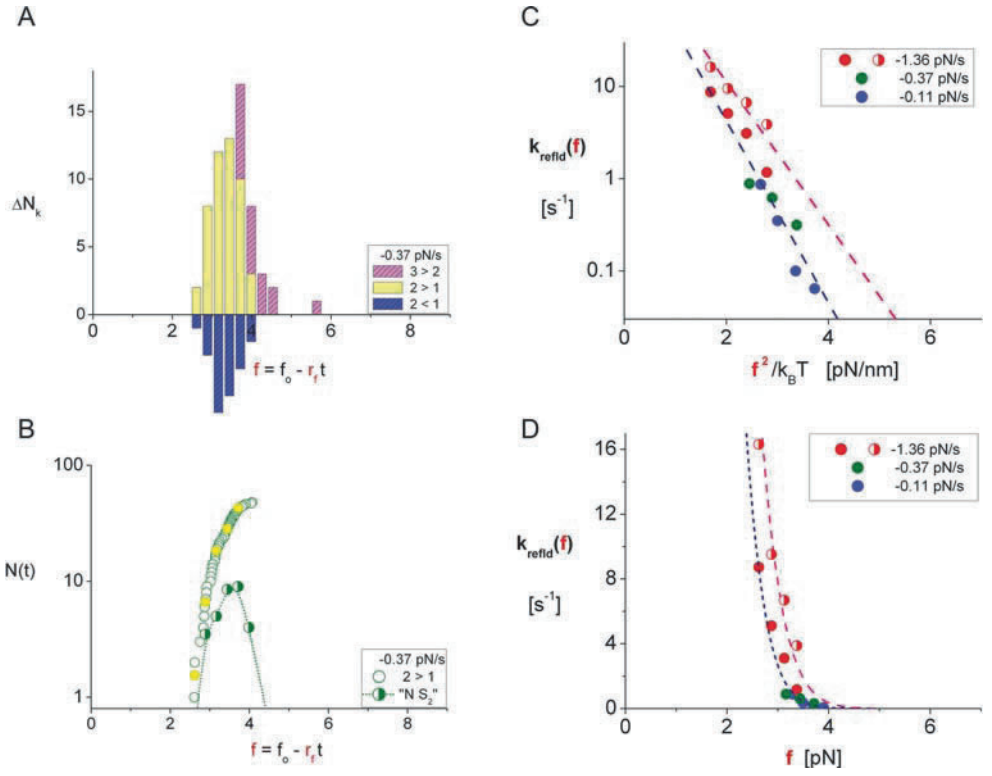


Figure 20.11. Rates of refolding (previously unfolded) domains in poly-R16/I27 proteins under *negative* ramps of force. Needed to assay the rates of refolding from level "2" to "1" in Fig. 9, histograms in **A** (plotted as forces) illustrate refolding and unfolding transitions that determine the un-normalized probability for being in a state with two unfolded and two folded domains under a negative ramp of –0.37 pN/s. Yellow bins show the downward refolding transitions to the state of one unfolded domain ("2>1"); blue bins show the intermittent-upward unfolding transitions returning to the state of two unfolded domains ("2<1"). Red bins show the downward refolding transitions to the state of two unfolded domains ("3>2") that seed level "2". [Not shown for clarity is the histrogram ("3<2") of transient returns to the state of three unfolded domains.] The statistical array N(f) of forces for refolding from "2>1" is plotted in **B** along with the un-normalized probability "N S_2" for being in level "2" computed from summation of all the transitions to/from this level. Naively approximating the refolding as a two-state process, the array N(f) plotted in **B** becomes the un-normalized estimator for probability of two unfolded domains, which diminishes monotonically with increase in time corresponding to the decrease in force. Using the arrays N(f) and histograms of transitions ("2>1") obtained with three negative ramps, we have plotted values (closed red, green, blue circles) for the approximate ratios of probability density/probability on a logarithmic scale as a function of force squared (f^2/k_BT) **(C)** and on a linear-linear scale versus force in **D**. The approximate refolding rates follow a negatively sloped line with a zero force intercept implying a spontaneous refolding rate of ~400/s given two unfolded domains. Also plotted for comparison, the outcome (red dotted lines) derived from the full multilevel analysis (details to appear elsewhere) shows that the two-state approximation provides a lower bound for the refolding rates at all forces. Illustrated by the true ratios of probability density/probability (open red circles) from the fast unloading rate of –1.36 pN/s, the true refolding rates approach the two-state approximation once refolding rates greatly exceed the rate of transient unfolding from an intervening metastable state.

Acknowledgment

This work was supported by National Institutes of Health grants HL65333 and HL31579.

References

1. Evans E, Ritchie K, Merkel R (1995) Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces. *Biophys. J*. 68: 2580–2587.
2. Evans E, Heinrich V, Leung A, Kinoshita K (2005) Nano-to-micro scale dynamics of P-selectin detachment from leukocyte interfaces: I. Membrane separation from the cytoskeleton. *Biophys. J*. 88: 2288–2298.
3. Halvorsen KA (2007) Probing weak single molecule interactions: development and demonstration of a new instrument. PhD dissertation, College of Engineering, Boston University, Boston MA.
4. Wong WP (2006) Exploring single-molecule interactions through 3D optical trapping and tracking: from thermal noise to protein refolding. PhD dissertation, Department of Physics, Harvard University, Cambridge MA.
5. Heinrich V, Wong WP, Halvorsen K, Evans E (2008) Imaging biomolecular interactions by fast three-dimensional tracking of laser-confined carrier particles. *Langmuir* 24: 1194–1203.
6. Williams P, Evans E (2002) In *Physics of Bio-Molecules and Cells*, *Les Houches: Ecoles d'Ete de Physique Theorique*, Vol. 75 (Springer, New York), pp. 186–203.
7. Lupher ML Jr, Harris EA, Beals CR, Sui L-M, Liddington RC, Staunton DE (2001) Cellular activation of leukocyte function-associated antigen-1 and its affinity are regulated at the I domain allosteric site. *J. Immunol*. 167: 1431–1439.
8. Beals CR, Edwards AC, Gottschalk RJ, Kuijpers TW, Staunton DE (2001) CD18 activation epitopes induced by leukocyte activation. *J. Immunol*. 167: 6113–6122.
9. Huth JR, Olejniczak ET, Mendoza R, Liang H, Harris EA, Lupher ML Jr., Wilson AE, Fesik SW, Staunton DE (2000) NMR and mutagenesis evidence for an I domain allosteric site that regulates lymphocyte function-associated antigen 1 ligand binding. *Proc. Natl. Acad. Sci. USA* 97: 5231–5236.
10. Evans E, Ritchie K (1997) Dynamic strength of molecular adhesion bonds. *Biophys. J*. 72: 1541–1555.
11. Evans E (2001) Probing the relation between force – lifetime – and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct*. 30: 105–128.
12. Evans E, Williams P (2002) In *Physics of Bio-Molecules and Cells. Les Houches: Ecoles d'Ete de Physique Theorique*, Vol. 75 (Springer, New York), pp. 145–185.
13. Labadia ME, Jeanfavre DD, Caviness GO, Morelock MM (1998) Molecular regulation of the interaction between leukocyte function-associated antigen-1 and soluble ICAM-1 by divalent metal cations. *J. Immunol*. 161: 836–842.
14. Randles LG, Rounsevell RWS, Clarke J (2007) Spectrin domains lose cooperativity in forced unfolding. *Biophys. J*. 92: 571–577.
15. Forman JR, Clarke J (2007) Mechanical unfolding of proteins: insights into biology, structure and folding. *Curr. Opin. Struct. Biol*. 17: 58–66.
16. Scott KA, Batey S, Hootan KA, Clarke J (2004) The folding of spectrin domains I: Wild-type domains have the same stability but very different kinetic properties. *J. Mol. Biol*. 344: 195–205.

21

Single-Molecule Recognition: Extracting Information from Individual Binding Events and Their Correlation

Cheng Zhu, Veronika I. Zarnitsyna, Wei Chen, and Krishna Sarangapani

Abstract Communication between cells via their surface receptors can be studied by ultrasensitive single-molecule techniques designed to operate with two apposing surfaces. Two-dimensional (2D) kinetics measured with this approach is physically distinct from and biologically more relevant than three-dimensional kinetics with at least one interacting molecule in the fluid phase. Two assays for measuring 2D kinectics are reviewed: the adhesion frequency assay and the thermal fluctuation assay. Both techniques allow one to extract force-free 2D kinetics parameters. The adhesion frequency assay measures the binding frequency as a function of contact duration, which is nonlinearly fitted with a probabilistic model. The thermal fluctuation assay detects bond formation/dissociation by monitoring the reduction/resumption of the thermal fluctuation of the force sensor. Both techniques record the outcomes in a time sequence of events. Different events in the sequence can be independent from or correlated between each other, depending on the molecular interaction. The strategy for analyzing statistical properties of the sequence by a mathematical model to test whether the i.i.d. assumption (i.e., all measurements in the sequence are identical yet independent from each other) is valid in the interaction in question is reviewed.

21.1. Introduction

Communication between cells with each other and their environment is mediated via specific molecules on their surfaces called receptors. Although unraveled genomes for key organisms have provided a lot of information about different surface molecules as

C. Zhu, K. Sarangapani • Coulter Department of Biomedical Engineering and Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
V. I. Zarnitsyna • Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
W. Chen • Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

P. Hinterdorfer, A. van Oijen (eds.), *Handbook of Single-Molecule Biophysics*,
DOI 10.1007/978-0-387-76497-9_21,

corresponding gene products, only a fraction of them have a function assigned, and even fewer have been studied in detail. Traditionally, receptor–ligand interactions are characterized by ensemble assays with at least one of the molecules purified from the cell membrane and added to the fluid phase, namely, three-dimensional (3D) assays, e.g., using the surface plasmon resonance technique. This interaction is described in terms of a chemical reaction through receptor and ligand concentrations, kinetic rates, and binding affinity. While sufficient for a proper description in homogeneous solution, these properties by themselves are often insufficient for determining the interaction of the same molecules residing in their natural environment on the cell membranes. This is because cell surface molecules can be clustered, partitioned in cell surface structures such as membrane rafts, or linked to cytoskeleton directly or via other docking molecules, which can impact binding by signaling that changes molecular conformations. On the cell, molecules of interest may be abundant but inactive due to specific cellular environment, or the same molecules may exhibit different activities in different areas or times. These features make ensemble assays insufficient, even if they are carried out on the cell surface, because the properties so measured represent ensemble averages only. It also calls for single-molecule studies because these experiments probe molecular interactions one by one, thereby allowing for measurement of not only average properties, but also their distributions.

Another important feature of receptor-mediated cell adhesion is that molecular interactions take place in two dimensions (2D) as both receptors and ligands are anchored to the respective surfaces of two apposing cells or a cell and a substrate surface. This situation is ideal for single-molecule force techniques because the sensitive force probes can be functionalized with interacting molecules. A molecular interaction manifests as a mechanical force through a receptor–ligand bond that physically connects two surfaces, one of which can be the force probe (the other is referred to as the target surface in this chapter). Single-molecule biomechanical experiments with atomic force microscopy (AFM) or other ultrasensitive force techniques for measuring unbinding forces for dynamic force spectroscopy analysis or unbinding of protein domains are discussed in other chapters of this book. In this chapter, we will focus on extracting other information from the binding events. Specifically, we will discuss an adhesion frequency assay (Chesla et al. 1998) and a thermal fluctuation assay (Chen et al. 2008) along with the analyses that allow for extracting kinetic information of the receptor–ligand interaction from these experiments. Both assays obtain the information from sequential binding events. Statistical analysis of the data has previously been based on the assumption that all measurements in a sequence are identical yet independent from each other, i.e., the i.i.d. assumption. If this assumption is violated, i.e., measurements from different positions in a test sequence are correlated, data should be segregated into subgroups for separate analyses. Thus, we will also discuss a model that allows us to check the validity of, and to measure the deviation from the i.i.d. assumption (Zarnitsyna et al. 2007).

21.2. Adhesion Frequency Assay

The simplest mechanical measurement is the detection of the presence of adhesion mediated by specific receptor–ligand bond(s) at a given time, regardless of how many bonds are involved, how long they last, and how big a force is required to detach the adhesion. It is desirable, but not required, that the observed adhesion is mediated by a single bond. To

achieve single-bond measurement requires the use of a force probe with sufficient sensitivity for single-bond detection and experimental conditions that limit multibond interactions, e.g., using low site densities (compared to the 2D K_d) of receptors and ligands on the respective surfaces such that binding events become rare, which is a necessary but not a sufficient condition (Zhu et al. 2002). Although this simplest measurement produces only a random binary readout of either adhesion (scored $X_i = 1$ for the ith test) or no adhesion (scored $X_i = 0$ for the ith test), it can be used to estimate the probability of adhesion, e.g., from the frequency of adhesion, i.e., the average of adhesion scores, from repeated measurements. When adhesion becomes infrequent and requires a probabilistic description for its presence, the number of molecular bonds that mediate adhesion must be low and should follow a Poisson distribution (Piper et al. 1998):

$$P_n = \frac{\langle n \rangle^a}{n!} \exp(-\langle n \rangle) \tag{21.1}$$

where p_n is the probability of having n bonds and $<n>$ is the average number of bonds, which can be smaller than 1 under infrequent-adhesion condition. The probability of adhesion P_a is

$$P_a = 1 - p_0 = 1 - \exp(-\langle n \rangle) \tag{21.2}$$

The average number of bonds is a statistic equivalent to the density of bonds, which increases with increasing densities of receptors and ligands due to mass action and changes during a kinetic process depending on time and the kinetic rate constants. For a simple bimolecular interaction,

$$R + L \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} B \tag{21.3}$$

where R, L, and B denote, respectively, the receptor, ligand, and bond, k_{+1} and k_{-1} denote the respective on- and off-rate constants, $<n>$ obeys the following kinetic equation and initial condition:

$$\frac{d\langle n \rangle}{dt_e} = m_r m_l A_c k_{+1} - k_{-1} \langle n \rangle \quad \text{and} \quad \langle n \rangle\big|_{t_e=0} = 0 \tag{21.4}$$

where A_c is the contact area and t_c is the contact duration. The solution is

$$\langle n \rangle = m_r m_l A_c K_a [1 - \exp(-k_{-1} t_c)] \tag{21.5}$$

where $K_a = k_{+1}/k_{-1}$ is the binding affinity. Substituting Eq. (21.5) into Eq. (21.2) yields (Chesla et al. 1998)

$$P_a = 1 - \exp\{-m_r m_l A_c K_a [1 - \exp(-k_{-1} t_c)]\} \tag{21.6}$$

Unlike 3D assays, where soluble molecules in the fluid phase can diffuse to the vicinity of the counter-molecules, 2D assay requires that the force probe and the target surface be brought into contact because interaction would not be physically possible if the two surfaces are separated by distances greater than the span of the receptor–ligand complex. Thus, the initiation of interaction is, in a sense, "staged" by the experimenter who puts the two surfaces into contact and controls the area and duration of the contact. Equation (21.6) states that the likelihood of

observing adhesion, P_a, depends on the time when the observation is made relative to the time when the contact is made, i.e., the contact duration, t_c. If the adhesion frequency is measured over a range of contact durations, fitting Eq. (21.6) to the measured P_a vs. t_c binding curve then allows estimation of the 2D kinetic rates and binding affinity, provided that the receptor and ligand densities are measured from independent experiments.

A set of example data of the adhesion frequency assay for measuring 2D binding kinetics is shown in Figure 21.1 (Huang et al. 2007). The receptor–ligand pair studied was CD8, a coreceptor expressed on T lymphocytes from monoclonal T cell receptor (TCR) transgenic mice, and a noncognate peptide bound to mouse major histocompatibility complex (pMHC) molecule coated on the surface of a human red blood cell (RBC). The RBC was aspirated by a micropipette via a small suction pressure (see Figure 21.2), which formed an adhesion sensor as originally developed by Evans (Evans et al. 1991). Note that a single set of kinetics parameters enable the model to best fit the data with varied surface densities of receptors and ligands.

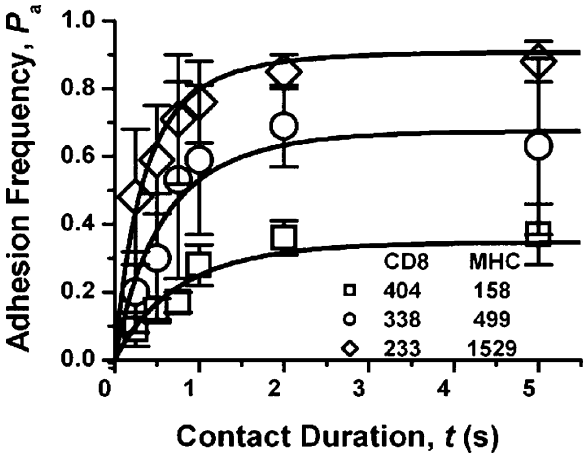


Figure 21.1. Adhesion frequency vs. contact duration plots. Data (points) from three sets of site densities (indicated), each obtained from 15 to 29 pairs of CD8+ T cells and noncognate pMHC-coated red blood cells contacting 50 times each, were presented as mean ± SEM at each contact duration. Equations, (21.6) was globally fit (curves) to all three data sets with $k_{-1} = 1.12 \pm 0.15\ s^{-1}$ and $A_c K_a = 6.66 \pm 0.36 \times 10^{-6}\ \mu m^4$. [Reproduced from Huang et al. (2007) with permission.]

Although the adhesion frequency assay was originally proposed using the micropipette technique (Figure 21.2) (Chesla et al. 1998), it can be implemented using other ultrasensitive force techniques, such as optical tweezers (Thoumine et al. 2000), biomembrane force probe (BFP; see Figure 21.10 in Section 21.3) (Chen et al. 2008), and AFM (Figure 21.3). The receptor–ligand pairs studied in the AFM experiment were L- or P-selectin reconstituted in glass-supported lipid bilayer (Marshall et al. 2003) interacting with an endoglycan–immunoglobulin chimeric molecule captured on the AFM cantilever tip by protein G (Sarangapani et al. 2004).

Thus, despite the simple binary outcome of each single-molecule biomechanical testing, proper design of the experiment and analysis of the data result in useful information.

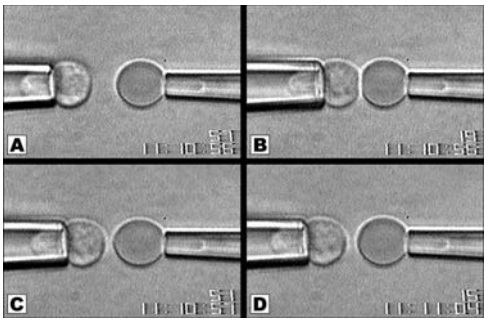


Figure 21.2. Micropipette-aspirated red blood cell (RBC) adhesion sensor. (A, B) A T cell was aspirated by a micropipette (*left*) and driven by a computer-programmed piezoelectric translator to make a controlled contact with a RBC coated with pMHC ligand, held stationary by another pipette (*right*). At the end of the contact phase, the translator retracted the left pipette to the starting position. (C) If adhesion results, the RBC apex would be stretched. (D) If adhesion is not present at the end of the contact, the RBC would restore its shape without extension. Images in panels C and D are chosen at the same instant after retraction. [Reproduced from Zarnitsyna et al. (2007) with permission.]

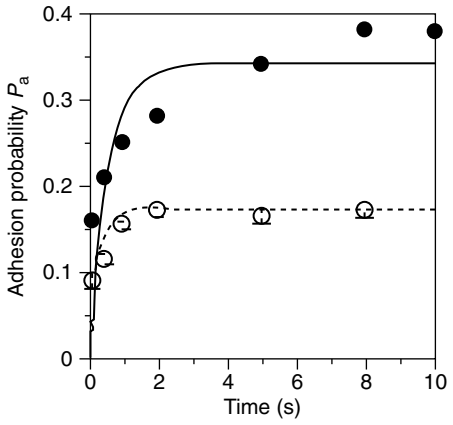


Figure 21.3. Adhesion probability vs. contact duration plots. Endoglycan–immunoglobulin captured on an atomic force microscope cantilever tip precoated with protein G (Sarangapani et al. 2004) was driven to repeatedly contact L-selectin (○) or P-selection (●) reconstituted in glass-supported lipid bilayer (Marshall et al. 2003). Data (points) were fitted (curves) by Eq. (21.6) with $k_{-1} = 3.55\ \mathrm{s}^{-1}$ and $m_t m_1 A_c K_a = 0.19$ for L-selectin and $k_{-1} = 1.57\ \mathrm{s}^{-1}$ and $m_t m_1 A_c K_a = 0.42$ for P-selectin.

21.3. Thermal Fluctuation Assay

The adhesion frequency assay discussed in the preceding section extracts kinetic information from the dependence of adhesion frequency P_a on contact duration t_c (Chesla et al. 1998). Adhesion is measured mechanically by separating the force probe (pressurized RBC or AFM cantilever) from the target to detect the presence of a receptor–ligand bond or bonds that connect the two surfaces at the end of a contact but not when a bond forms or dissociates during the contact. Therefore, kinetics of molecular interaction must be inferred from fitting a model, e.g., Eq. (21.6), to the P_a vs. t_c data (Figure 21.1 and 21.3) (Chesla et al. 1998). By

comparison, a recently developed thermal fluctuation assay can pinpoint the association and dissociation events at the single-bond level during the contact period without separating the two surfaces (Chen et al. 2008).

The idea of the thermal fluctuation assay is as follows. Due to their ultrasensitivity, force probes used for single-molecule experiments are usually susceptible to thermal fluctuations. Indeed, thermal fluctuation analysis is commonly used to calibrate the spring constant of the force probe, k_p, based on the equipartition theorem (Hutter and Bechhoefer 1993)

$$\frac{1}{2}k_P\sigma^2 = \frac{1}{2}k_B T \tag{21.7}$$

where σ is the standard deviation of the thermally excited random displacements of the force probe, k_B is the Boltzmann constant, and T is the absolute temperature. Formation of a molecular bond spanned across the gap between the force probe and the target physically connects the two surfaces together, which reduces the thermal fluctuation of the force probe. This is because bond formation adds mechanically the molecular spring k_m to the force probe spring k_p, thereby resulting in a stiffer system spring k_s that is equivalent to two springs in parallel (Wu et al. 2005):

$$k_s = k_p + k_m \tag{21.8}$$

Indeed, this idea has been used to measure the molecular elasticity of L- and P-selectins from the decreased thermal fluctuations in an AFM cantilever caused by bond formation (Marshall et al. 2006). In this work, the AFM cantilevers have a nominal spring constant of $k_p \approx 10$ pN/nm and the molecular spring constants are $k_m \approx 4$ and 1 pN/nm for L- and P-selectins, respectively (Marshall et al. 2006). Adding a molecular spring constant of this magnitude in parallel to the cantilever spring constant results in a 40% or 10% increase in the system spring constant and a 40% or 10% decrease in the thermal fluctuations. This is shown in Figure 21.4, where cantilever deflection versus time (Figure 21.4A and B), 15-point sliding deviation of the cantilever deflection versus time (Figure 21.4C and D), and their corresponding histograms (Figure 21.4E and F) for L-selectin (Figure 21.4A,C and E) and P-selectin (Figure 21.4B,D and F), are respectively shown. The differences in the deflection fluctuations (as directly visualized or measured by sliding standard deviations) between the free and bound cantilevers are more distinct for L-selectin (Figure 21.4A and C) than for P-selectin (Figure 21.4B and D). Two separate histograms are overlaid in each panel, one corresponds to the thermal fluctuations of a free cantilever and the other to those of a cantilever linked to the surface via a selectin–ligand bond (Figure 21.4E and F). Using Eq. (21.8), we can determine the molecular spring constants for L-selectin and P-selectin respectively from the differences in the peak locations of the two histograms from the respective panels. However, there are significant overlaps between the subpopulations under the two peaks due to the small differences in the spring constants of the free and bound cantilevers. The difference is greater for L-selectin than for P-selectin, resulting in a more distinct separation of the two subpopulations in the former case (Figure 21.4E) than the latter case (Figure 21.4F) because restraining the cantilever tip by the stiffer L-selectin increases the system spring constant more than that by the softer P-selectin. However, the significant overlaps prevent the events of bond formation and dissociation from being detected reliably with sufficient temporal resolution, i.e., without averaging over long periods.

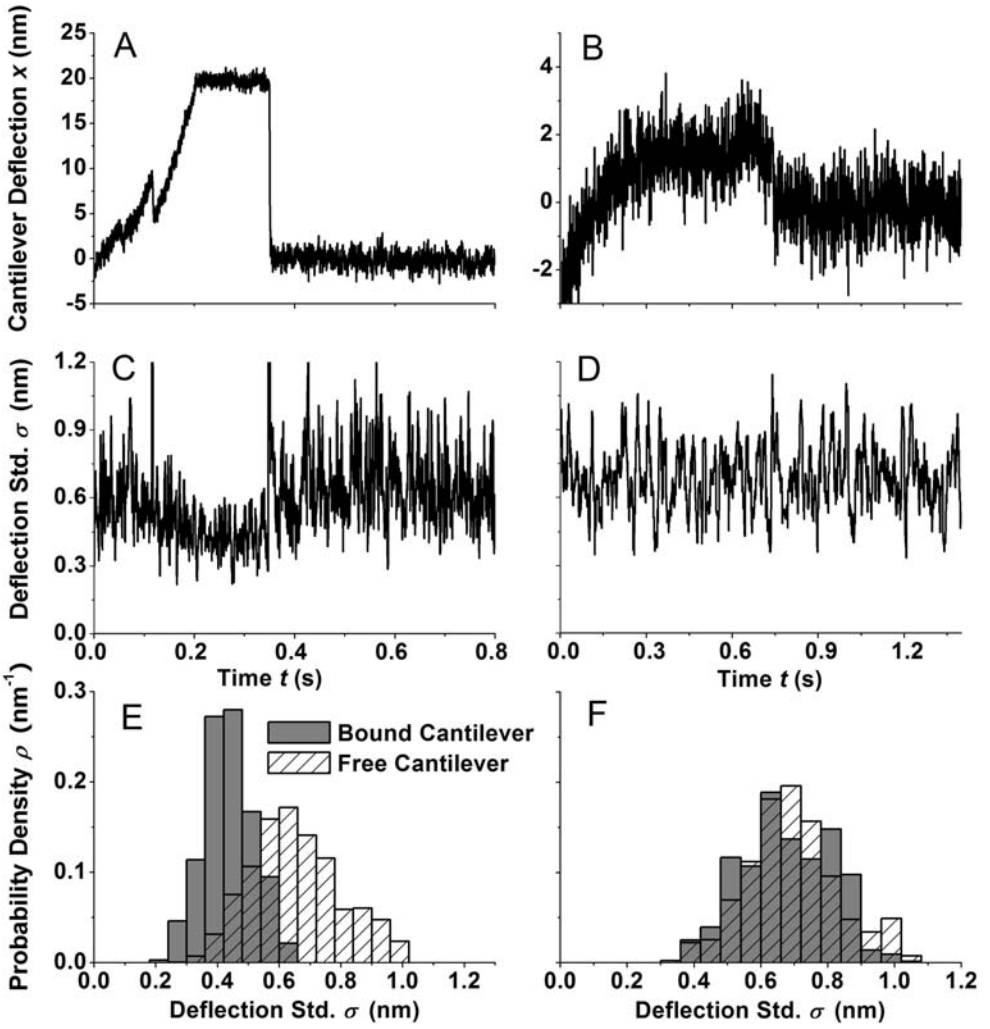


Figure 21.4. Comparison of thermal fluctuations of bound and free cantilever. Cantilever deflection x vs. time t curve of L-selectin/PSGL-1 interaction (A) or P-selectin/endoglycan interaction (B) in an experiment that measures bond lifetime at constant force. (C, D) Sliding standard deviations σ of 15 consecutive points of the cantilever deflection data in panels A and B, respectively. (E, F) Histograms of the σ data in panels C and D, respectively.

This difficulty can be overcome by using softer force probes with much smaller spring constants than that of the molecules, such as optical tweezers or BFP (Chen et al. 2008). A BFP is a high-tech version of a micropipette system, which includes a probe bead glued to the apex of the aspirated RBC as a picoforce transducer (Figure 21.5, left). Observed under an inverted microscope, the position of the probe bead is tracked by a high-speed camera with 0.8-msec temporal resolution. With advanced image analysis software, the spatial resolution of the bead position tracking can be achieved at ~5 nm. A target bead is aspirated by an apposing micropipette (Figure 21.5, right). A piezo translator with capacitive feedback control brings the receptor-coated target bead to briefly contact the ligand-coated probe bead, retracts

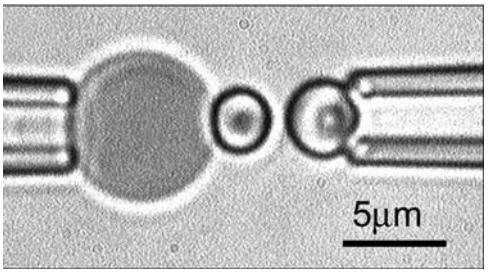


Figure 21.5. Photomicrograph of a biomembrance force probe. A micropipette-aspirated red blood cell with a bead (probe) glued to its apex (*left*) was aligned against another bead (target) aspirated by another pipette (*right*). The right pipette was driven by a computer-programmed piezoelectric translator to move in a repcated approach-push-retract-hold-return test cycle. The left pipette was held stationary, but the position of the probe was tracked by a high speed camera and analyzed by advanced tracking software. [Reproduced from Chen et al. (2008) with permission.]

it to a desired position with 1nm precision, holds it there for a given duration, and returns it to the original position. This cycle is repeated many times to acquire an ensemble of data for statistical analysis as in the adhesion frequency assay. The position of the probe bead is continuously recorded and analyzed for detection of events of formation and dissociation of receptor–ligand bonds (Chen et al. 2008).

Two such test cycles are exemplified Figure 21.6A and B for PSGL-1 interacting with L-selectin and P-selectin, respectively. The BFP probe was initially pushed 150 nm by the target, returned to ∼0 nm after the target was retracted, and held to allow contact with the probe via thermal fluctuations but not by compression, where its mean position may be pulled toward the target by bond formation, and remained at ∼0 nm (Figure 21.6A), when the target was retracted again (when no bond was present at that time) or was pulled ∼90 nm away by the retracting target (when a bond or bonds were present at that time) and then sprung back to ∼0 nm after adhesion between the probe and target was ruptured (Figure 21.6B). Thermal fluctuations were measured by a sliding standard deviation of 15 consecutive points from the time course data (Figure 21.6C and D). For a BFP spring constant of $k_p = 0.15$ pN/nm, an expected standard deviation of ∼5.2 nm is predicted from the equipartition theorem for a free probe at room temperature, Eq. (21.7). Pressing the target against the probe suppressed the thermal fluctuations. Nonrandom long-distance travels driven by the piezoelectric translator resulted in artificially large sliding standard deviations when the probe was pushed or pulled by the target. After discarding these, three intervals (arrows) were still clearly seen in Figure 21.6A where thermal fluctuations lower than 3 nm (horizontal solid line in Figure 21.6C and D, which is a set threshold of one standard deviation lower than that expected for a free probe when the probe was neither pushed nor pulled), suggesting the presence of a bond or bonds in these intervals. Indeed, the mean deflections in these periods were shifted upward due to bond formation across a mean gap distance that was slightly larger than the length of the L-selectin/PSGL-1 complex. The threshold method also allowed for determination of two periods in Figure 21.6*D* in which a bond or bonds were present despite the fact that no upward shift of the mean deflection was observed in Figure 21.6*B*, presumably because the mean gap distance between the two surfaces was no larger than the length of the L-selectin/PSGL-1 complex. The suggested absence or presence of a bond or bonds immediately prior to the target retraction was confirmed by the observation that the probe was not (Figure 21.6*A*) or was (Figure 21.6*B*) pulled by the retracting target.

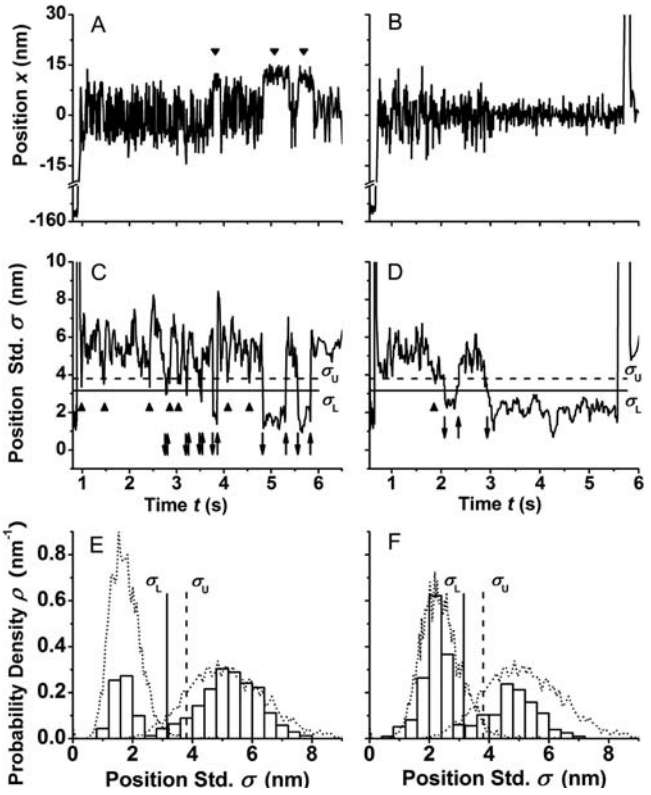


Figure 21.6. Thermal fluctuation method. (A, B) Horizontal position x of the right edge of the probe is plotted vs. time t for a representative test cycle measuring interactions of PSGL-1 coated on the probe with L- (A) or P-selectin (B) coated on the target. Two periods of high positions in panel A are indicated by arrowheads. (C, D) Sliding standard deviations σ of 15 consecutive points of the position data in panels A and B, respectively. (E, F) Histograms of the σ data in panels C and D (*bars*), respectively. Also superimposed on each panel are two histograms of σ values calculated from $x(t)$ data of two unencumbered probes recorded for the same duration (*dotted curves*). One unencumbered probe had the same spring constant of $k_p = 0.15$ pN/nm as the probe used to acquire the data in panels C and D. The other unencumbered probe had spring constant of $k_p = 1.7$ (***E***) or 0.8 (***F***) pN/nm. All histograms were normalized to have a unity area. The vertical dashed line $\sigma_U = 3.8$ nm on each panel is one standard deviation (1.3 nm) to the left from the peak at 5.1 nm. The vertical solid line $\sigma_L = 3.15$ nm on each panel is 1.5 standard deviation to the left from the same peak. These thresholds are marked in panels C and D as horizontal lines to identify bond association/dissociation events, which are marked by the respective down and up arrows. Arrowheads indicate intervals deemed indeterminate for whether they corresponded to free or bound probes because data lay between the two thresholds. [Reproduced from Chen et al. (2008) with permission.]

Figure 21.6 is similar to Figure 21.4; however, the much softer spring constant of the BFP compared to the AFM resulted in much clearer changes in the deflection fluctuations when a bond is formed or dissociates (Figure 21.6A and B vs. Figure 21.4A and B), which manifests as a sudden decrease or increase in the standard deviations (Figure 21.6C and D vs. Figure 21.4C and D), enabling us to pinpoint the time when bond formation or dissociation occurs even when the force probe is placed so close to the target that bond dissociation does not result in a significant shift in the mean displacement (Figure 21.6B vs. Figure 21.6A and Figure 21.4A and B). The two peaks in each of the standard deviation histograms corresponding to the free and bound force probe are clearly separated with

minimal overlaps (Figure 21.6E and F vs. Figure 21.4E and F). Similar to the observation from Figure 21.4E and F, the displacement standard deviations that resulted from bond formation are shifted further left for L-selection than for P-selectin. This is consistent qualitatively with the fact that P-selectin has a smaller molecular spring constant than L-selectin and the amounts of left shift agree quantitatively with the spring constants of the BFP, L-selectin, and P-selectin according to Eqs. (21.7) and (21.8). This suggests that the thermal fluctuation assay can distinguish the type of bonds in addition to identifying the bond formation and dissociation events. In addition, these data suggest that the number of bonds present is no more than one because multiple bonds result in multiple springs in parallel, which should further decrease the thermal fluctuations, which was not observed.

A series of tests were performed by Chen et al. (2008) to validate the thermal fluctuation assay. The correlation (or lack thereof) between the two methods of determining the presence of a bond—by the target pulling during its retraction and by the reduced thermal fluctuations immediately prior to the retraction—can be used to test the reliability of the thermal fluctuation method to determine the presence of a bond at a particular moment. In 725 tests similar to those shown in Figure 21.6, it was found that ∼10% were too close to call because their thermal fluctuations immediately prior to the target retraction were between the upper and lower bounds chosen to separate a free probe from a probe linked to the target by a bond (e.g., at the level similar to those indicated by the arrowhead in Figure 21.6A). In the remaining 90% of the tests, >96% of the times the thermal fluctuation method reported correctly either having or not having a bond as confirmed by the pulling method (Figure 21.7,two diagonal bars). False-positive events (scored as having a bond by the thermal fluctuation method but not confirmed by the pulling method) and false-negative events (scored as not having a bond by the thermal fluctuation method but shown to have a bond by the pulling method) were <3.5% (Figure 21.7, two off-diagonal bars). These results support the reliability of determining bond formation and dissociation events based on the reduction and resumption of thermal fluctuations.

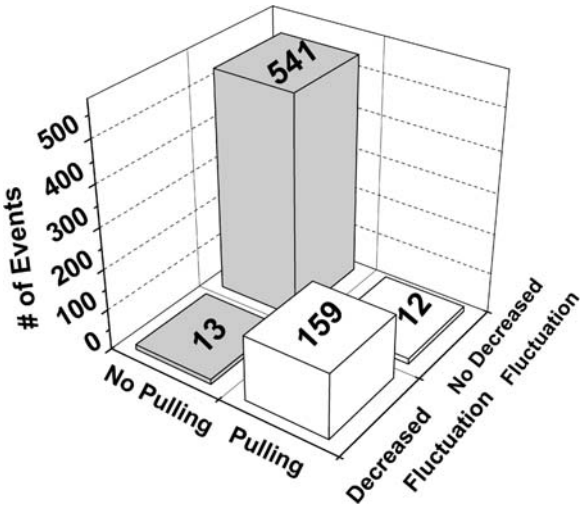


Figure 21.7. Correlation between two methods for determining the presence of a bond. [Reproduced from Chen et al. (2008) with permission.]

The ability to pinpoint when bond formation and dissociation take place enables measurement of 2D receptor–ligand binding kinetics. The period from the instant of dissociation of an existing bond to the instant of formation of the next bond is the waiting time t_w, and the period from the instant of bond formation to the instant of bond dissociation is the bond lifetime t_1. Their statistics contain information regarding on- and off-rates because the faster the on-rate (or off-rate), the shorter is the waiting time (or lifetime). In addition, the higher the densities of receptors and ligands on the respective surfaces of the force probe and the target, the shorter is the waiting time. Assuming first-order kinetics of irreversible association of single bonds governed by

$$\frac{dP_a}{dt_w} = m_t m_l A_c k_{+1}(1 - P_a) \quad \text{and} \quad P_a|_{t_w=0} = 0 \tag{21.9a,b}$$

we can analyze a pooled collection of waiting times by the solution

$$P_a = 1 - \exp(-m_t m_l A_c k_{+1} t_w) \tag{21.9c}$$

Here $1 - P_a$ is the probability for no bond to form during the interval $[0, t]$. Taking the natural log of $(1 - P_a)$ linearizes the exponential waiting-time distribution given by Eq. (21.9c). The negative slope of the ln(# of events with a waiting time $\geq t$) vs. t_w plot provides an estimate for $m_t m_l A_c k_{+1}$ (Figure 21.8A). Similarly, assuming first-order kinetics of irreversible dissociation of single bonds governed by

$$\frac{dP_b}{dt_1} = k_{-1} P_b \quad \text{and} \quad P_b|_{t_1=0} = 1 \tag{21.10a,b}$$

we can analyze a pooled collection of bond lifetimes by the solution

$$P_b = \exp(-k_{-1} t_1) \tag{21.10c}$$

Here P_b is the probability for a bond formed at time 0 to remain bound at time t. Taking the natural log of P_b linearizes the exponential bond lifetime distribution given by Eq. (21.10c).

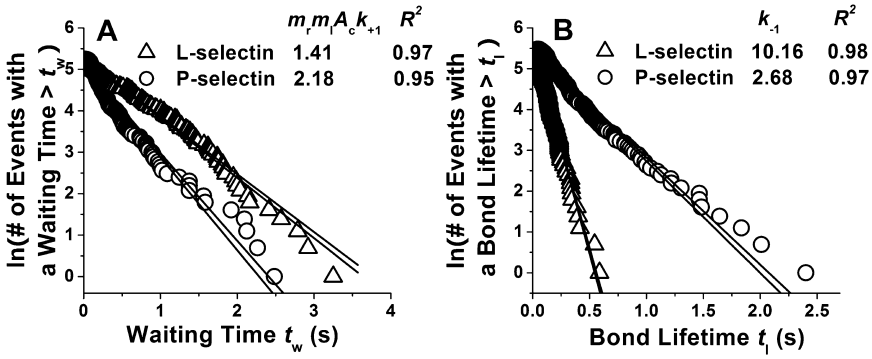


Figure 21.8. (A) Exponential distributions of pooled bond waiting times and (B) bond lifetimes of PSGL-1 interacting with L- and P-selectin, respectively. [Reproduced from Chen et al. (2008) with permission.]

Thus, the negative slope of the ln(# of events with a lifetime $\geq t$) vs. t_1 plot provides an estimate for k_{-1} (Figure 21.8B). Note that Eqs. (21.9c) and (21.10c) each contains only a single fitting parameter, which makes their respective curve fitting of the waiting-time distribution (Figure 21.8A) and lifetime distribution (Figure 21.8B) more robust and the best-fit parameters more reliable.

The linear appearance of the data seen in Figure 21.8 corroborates the observation in Figure 21.6 and supports the hypothesis that formation and dissociation of single bonds were measured in these experiments. If the negative slopes of the linear fits to the data in Figure 21.8A and B indeed represent respective cellular on-rates and off-rates, then the former should increase linearly with, and the latter should be independent of, the site densities of the receptors and ligands, provided that observed interactions reflect predominately single bonds. Chen et al. (2008) tested this prediction by measuring L-selectin–PSGL-1 interaction kinetics using four different site densities. The cellular association rate constant, $m_t m_l A_c k_{+1}$, was found to be proportional to the site densities of L-selectin and PSGL-1 (Figure 21.9A), as expected from second-order forward reaction (Eqs. 21.3 and 21.9a). The slope of the linear fit in Figure 21.10A gives the average on-rate, $A_c k_{on}$, of 5.9 $\times$ 10^{-5} $\mu m^4 s^{-1}$. The off-rate constant was found to be independent of the site densities (Figure 21.9B), as expected from Eq. (21.10a). The mean off-rate, $<k_{-1}>$, is 10.2 s^{-1}.

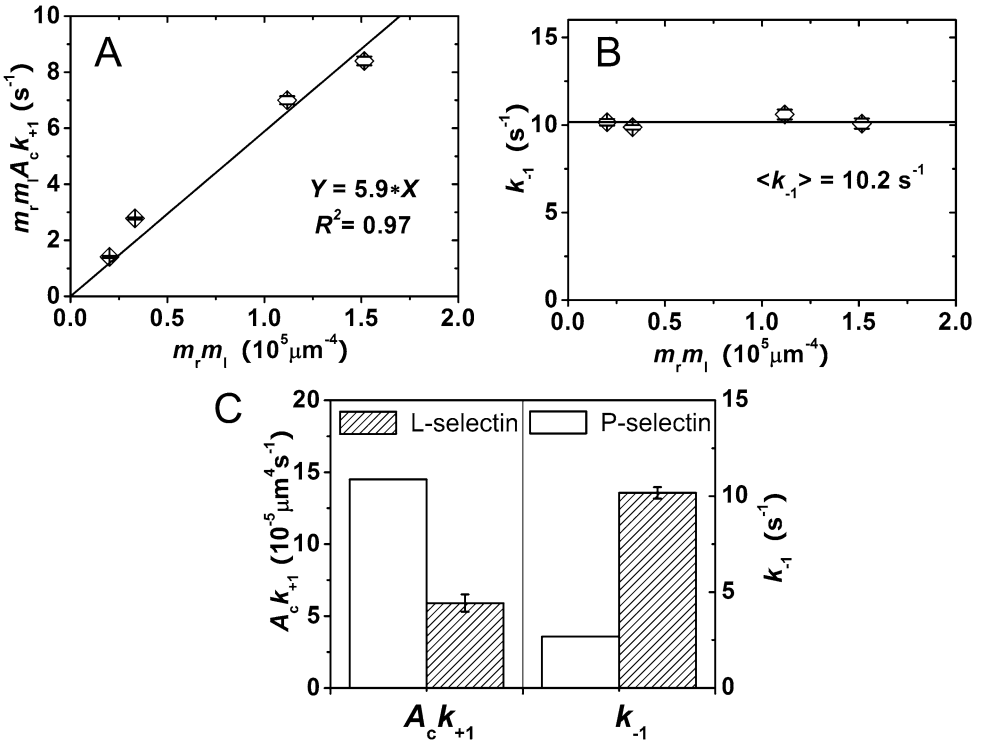


Figure 21.9. Kinetic parameters of PSGL-1 interacting with L- and P-selectins, respectively. [Reproduced from Chen et al. (2008) with permission.]

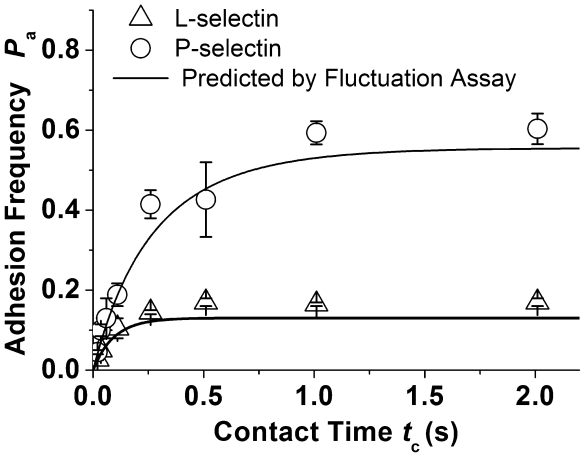


Figure 21.10. Comparison of the thermal fluctuation assay and the adhesion frequency assay. [Reproduced from Chen et al. (2008) with permission.]

The validity of the thermal fluctuation assay can be tested by the sensitivity of the kinetic parameters estimated to the molecular interactions assayed, e.g., L-selectin vs. PSGL-1 and P-selectin vs. PSGL-1. Histograms of waiting times and lifetimes of these two interactions at comparable site densities are compared in Figure 21.8A and B. It clearly shows that P-selectin has a faster on-rate but a slower off-rate to PSGL-1 than does L-selectin. The kinetic rates are compared in Figure 21.9C, and are consistent with previous finding that P-selectin has much higher affinity and slower off-rate to PSGL-1 than does L-selectin (Nicholson et al. 1998).

The thermal fluctuation method can be further validated by comparing the 2D kinetic rates measured by this method with those measured by the adhesion frequency assay (Chesla et al. 1998), which has been extensively used to determine many receptor–ligand interactions. As described in the previous section, rather than measuring rupture forces, the adhesion frequency assay estimates the likelihood of adhesion, or adhesion probability, P_a, from the frequency of adhesion enumerated from a large number of repeated controlled contacts. P_a is related to the contact time t_c through a probabilistic model described by Eq. (21.6) in Section 21.2 (Chesla et al. 1998).

Using the same BFP, same reagents, same site densities, and experiments prepared in the same way as those in the thermal fluctuation method, Chen et al. (2008) measured adhesion frequencies in a range of contact durations for both L-selectin and P-selectin interacting with PSGL-1. The P_a vs. t_c data (points) are shown in Figure 21.10. Frequencies of adhesion mediated by PSGL-1 interacting with L-selectin (Δ) or P-selectin ($\circ$) were measured at indicated contact times (points, mean $\pm$ s.e.m. of three probe-target pairs) from 100 test cycles per probe–target pair. Theoretical adhesion frequencies as functions of contact time were predicted (curves) by Eq. 21.6 using the kinetic rates from Figure 21.9C and molecular densities measured from independent experiments ($m_t m_l = 0.2$ and 0.15×10^5 μm^{-4} for the L- and P-selectin cases, respectively). It is evident that the predictions agree with the data reasonably well for both the L-selectin and P-selectin cases (Figure 21.10), further supporting the validity of the thermal fluctuation method.

We have used kinetic measurements to illustrate the thermal fluctuation assay. Whereas these measurements can also be performed using the adhesion frequency assay (Section 21.2), it has to acquire the information on whether a bond is present at a given time by terminating the contact at that time. By comparison, the thermal fluctuation assay can additionally monitor continuously when a bond forms and dissociates without any experimental intervention This greatly enhances the quantity, quality, and reliability of the information obtained, which makes kinetic measurements much simpler and more robust.

21.4. Analysis of Correlation among Outcomes from Sequential Tests

The adhesion frequency assay and thermal fluctuation assay discussed in the previous sections are similar to other single-molecule biomechanical measurements discussed in other chapters of this book, in that they are inherently stochastic. This is because molecular events (e.g., unfolding of a globular protein domain or unbinding of a receptor–ligand bond) are determined not only by the weak noncovalent interactions (within a single molecule or between two interacting molecules), but also by thermal excitations from the environment. For example, in a particular adhesion test both positive and negative outcomes are possible. When adhesion does occur, its detachment force or lifetime can be any positive value. The stochastic nature of the data demands a large number of measurements for statistical analysis. For example, to estimate the adhesion probability requires averaging a large number of adhesion scores (Chesla et al. 1998). To estimate the probability distribution of single-bond unbinding forces or lifetimes requires histogram analysis of a large number of measurements. These are most often obtained by sequentially repeating the measurement many times using the same force probe and target to yield a sequence of random numbers $\{X_i, i = 0, 1,\ldots\}$, e.g., random sequence of 0's ($X_i = 0$ for no adhesion) and 1's ($Xi = 1$ for an adhesion) in the case of the adhesion frequency assay. Without high-throughput techniques to enable simultaneous parallel testing, the sequential testing protocol is much more efficient than the consecutive parallel testing protocol that uses a large number of force probe–target surface pairs but tests each pair just once, e.g., for adhesion/no adhesion, which would be very time-consuming.

A crucial assumption that allows probability to be calculated from an ensemble of measurements is that all measurements are identical yet independent from each other, i.e., they are i.i.d. random variables. It seems reasonable to accept the i.i.d. assumption for measurements collected from parallel tests. It is also a common practice to accept the same assumption for measurements collected from sequential tests. But is this assumption always valid? Can it be tested? How can the degree of deviation from the i.i.d. assumption or the level of correlation among different measurements be measured? What may be the reasons for the i.i.d. assumption to be violated? Can a violation of the i.i.d. assumption reveal intrinsic properties of the interacting cells or molecules under study? In this section we describe a recently published analysis that models the sequential tests as a one-step Markov process to extract information from the correlation of the binding events in the sequence, which has addressed some of these questions (Zarnitsyna et al. 2007).

Zarnitsyna et al. (2007) analyzed data from the adhesion frequency assay by developing a model for size distribution of consecutive adhesion events expected for a one-step Markov process. Fitting the model to experimental data allows one to quantify the magnitude and direction of deviation from the i.i.d. assumption in terms of a "memory" index. Here, memory represents the ability of the system to remember the result of the previous test as evidenced

by a change in the likelihood of outcome in the subsequent test. It was found that nature has provided examples for all three theoretically possible scenarios: no memory, positive memory, and negative memory.

To illustrate this analysis, let us consider the micropipette adhesion frequency data for three receptor–ligand systems shown in Figure 21.11: lymphocyte function–associated antigen 1 (LFA-1) interacting with intercellular adhesion molecule 1 (ICAM-1) (Figure 21.11A), TCR/pMHC interaction (Figure 21.11B), and homotypical C-cadherin interaction (Figure 21.11C). The result of n (= 50) repeated tests is a random sequence whose value X_i at the ith position is either 0 or 1. In previous analysis (Chesla et al. 1998), the running adhesion frequency was defined as $F_\mathrm{i} = (X_1 + X_2 + \ldots + X_\mathrm{i})/i$ $(1 \leq i \leq n)$. Three running adhesion frequencies F_i vs. test cycle index i plots are shown in Figure 21.11A–C. F_i fluctuates when i is small because of small–number statistics, but it should approach an asymptotic curve as i approaches n unless the sequence is too short. For sequences of sufficient length, F_n is the average adhesion frequency, which is the best estimator for the probability of adhesion in each test if the i.i.d. assumption holds, i.e., the obtained sequence could be described by a Bernoulli process. For the three molecular systems studied, F_i approaches a plateau equal to the averaged adhesion probability P_a.

Another way to visualize the sequences in Figure 21.11 A–C is to plot the nonzero $X_\mathrm{i}F_\mathrm{n}$ vs. i (Figure 21.11D–F). The zero scores for no-adhesion events have been omitted for clarity. Three molecular systems are shown: LFA-1/ICAM-1 interaction (Figure 21.1A,D), TCR/pMHC interaction (Figure 21.1B,E), and homotypic interaction between C-cadherins (Figure 21.1C,D). For each interaction, three sequences were obtained using three pairs

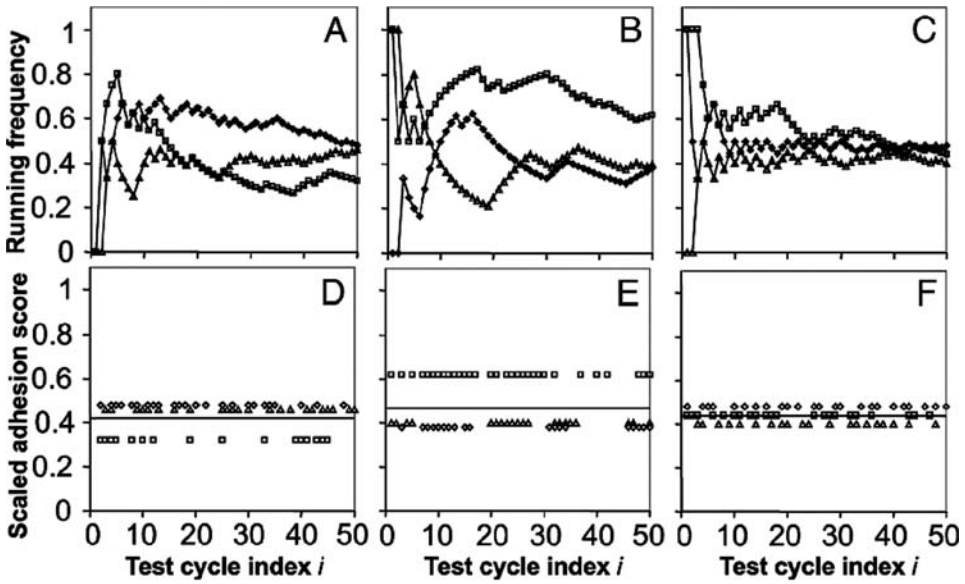


Figure 21.11. Running adhesion frequencies (A–C) and scaled adhesion score sequences (D–F) plotted vs. the test cycle index i. The measured adhesions were mediated by LFA-1/ICAM-1 interaction (A, D), TCR/pMHC interaction (B, E), or homotypic C-cadherin interaction (C, F). Matched symbols are used to show corresponding test sequences in panels A–C with those in panels D–F. Each sequence represents n (= 50) repeated tests using a single pair of cells performed under identical experimental conditions. The average of the three averaged adhesion frequencies is shown as a solid line on each of panels *E–H*. [Reproduced from Zarnitsyna et al. (2007) with permission.]

of cells tested under the same conditions. Symbols in Figure 21.11D–F match those in Figure 21.11A–C. To compare different receptor–ligand interactions, data sets were chosen to have similar mean adhesion levels (solid lines, Figure 21.11D–F).

Three distinct behaviors seem apparent even at a brief glance at the adhesion score sequences. Compared to those for the LFA-1/ICAM-1 interaction (Figure 21.11D), the sequences for the TCR/pMHC interaction (Figure 21.11E) appear more "clustered," whereas those for the C-cadherin interaction (Figure 21.11F) are less "clustered." Here, cluster refers to consecutive adhesion events without interruption by no-adhesion events. In Figure 21.11A–C, clustering manifests as uninterrupted ascending segments in the running adhesion frequency curves.

Closer inspection of the scaled adhesion scores in Figure 21.11D–F reveals that they are clustered at different sizes. It seems intuitive that, for a given "cluster size" m (i.e., m consecutive adhesions), the number of times it appears in an adhesion score sequence contains statistical information about that sequence. This intuition is supported by the data in Figure 21.12, which show the cluster size distribution enumerated from the adhesion score sequences in Figure 21.11. Compared to the distribution for the LFA-1/ICAM-1 interaction (Figure 21.12A), the distribution for the TCR/pMHC interaction (Figure 21.12B) has more clusters of large sizes, whereas the distribution for the C-cadherin interaction (Figure 21.12C) has more single-adhesion events surrounded by no-adhesion events.

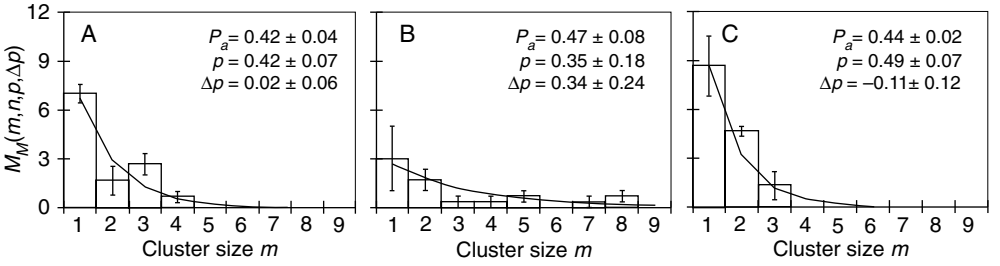


Figure 21.12. Cluster size distribution. Number of clusters of various sizes enumerated from Figure 21.11 (*bars*) for TCR/pMHC interaction (A), LFA-1/ICAM-1 interaction (B), or homotypic C-cadherin interaction (C) was fitted by Eq. 21.12 (*curves*). Both data and the best-fit parameters (legends) are presented as mean ± s.e.m. [Reproduced from Zarnitsyna et al. (2007) with permission.]

To quantify the differences among the three cases in Figure 21.11, a formula was derived to express the number M_B of clusters of size m expected in a sequence described by a Bernoulli process of length n and probability p for the positive outcome in each test:

$$M_B(m,n,p) = \begin{cases} [n-(m+1)]p^m(1-p)^2 + 2p^m(1-p) & m < n,\ m \geq 1 \\ p^n & m = n \end{cases} \tag{21.11}$$

Here p is the probability to have adhesion in the next test (i.e., $X_{i+1} = 1$) under the condition that the present adhesion test is not successful (i.e., $X_i = 0$). The first term in the upper branch on the right-hand side of Eq. (21.11) represents summation over the probabilities of having a cluster of size m in all possible positions, i.e., for clusters starting at $i = 2$ to $i = n - (m - 1)$. Clusters starting from X_1 or ending with X_n are accounted for by the second term in the upper branch. The lower branch of the formula accounts for the sequence of all 1's. It becomes apparent from the derivation that Eq. (21.11) assumes equal probability for the cluster to take any position in the sequence. This can be thought of as a stationary assumption.

The total number of positive adhesion scores in the entire sequence can be calculated by multiplying Eq. (21.11) by m and then summing over m from 1 to n. It can be shown by direct calculation that $\sum mM_{\mathrm{B}}(m,n,p) = np$. Here np is the expected total number of adhesion events. This is predicted and shows that Eq. (21.11) is self-consistent.

Equation (21.11) can be extended to the case of Markov sequence by including a single-step memory:

$$M_{\mathrm{M}}(m,n,p\Delta p) = \begin{cases} p(p+\Delta p)^{m-1}\left[1 - p\frac{1-\Delta p^{n-m}}{1-\Delta p} + (1-p-\Delta p)\left(n - m - \frac{p}{1-\Delta p}(n-(m+1) - \Delta p\frac{1-\Delta p^{n-(m+1)}}{1-\Delta p})\right)\right] & m < n \\ p(p+\Delta p)^{n-1} & m = n \end{cases} \tag{21.12}$$

where M_{M} is the number of clusters of size m expected in a Markov sequence of n repeated tests. Here p and $p + \Delta p$ are the respective probabilities for binding in the next test if the present test results no binding and binding, respectively. Thus, Δp represents an increment (can be positive or negative) of binding probability in the future test due to the occurrence of binding in the present test and can be thought of as a memory index.

Setting $\Delta p = 0$ reduces Eq. (21.12) to (21.11), as required. Equation (21.12) has another special case at $\Delta p = 1 - p$, when it simplifies to $M_{\mathrm{M}}(m,n,p,1–p) = p(1 - p)^{\mathrm{n–m}}$, which describes the situation where the first adhesion event in the sequence would increase the probability for adhesion in the next test to 1, leading to continuous adhesion for all tests until the end of the experiment.

Equation (21.12) was fit (curves) to the cluster size distributions (bars) in Figure 21.12, which were measured from the data in Figure 21.11. A positive memory index $\Delta p = 0.34$ was returned from fitting the data of the TCR/pMHC interaction (Figure 21.12B), indicating memory with positive correlation. Note that the binding probability increment $\Delta p = 0.34$ is almost equal to p ($= 0.35$), suggesting very strong impact of the past on the future. By comparison, fitting the data of the LFA-1/ICAM-1 interaction (Figure 21.12A) returned a 17-fold smaller Δp that is not statistically different from zero, indicating the lack of memory and correlation in this case. Fitting the C-cadherin data (Figure 21.12C) returned a negative Δp (= –0.11), indicating memory with negative correlation.

The results shown in Figure 21.12A–C are further supported by tenfold more data obtained by varying the experimental conditions for each receptor–ligand interaction. Different contact times and/or receptor–ligand densities resulted in different average adhesion frequencies P_{a}. The memory index Δp was obtained by fitting Eq. (21.12) to an experimental cluster size distribution and is plotted in Fig. 21.13 (solid bars) along with Δp estimated from direct calculation described by Eq. (21.13) (open bars),

$$\Delta p = p_{11} - p_{01} = n_{11}/(n_{11} + n_{10}) - n_{01}/(n_{01} + n_{00}) \tag{21.13}$$

where n_{ij} is the number of $i \rightarrow j$ transitions directly calculated from the adhesion score sequence and p_{01} is the probability to have adhesion (i.e., $X_{\mathrm{i+1}} = 1$) under the condition that the previous adhesion test was not successful (i.e., $X_{\mathrm{i}} = 0$). It is also given the short notation p by definition. p_{11} is the probability to have adhesion in the next test if the previous test also results in adhesion.

Comparable results were obtained by both approaches for all P_{a} values tested for all three systems exhibiting qualitatively distinct behaviors. Δp values for the LFA-1/ICAM-1 system are not statistically significantly ($P \geq 0.23$) different from zero except one ($P = 0.06$).

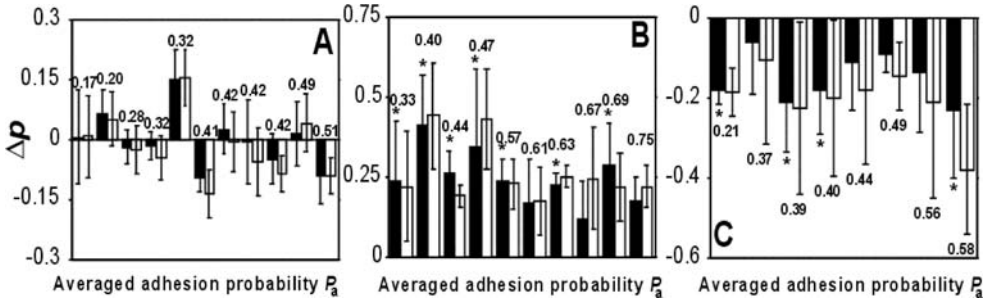


Figure 21.13. Memory in adhesion mediated by three interactions. The memory index Δp is determined by fitting the measured number of clusters with Eq. 21.12 (*solid bars*) or by direct calculation using Eq. 21.13 (*open bars*) for a range of averaged adhesion probabilities (indicated) for adhesion mediated by LFA-1/ICAM-1 (A), TCR/pMHC (B), or C-cadherin (C) interactions. Data are shown as mean ± s.e.m. Statistical significant ($P \leq 0.05$) differences of Δp values from zero using the *one-sample z-test* are marked with asterisks (*). [Reproduced from Zarnitsyna et al. (2007) with permission.]

By comparison, a vast majority of the Δp values for the TCR/pMHC system are statistically significantly greater than zero, whereas half of the Δp values for the C-cadherin system are statistically significantly less than zero, which are marked with asterisks (*) on the top of corresponding solid bars to indicate $P \leq 0.05$.

The foregoing analysis shows that the validity of the i.i.d. assumption can be tested using the cluster size distribution model, Eq. (21.12). This test reveals that the i.i.d. assumption is valid for some but not all consecutive single-molecule biomechanical measurement sequences obtained by the same adhesion frequency assay from the same laboratory using the same equipment but for different molecular interactions. The degree of deviation from the i.i.d. assumption, or the level of correlation among sequential measurements, can be quantified using a memory index Δp. It is called a memory index because correlation among different tests in a time sequence represents the impact of the past on the future. This includes at least three aspects: (1) the magnitude (i.e., to what extent the past memory impacts the future), (2) the direction (i.e., whether the impact is positive or negative), and (3) the duration (i.e., how long the memory lasts). The duration can be quantified by varying the time between two consecutive tests, which was 0.5 sec in the experiments analyzed here. It seems reasonable to suspect that the memory may fade away if this time is prolonged. Another question is how long ago a previous test may still have an impact. Here the simplest scenario is considered in which only the immediate past test is assumed to influence the next test. Relaxing this assumption can include more general scenarios to allow influences from tests further upstream, which requires multistep memories.

There can be multiple reasons for the i.i.d. assumption to be violated, so measurement of the memory index can be used to probe intrinsic properties of the interacting cells or molecules. At the level of molecular interactions, adhesion memory likely reflects kinetic processes triggered by the measured binding events. The hypothesis that memory may reflect intrinsic properties of the interacting cells or molecules is supported by the data that fixing the T cells reduced Δp to zero. The mathematical model for the adhesion frequency assay predicts that the average number of bonds is $\sim$1 when P_a has mid-range values (Chesla et al. 1998). The data in Figure 21.13 thus suggest that a single TCR/pMHC bond engaged less than 1 sec is sufficient to generate significant memory. It has been shown that calcium responses

(Li et al. 2004) or cytotoxic activities (Sykulev et al. 1996) in T cells can be triggered by engagement of TCR with even a single pMHC. An intriguing question is whether these two forms of exquisite sensitivity are related or are two manifestations of the same mechanism.

Like the TCR/pMHC interaction, the homotypic interaction between C-cadherins mediates adhesion as well as signaling. Contrary to the TCR/pMHC interaction, engagement of C-cadherin in the previous test downregulates the probability of adhesion in the next test, which is also intriguing. The damping of receptor responsiveness could reflect a physiological feedback mechanism that protects against both acute and chronic receptor overstimulation (Ferguson 2001).

It is reasonable to suggest that adhesion "memory" likely reflects transient structural alterations triggered by binding that affect not only the probability of the next adhesion, but also other characteristics of the receptor–ligand bond, such as the waiting time between the dissociation of the previous bond and the formation of the next bond and the lifetime and unbinding force of the next bond. These predictions can be tested using the thermal fluctuation assay discussion in Section 21.3. Of note, for interactions in which nonzero memory indices are found from analysis of the sequences of binary adhesion scores, measurements such as unfolding or unbinding forces, time to unfold a globular protein domain, or time to unbind a receptor–ligand bond, as discussed in other chapters in this book, should be segregated into subgroups according to the outcomes of the previous tests for separate analysis.

Acknowledgments

We thank our coworkers in the work reported in (Huang et al. (2007), Chen et al. (2008), and Zarnitsyna et al. (2007), who contributed the original data discussed here. This work was supported by NIH grants A103828 and A1044902.

References

Chen, W., E. A. Evans, R. P. McEver, and C. Zhu (2008). "Monitoring receptor-lignad interactions between surfaces by thermal fluctuations." *Biophys J* 94(2): 694–701.

Chesla, S. E., P. Selvarai, and C. Zhu (1998). "Measuring two-dimensional receptor-ligand binding kinetics by micropipette." *Biophys J* 75(3): 1553–72.

Evans, E., D. Berk, and A. Leung (1991). "Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments." *Biophys J* 59(4): 838–48.

Ferguson, S. S. (2001). "Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling." *Pharmacol Rev* 53(1): 1–24.

Huang, J., L. J. Edwards, B. D. Evavold, and C. Zhu (2007). "Kinetics of MHC-CD8 interaction at the T cell membrane." *J Immunol* 179(11): 7653–62.

Hutter, J. L., and J. Bechhoefer (1993). "Calibration of atomic-force microscope tips." *Rev. Sci Instrum* 64: 1868–1873.

Li, Q. J., A. R. Dinner, S. Qi, D. J. Irvine, J. B. Huppa, M. M. Davis, and A. K. Chakraborty (2004). "CD4 enhances T cell sensitivity to antigen by coordinating Lck accumulation at the immunological synapse." *Nat Immunol* 5(8): 791–9.

Marshall, B. T., M. Long, J. W. Piper, T. Yago, R. P. McEver, and C. Zhu (2003). "Direct observation of catch bonds involving cell-adhesion molecules." *Nature* 423(6936): 190–3.

Marshall, B. T., K. K. Sarangapani, J. Wu, M. B. Lawrence, R. P. McEver, and C. Zhu (2006). "Measuring molecular elasticity by atomic force microscope cantilever fluctuations." *Biophys J* 90(2): 681–92.

Nicholson, M. W., A. N. Barclay, M. S. Singer, S. D. Rosen, and P. A. van der Merwe (1998). "Affinity and kinetic

analysis of L-selectin (CD62L) binding to glycosylation-dependent cell-adhesion molecule-1." *J Biol Chem* 273(2): 763–70.

Piper, J. W., R. A. Swerlick, and C. Zhu (1998). "Determining force dependence of two-dimensional receptor-ligand binding affinity by centrifugation." *Biophys J* 74(1): 492–513.

Sarangapani, K. K., T. Yago, A. G. Klopocki, M. B. Lawrence, C. B. Fieger, S. D. Rosen, R. P. McEver, and C. Zhu (2004). "Low force decelerates L-selectin dissociation from P-selectin glycoprotein ligand-1 and endoglycan." *J Biol Chem* 279(3): 2291–8.

Sykulev, Y., M. Joo, I. Vturina, T. J. Tsomides, and H. N. Eisen (1996). "Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response." *Immunity* 4(6): 565–71.

Thoumine, O., P. Kocian, A. Kottelat, and J. J. Meister (2000). "Short-term binding of fibroblasts to fibronection: optical tweezers experiments and probabilistic analysis." *Eur Biophys J* 29(6): 398–408.

Wu, J., Y. Fang, D. Yang, and C. Zhu (2005). "Thermo-mechanical responses of a surface-coupled AFM cantilever." *J Biomech Eng* 127(7): 1208–15.

Zarnitsyna, V. I., J. Huang, F. Zhang, Y. H. Chien, D. Leckband, and C. Zhu (2007). "Memory in receptor-ligand-mediated cell adhesion." *Proc Natl Acad Sci USA* 104(46): 18037–42.

Zhu, C., M. Long, S. E. Chesla, and P. Bongrand (2002). "Measuring receptor/ligand interaction at the single-bond level: experimental and interpretative issues." *Ann Biomed Eng* 30(3): 305–14.

Subject Index