N

N-Acetylmuramide Glycanhydrolase

► Lysozyme – Computational Studies

NADH Dehydrogenase

► NADH-Ubiquinone Oxidoreductase (Complex I)

NADH-Ubiquinone Oxidoreductase (Complex I)

Ulrich Brandt and Volker Zickermann
Molecular Bioenergetics Group, Medical School, Cluster of Excellence Frankfurt "Macromolecular Complexes," Center for Membrane Proteomics, Goethe-University, Frankfurt am Main, Germany

Synonyms

NADH dehydrogenase; NDH-1; Respiratory complex I

Definition

Proton-pumping NADH-ubiquinone oxidoreductase (complex I) is a very large membrane protein complex that is found in many bacteria and almost all eukaryotes (Brandt 2006). In the mitochondrial respiratory chain it couples the transfer of two electrons from NADH to ubiquinone to the vectorial translocation of four protons across the inner membrane of the organelle and thereby generates up to 40% of the driving force for ATP synthesis by ATP synthase. Complex I is a source of reactive oxygen species and is involved in the pathogenesis of encephalomyopathies and neurodegenerative diseases.

Basic Characteristics

Structure and Functional Modules

The minimal form of complex I represented by the enzyme from prokaryotes typically comprises 14 *central* subunits that are conserved throughout species. In eukaryotic complex I up to 31 additional *accessory* subunits have been identified (Carroll et al. 2006). The central subunits can be assigned to three functional modules (Table 1, Fig. 1) that harbor the bioenergetic core functions of reduced ► nicotinamide adenine dinucleotide (NADH) oxidation (N-module), quinone reduction (Q-module), and proton pumping (P-module) (Brandt 2006). The N-module and the Q-module contain all redox prosthetic groups, one ► flavin mononucleotide (FMN) molecule, and eight canonical iron-sulfur clusters. In most eukaryotes the hydrophobic central subunits ND1-ND6 and ND4L of the P-module are encoded by mitochondrial DNA.

A common L-shaped overall structure with a hydrophilic peripheral arm and a hydrophobic membrane arm has been determined for prokaryotic and eukaryotic complex I by electron microscopy. X-ray crystallographic analysis provided low resolution structural models for the complete enzyme complexes of the eubacterium *Thermus thermophilus* and the aerobic yeast *Yarrowia lipolytica* and for the major part of

G.C.K. Roberts (ed.), *Encyclopedia of Biophysics*, DOI 10.1007/978-3-642-16712-6,

NADH-Ubiquinone Oxidoreductase (Complex I), Table 1 Complex I comprising 14 central subunits arranged in three functional modules

Subunit symbol						
E. coli	*T. thermophilus*	*Y. lipolytica*	Bovine	Human	Redox centers[a]	Module
NuoG	Nqo3	NUAM	75-kDa	NDUFS1	N1b, 2 × $[Fe_4S_4]$[b]	N
NuoF	Nqo1	NUBM	51-kDa	NDUFV1	FMN, N3	N[c]
NuoE	Nqo2	NUHM	24-kDa	NDUFV2	N1a	N
NuoD[d]	Nqo4	NUCM	49-kDa	NDUFS2		Q[e]
NuoC[d]	Nqo5	NUGM	30-kDa	NDUFS3		Q
NuoI	Nqo9	NUIM	TYKY	NDUFS8	2 × $[Fe_4S_4]$	Q
NuoB	Nqo6	NUKM	PSST	NDUFS7	N2	Q[e]
NuoH	Nqo8	ND1	ND1	ND1		P_P
NuoN	Nqo14	ND2	ND2	ND2		P_P
NuoA	Nqo7	ND3	ND3	ND3		P_P
NuoK	Nqo11	ND4L	ND4L	ND4L		P_P
NuoJ	Nqo10	ND6	ND6	ND6		P_P
NuoM	Nqo13	ND4	ND4	ND4		P_D
NuoL	Nqo12	ND5	ND5	ND5		P_D

[a]The nomenclature and assignment of several iron-sulfur clusters is discussed controversially
[b]Some bacterial complexes contain an additional iron-sulfur cluster (N7) in this subunit
[c]NADH oxidation site
[d]Subunits NuoC and NuoD are fused to a single polypeptide in *E. coli*
[e]Ubiquinone binding and reduction

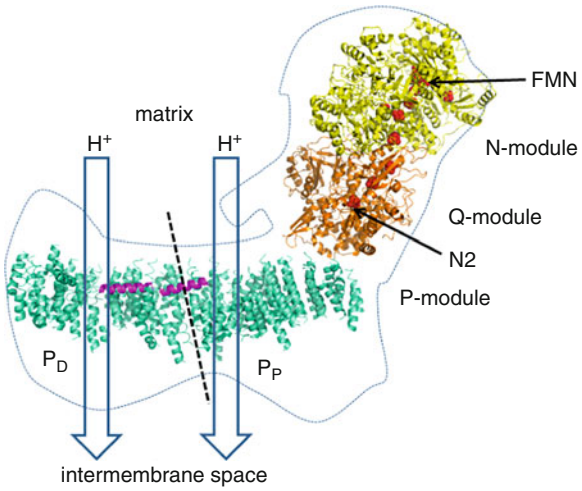
the membrane arm of complex I from *Escherichia coli* (Efremov et al. 2010; Hunte et al. 2010). The X-ray structure of the peripheral arm of complex I from *T. thermophilus* was solved at up to 3.1 Å resolution.

The N-module resides in the distal and the Q-module in the proximal part of the peripheral arm. The primary electron acceptor FMN is connected via a linear chain of iron-sulfur centers with the ubiquinone reduction site (Tocilescu et al. 2010). The proton translocation function is associated with the membrane arm that is divided into a proximal (P_P) and a distal (P_D) module.

Coupling of Redox Chemistry and Proton Translocation

The function of complex I is to convert the redox energy of the electron transfer from NADH to ubiquinone into a chemiosmotic membrane potential. In oxidative phosphorylation this proton motive force drives ATP synthesis. The ubiquinone reductase activity is inhibited by a large number of chemically diverse compounds like Rotenone, Piericidine, and DQA (2-decyl-4-quinazolinyl amine). Note that some bacteria use menaquinone instead of ubiquinone as a substrate.

The primary acceptor for electrons donated by NADH is FMN. Electron transfer to ubiquinone is mediated by a linear chain of seven iron-sulfur clusters. One additional cluster is situated near the FMN cofactor and may serve a special function in protection against ▶ reactive oxygen species (ROS) formation at the flavin site. ▶ EPR spectroscopy resolves the signatures of two binuclear and four tetranuclear clusters while two tetranuclear clusters are EPR silent. The assignment of several EPR signatures to individual clusters in the structure of complex I is a matter of debate. N2, the terminal cluster of the electron transfer pathway, is characterized by a pH-dependent midpoint ▶ redox potential ($E_{m7} = -150$ mV). The EPR visible clusters in the electron transfer pathway upstream of N2 are nearly isopotential with E_{m7} values around −250 mV. Cluster N2 is the immediate electron donor for ubiquinone and several lines of evidence indicate that the two electron reduction of ubiquinone releases the energy that drives proton translocation by complex I. Remarkably, recent structural data located cluster N2 ~30 Å above the membrane plane while the proton-pumping sites have been allocated to the proximal and distal membrane arm domain. In conclusion the spatial separation of redox chemistry and proton



NADH-Ubiquinone Oxidoreductase (Complex I), Fig. 1 *Structural model for mitochondrial complex I from Y. lipolytica and coupling of redox chemistry and proton translocation*. The peripheral arm contains the N-module (*yellow*) and the Q-module (*orange*). A linear chain of iron-sulfur clusters (*red spheres*) connects the NADH oxidation site with the ubiquinone reduction site. Cluster N2 is the immediate electron donor for ubiquinone and is situated 30 Å above the membrane plane. The membrane arm is subdivided (*black dashed line*) into a distal and proximal P-module (P_P and P_D, *cyan*). The redox chemistry is the driving force for proton translocation but is spatially separated from the putative proton translocation sites. A conformational coupling mechanism involving long range energy transfer via the long helical transmission element shown in magenta is discussed. The approximate contour line of the electron density map tentatively assigned to a complex I monomer is indicated by a *dashed blue line*

translocation implies that long range energy transfer plays a central role in the coupling mechanism of complex I (Zickermann et al. 2009). A long helical structure laterally associated with the membrane arm connects the P_P and P_D domain in prokaryotic and eukaryotic complex I and has been suggested to mediate energy transmission in a conformational coupling mechanism (Fig. 1) (Efremov et al. 2010; Hunte et al. 2010).

Evolution and Variants of Complex I

Complex I is a modular assembly of three preexisting structural and functional entities (Friedrich and Scheide 2000). Subunits of the N-module are homologous to the NAD^+ reducing [NiFe] hydrogenase of *Ralstonia eutropha*. In the Q-module the two subunits that construct the ubiquinone binding site correspond to subunits of soluble [NiFe] hydrogenases. In the P-module four polypeptides are related to subunits of bacterial monovalent cation/proton antiporters encoded by the *mrp* operon (Mathiessen and Hägerhall 2003).

Enzyme complexes related to complex I but with different electron input modules are found in archaea, in a few bacterial species, e.g., *Helicobacter pylori*, and in cyanobacteria and chloroplasts.

Complex I in Disease

Mutations in the mitochondrial genome are known to cause a number of encephalomyopathies. Leber's Hereditary Optic Neuropathy (LHON) is a prominent example that is caused by specific point mutations mainly in the genes for subunits ND1, ND4, and ND6. The typical clinical picture of nuclear mutations in complex I genes or assembly factors is Leigh syndrome, a severe, early onset neuromuscular disorder (Janssen et al. 2006).

Poisoning with the neurotoxin 1-methyl-4-phenyl 1,2,3,4 tetrahydropyridine (MPTP) was shown to reproduce Parkinson's disease in humans. The compound is converted into 1-methyl-4-phenylpyridinium (MPP^+) by monoamine oxidase. MPP^+ is selectively taken up into dopaminergic neurons by the dopamine transporter and triggers apoptotic cell death by inhibiting complex I.

Cross-References

- ▶ Electron Transfer Through Proteins
- ▶ Electron Transport Chains in Bacteria
- ▶ EPR Spectroscopy: General Principles
- ▶ Iron–Sulfur Clusters
- ▶ Membrane Proteins: Structure and Organization
- ▶ Mitochondrial Electron Transport
- ▶ Nicotinamide Adenine Dinucleotide (NAD)
- ▶ Oxidative Stress
- ▶ Paramagnetic Metalloproteins
- ▶ Reactive Oxygen Species
- ▶ X-Ray Diffraction and Crystallography of Oligosaccharides and Polysaccharides

References

Brandt U. Energy-converting NADH: quinone oxidoreductase (complex I). Annu Rev Biochem. 2006;75:69–92.

Carroll J, Fearnley IM, Skehel M, Shannon RJ, Hirst J, Walker JE. Bovine complex I is a complex of 45 different subunits. J Biol Chem. 2006;281:32724–7.

Efremov RG, Baradaran R, Sazanov LA. The architecture of respiratory complex I. Nature. 2010;465:441–5.
Friedrich T, Scheide D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multi subunit hydrogenases. FEBS Lett. 2000;479:1–5.
Hunte C, Zickermann V, Brandt U. Functional modules and structural basis of conformational coupling in mitochondrial complex I. Science. 2010;329:448–51.
Janssen RJ, Nijtmans LG, van den Heuvel LP, Smeitink JAM. Mitochondrial complex I: structure, function and pathology. J Inherit Metab Dis. 2006;29:499–515.
Mathiessen C, Hägerhall C. The "antiporter module" of respiratory chain Complex I includes the MrpC/NuoK subunit – a revision of the modular evolution scheme. FEBS Lett. 2003;549:7–13.
Tocilescu MA, Zickermann V, Zwicker K, Brandt U. Quinone binding and reduction by respiratory complex I. Biochim Biophys Acta. 2010;1797:1883–90.
Zickermann V, Kerscher S, Zwicker K, Tocilescu MA, Radermacher M, Brandt U. Architecture of complex I and its implications for electron transfer and proton pumping. Biochim Biophys Acta. 2009;1787:574–83.

NALMS (Nanometer-Localized Multiple Single-Molecule Fluorescence Microscopy)

► Single-Molecule High-Resolution Imaging with Photobleaching (SHRImP)

Nanodroplet Confinement

Lori S. Goldner
Department of Physics, University of Massachusetts, Amherst, MA, USA

Synonyms

Single-molecule methods

Definition

A method for isolating, confining, and manipulating individual molecules for study.

Basic Characteristics

Aqueous droplets in oil, with diameters from 100 nm to 100 μm and volumes from attoliters to nanoliters, are used as nanoreactors for small quantities of reagent, down to the level of single molecules. Above a femtoliter, droplet reactors can be easily formed and manipulated in microfluidic devices and have found wide use in industry. Below a femtoliter, droplets provide a convenient environment for single-molecule sensitive chemistry and biophysics.

Single-molecule sensitive measurements of structural transformations, reaction dynamics and binding all require that a molecule under study be measured for some amount of time. This requires a method for locating or distinguishing individual molecules, and for confining them in a detection region. Depending on the measurement scheme, various methods have been devised for distinguishing or localizing biomolecules, including surface binding, surface adsorption, or confinement in a porous material. Additionally, molecular confinement in surface-tethered liposomes (Boukobza et al. 2001) offers a less perturbative and more reproducible environment for the confined molecule than either direct molecular tethering or adsorption.

Confinement of single molecules to aqueous nanodroplets in oil is a simple and sometimes convenient alternative to other methods for confining or localizing molecules (Goldner et al. 2010; Reiner et al. 2006; Tang et al. 2008). Droplets are easily identified and tracked in an optical microscope, and can be manipulated optically or physically to remain in a detection volume. Droplets offer an advantage that they can be made to coalesce on contact, without loss of hydrophilic contents, thereby providing a convenient means for fast mixing. In the absence of surfactant, droplet coalescence is intrinsically fast, and the timing of mixing will be limited by diffusion in the droplets. Diffusional mixing can occur in less than 1 ms for subfemtoliter droplets.

More generally, nanodroplets offer convenient compartmentalization and localization useful in many chemical applications. Unlike nature's own cellular compartments, droplets can be engineered to be stable under conditions not commensurate with living systems. The use of emulsions makes massively parallel high-throughput measurement of bioreactions possible; a single microliter of sample can be aliquoted into 10^6–10^9 separate bioreactors. Confinement to

a small reaction volume leads to higher reaction rates than in the bulk, both because reactant concentrations can be made arbitrarily high as droplets become arbitrarily small, and because diffusion-limited reaction rates increase as volume decreases.

The utility of single aqueous droplets in oil for the study of single molecule reaction kinetics has long been recognized. Almost 50 years ago Rotman (1961) demonstrated that the activity of single β-D galactosidase molecules confined to aqueous droplets in silicone oil could be measured by monitoring the increase in fluorescence of a fluorogenic substrate over time. While Rotman did not detect individual turnovers, his technique represents perhaps the first time that single enzyme kinetics could be measured using fluorescence. More recently, chymotrypsin kinetics have been studied using a similar technique (Lee and Brody 2005). In both these works, droplets were used to confine individual molecules and indirectly study their kinetics using fluorogenic substrates. Most recently, fluorescence and fluorescence resonance energy transfer has been detected directly from individual biomolecules confined to single nanodroplets (Reiner et al. 2006; Tang et al. 2008).

Many applications have been found in recent years for droplet compartmentalization of biological reagents, typically in picoliter to nanoliter droplets. One example includes use of single DNA templates in applications of directed evolution, where droplet confinement is a convenient method of linking genotype and phenotype. Reviews can be found in Refs. (Griffiths and Tawfik 2006; Kelly et al. 2007; Leamon et al. 2006; Song et al. 2006; Taly et al. 2007). Droplet confinement has also proven useful in the amplification of single DNA templates; a recent review of microfluidic DNA amplification can be found in Zhang and Ozdemir (2009).

In recent years, the development of droplet fluidic devices is facilitating many more applications of droplet-confined chemistry. Microfluidics have the advantage of creating very monodisperse droplets, with rates of droplet formation up to 3,000/s now typical. Furthermore, by confining reactants to a droplet, dispersion that occurs by virtue of a parabolic flow field in continuous flow devices is eliminated. Many of the applications of droplet fluidic devices are discussed in Refs. (Griffiths and Tawfik 2006; Kelly et al. 2007; Leamon et al. 2006; Song et al. 2006; Taly et al. 2007).

Cross-References

- ▶ Single-Molecule Fluorescence Resonance Energy Transfer
- ▶ Single-Molecule Methods
- ▶ Surface-Tethered Lipid Vesicles

References

Boukobza E, Sonnenfeld A, Haran G. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. J Phys Chem B. 2001;105:12165–70.

Goldner LS, Jofre AM, Tang JY. Droplet confinement and fluorescence measurement of single molecules. Methods Enzymol. 2010;472:61–88.

Griffiths AD, Tawfik DS. Miniaturising the laboratory in emulsion droplets. Trends Biotechnol. 2006;24:395–402.

Kelly BT, Baret JC, Taly V, Griffiths AD. Miniaturizing chemistry and biology in microdroplets. Chem Commun. 2007; pp 1773–88.

Leamon JH, Link DR, Egholm M, Rothberg JM. Overview: methods and applications for droplet compartmentalization of biology. Nat Methods. 2006;3:541–3.

Lee AI, Brody JP. Single-molecule enzymology of chymotrypsin using water-in-oil emulsion. Biophys J. 2005;88:4303–11.

Reiner JE, Crawford AM, Kishore RB, Goldner LS, Helmerson K, Gilson MK. Optically trapped aqueous droplets for single molecule studies. Appl Phys Lett. 2006;89:3.

Rotman B. Measurement of activity of single molecules of beta-D-galactosidase. Proc Natl Acad Sci USA. 1961;47:1981–91.

Song H, Chen DL, Ismagilov RF. Reactions in droplets in microflulidic channels. Angew Chem Int Edit. 2006;45:7336–56.

Taly V, Kelly BT, Griffiths AD. Droplets as microreactors for high-throughput biology. Chembiochem. 2007;8:263–72.

Tang J, Jofre AM, Lowman GM, Kishore RB, Reiner JE, Helmerson K, Goldner LS, Greene ME. Green fluorescent protein in inertially injected aqueous nanodroplets. Langmuir. 2008;24:4975–78.

Zhang YH, Ozdemir P. Microfluidic DNA amplification-A review. Anal Chim Acta. 2009;638:115–25.

Nanofiber Biosensors

- ▶ Nanowire Biosensors

Nanomaterials

- ▶ Nanoparticles for Drug and Gene Delivery

Nanomedicine

► Nanoparticles for Drug and Gene Delivery

Nanometer-Scale Apertures

► Zero-Mode Waveguides

Nanoparticles for Drug and Gene Delivery

Maya Thanou
Institute of Pharmaceutical Science, King's College London, London, UK

Synonyms

Nanomaterials; Nanomedicine

Definition

Particles with one or more dimensions of the order of 100 nm or less designed for the targeted delivery of drug and genetic therapies.

Nanoparticles for Therapeutic Applications

Today, there is a strong focus of nanotechnology application on certain challenging diseases. Cancer nanotechnology is a new field of interdisciplinary research cutting across biology, chemistry, engineering, and medicine aiming to lead major advances in cancer detection diagnosis and treatment (Ferrari 2005; Nie et al. 2007). The opportunity lies in the fact that for the first time, we are able to tackle cancer management needs and individualize therapies by developing personalized treatments. Ideally, clinical scientists should detect disease markers and at the same time formulate the nanoparticles targeting such markers to deliver specifically the drug or gene (Wang and Thanou 2010). For the first time, scientists can design and prepare multifunctional nanoparticles able to (a) detect disease, (b) image disease, (c) treat, and (d) monitor treatment progression. This can be done by assembling functional molecules in one nanosized particle. The application and efficiency of these nanoparticles in vivo will help enormously in the treatment of diseases.

Nanoparticles have been widely studied for drug and gene delivery in tumors. Most nanoparticles are expected to accumulate in tumors due to the enhanced permeation and retention effect (ERP) identified by Maeda as a means to target anticancer agents (Matsumura et al. 1986). Tumors are characterized by highly permeable blood vessels due to rapid and defected angiogenesis. In addition, tumors are characterized by dysfunctional lymphatic drainage that helps the retention of nanoparticles long enough to allow local nanoparticle disintegration and release of the drug in the vicinity of tumor cells. The phenomenon has been used widely to explain the efficiency of nanoparticle and macromolecular drug accumulation in tumors.

The first nanoparticles used to deliver cancer chemotherapy were the liposomes. Liposomes were discovered 40 years ago by Bangham (Cohen and Bangham 1972). Liposomes are vesicles with an aqueous interior surrounded by one or more concentric bilayers of phospholipids with a diameter ranging from 30 nm to several microns. Liposomes may vary in size, lipid composition, method of preparation, and particularly surface chemistry. Liposomes have been evolved through the years to a versatile carrier adapted each time to have a different functionality and serve a certain drug delivery purpose.

Liposomes, sized at the nanoscale, consist of a lipid bilayer surrounding a water core hosting the drug. The first studies to report the efficiency of liposomes as nanoparticles focused on the improvement of pharmacokinetics and biodistribution of the anthracycline drug doxorubicin. Doxil, the PEGylated liposomal doxorubicin, shows high efficiency due to improved pharmacokinetics as it has been shown to escape the reticuloendothelial system (RES), an important barrier in nanoparticle systemic circulation (Gabizon and Martin 1997). Polyethylene glycol is used with a number of nanoparticles as it improves colloidal stability and prevents uptake by the RES. PEG is usually added on the surface of nanoparticles to create the so-called steric stabilization effect where the PEG

molecules form a protective hydrophilic layer on the surface of nanoparticles that prevents interaction with each other (aggregation) and with blood components. As a result, grafting of PEG on the surface of nanoparticles reduces uptake by the macrophages of the mononuclear phagocyte system (MPS) and prolongs the blood circulation times (Papahadjopoulos et al. 1991).

Nanoparticles have been suggested for the treatment of diseases of the central nervous system (CNS) such as stroke Alzheimer's and Parkinson's disease. CNS diseases represent the largest and fastest growing area of unmet medical need. Several major types of nanoparticles have been widely used for drug delivery to the brain, such as liposomes, dendrimers, and polymeric micelles (Yang 2010).

Nanoparticles with functions such as imaging and drug delivery have been suggested for atherosclerosis and cardiovascular diseases (Lobatto et al. 2011).

The design of nanoparticles for drug and gene delivery is not trivial as several factors have to be taken into consideration. Primarily, the chemistry of the core component that will carry the drug or gene and the biocompatibility layers need to be designed considering the structural integrity and stability in biological fluids. Further similar to the process of product drug development, nanoparticle development needs to consider physicochemical issues related to the properties of the nanoparticles, biopharmaceutical issues related to the properties of the "bio-barriers," and pharmacological issues related to the site, time, and duration of nanoparticle's action. These nanoparticles have to be considered different to small and large molecular drugs introducing novel parameters for their design. FDA states that ADME (Administration, Distribution, Metabolism, and Excretion) studies need to be redesigned in the case of nanoparticles to take under consideration their aggregation and surface characteristics (Zolnik and Sadrieh 2009).

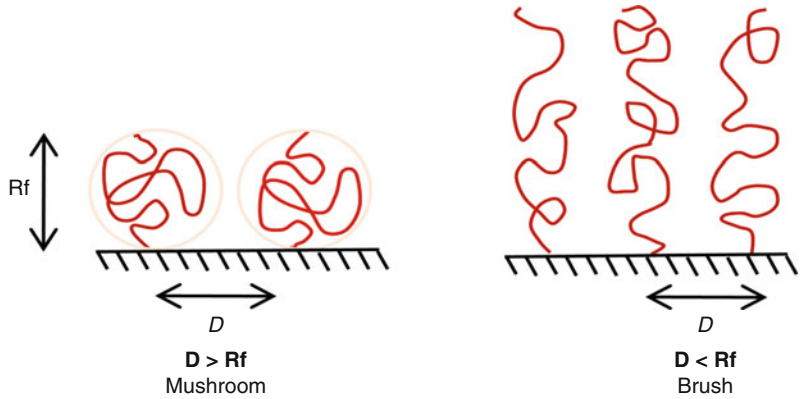
Nanoparticle Surface Constituents; the Stealth Layer

PEG layer (corona) is added onto any nanoparticle that is aimed to be administered intravenously. Either this is liposome, dendrimer, or inorganic nanoparticles, such a layer provides long circulation as it inhibits the accumulation of opsonins and dysopsonins and their uptake by the macrophages. There are currently no rules about the organization and the type of PEG on the surface of nanoparticles. However, there are some trends indicating that the PEG molecular weight and density on the nanoparticles are key features that will control PEG conformation on the surface of the nanoparticle and the circulation half-life ($t_{1/2}$). Such PEG layer can interfere with targeting ligands and inhibit them from interaction with the corresponding receptors.

Most nanoparticles require colloidal stability in vitro (in aqueous buffers, i.e., for storage) and in vivo (biological stability). Their protective layer is formed by the incorporation of a hydrophilic polymer layer. In this case, PEG is the mostly used material. PEG molecules are commonly adopted for the stealth function, due to their enhanced hydrophilicity and flexibility (Bergstrom et al. 1994). The amphiphilic nature of the PEG-lipids renders them into micellar structures, and their insertion into liposomal surfaces is perceived as a method of relieving micellar strain.

In liposomes, the incorporated PEG-lipid conjugates need to be able to render steric stability of the nanoparticles and be able to prevent surface adsorption of blood proteins. When presented on a liposomal surface, individual PEG chains exhibit a Flory dimension, Rf, which represents the volume that each flexible PEG cloud occupies (Fig. 1). Longer polymer chains have larger Rf, although studies have shown that the PEG density on a liposome surface is more important than the size of the polymer (Dos Santos et al. 2007). The Rf of a PEG_{2000} chain is approximately 5.6 nm and, in conjunction with grafting density, affects the resultant conformation of the PEG chains. Increasing the amount of PEG-lipids within a liposomal formulation increases the PEG grafting density, which ultimately reduces the distance, *D*, between each PEG molecule on the nanoparticle surface. When $D > Rf$, the PEG chains will self-assemble into a random-coil-like *mushroom* cloud. When $D < Rf$, the lateral pressure between the overcrowded PEG mushroom clouds will force the extension of the PEG chains into a *brush* conformation. The brush regime results from the increased lateral pressure between the PEG chains, which forces the extension of the polymers away from the surface, into more linear conformations.

Inorganic nanoparticles require different methods of coating or introducing the biocompatibility and colloidal stability layers. Generally, this is performed



Nanoparticles for Drug and Gene Delivery, Fig. 1 Representation of different PEG conformations, formed through their incorporation onto surfaces at different densities. When the distance between the PEG-lipids, *D*, is greater than the Flory diameter of the PEG (D > Rf), the polymer will collapse into a mushroom configuration. When D < Rf, the PEG chains will be extended from the surface to form a brush-like conformation

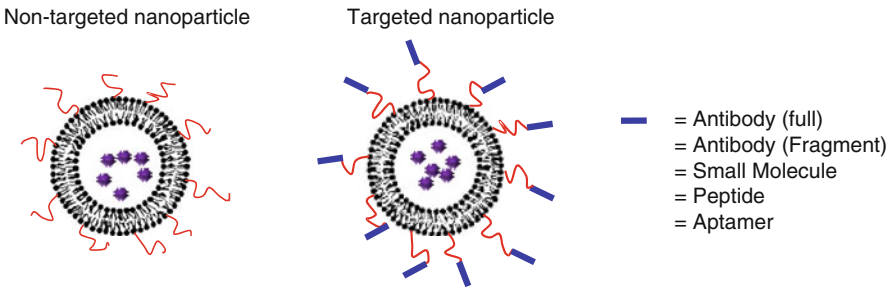
by the co-precipitation of the particles with various types of polymers or cross-linked polymers, a technique that can improve the particle monodispersity, an important parameter regarding their biological applications. Dextran, albumin, and PEG (Mw 5000) are all types of molecules that have been studies as coating materials for providing biocompatibility to iron oxide nanoparticles whereas gold nanoparticles provide a more suitable surface for PEG grafting. Thiol-PEG or bifunctional PEG can be added on the surface to provide colloidally stable and biocompatible gold nanoparticles (Shenoy et al. 2006).

Nanoparticle Surface Constituents; Targeting Ligands

The addition of a targeting moiety onto the surface of nanoparticles (Fig. 2) aims to increase selective cellular binding and internalization through receptor-mediated endocytosis. Without the incorporation of targeting ligands, nanoparticles rely on nonspecific interactions with cell membranes, which can be low when covered in a layer of PEG polymers. The targeting effect of the surface ligands is dependent not only on the nature of the ligand itself, but on a variety of factors that require cooperative optimization.

One challenge of targeting diseased tissue (i.e., tumors) is that the defective cells are often very similar in characteristics to its surrounding healthy tissue. To differentiate such cells, the ligands can be designed to have specificity for receptors that are overexpressed on cancerous cells, but are normally or minimally expressed on normal, healthy cells. These molecules should have high affinity to their cognate receptors, plus have innate abilities to induce receptor-mediated endocytosis. The targeting layer poses as the outmost exterior of the nanoparticle delivery system, where targeting ligands are generally presented on top of the stealth layer (Wang et al. 2008). Structures such as antibodies, antibody fragments, proteins, small molecules, aptamers, and peptides have all demonstrated abilities to induce nanoparticle targeting to cancer cells.

Antibodies against the HER-2 receptor, the transferrin receptor (TfR) and the prostate-specific antigen receptor, are both common examples of receptor targets, due to their overexpression of such receptors on cancer cells (Wagner et al. 1994). These antibodies generally exhibit strong interactions with corresponding receptors, with dissociation constants in the nanomolar range. Antibody fragments, consisting of only the Fab binding regions, have also been studied as targeting ligands (Tang et al. 2005). The advantages of such structures are that they are smaller and do not contain the Fc region of the antibody which can induce immunogenicity and antigenicity when present on liposome surfaces. Like antibodies and antibody fragments, the use of whole proteins is also commonly considered as targeting ligands for their increased affinity for target receptors. For example, the natural ligand for TfR, transferrin, binds to its receptor with a dissociation constant of around $K_d = 40$ nM. On overcoming systemic barriers to arrive in tumor vicinities, targeted liposomes have



Nanoparticles for Drug and Gene Delivery, Fig. 2 Representation of targeted and nontargeted liposomal nanoparticle systems. The targeted nanoparticles are functionalized with an outer layer of receptor-specific ligands, which can be of an antibody (full or fragment), small molecule, peptide, or aptamer

greater opportunities for binding to cancer cells if the ligand has naturally high affinities for the corresponding receptor.

Peptide sequences pose as even shorter and smaller versions of antibody and protein fragments. A prominent example of peptide ligands is the RGD peptide, identified through phage display to have high affinity to $\alpha_5\beta_3$ integrin receptors overexpressed on angiogenic vasculatures (Ruoslahti 1996). Other types of shorter ligands include small molecules folic acid specific for the folate receptors on ovarian cells.

Ligands chosen for receptor targeting should have the function of inducing receptor-mediated endocytosis. Depending on the type of ligand-receptor interaction, the rate of cellular internalization would differ. This is an important factor as rates of internalization could affect the accumulation of nanoparticle in targeted sites. Some ligand species, such as folate, have been shown to have fast internalization as well as cell surface recycling rates in cancer cells (Paulos et al. 2004). The ligands also need to be conjugated onto nanoparticles in an optimal fashion, as to maintain their affinities for their corresponding receptors. For example, the measured dissociation constants for optimally conjugated HER-2 antibodies were similar to free antibodies, around 12–15 nM, depending on the surface PEG density. The targeting layer is presented as the ultimate exterior of the nanoparticle, the surface that interacts primarily with cell membranes. Irrespective of their nature, the ligand must have the right conformation, high affinities for corresponding receptors, and be able to exhibit high rates of cellular internalization. Furthermore, the loading and presentation of the ligand on nanoparticle surfaces must be cooperative with the stealth layer.

It is important to incorporate PEG chains at densities that allow optimal coverage, which means inducing brush conformations on nanoparticle surfaces, as discussed earlier. As the brush conformation allows the extension of the PEG chains away from the nanoparticle surface, presenting targeting ligands on such a platform should allow increased interactions with corresponding receptors. Although the presentation of ligands on a PEG-brush surface enhances their binding to receptors, their improved exposure can decrease the nanoparticle lifetimes in circulation. Faster clearance from circulation generally corresponds to a higher and faster accumulation into the liver and the spleen (RES). In this example, the total amount of PEG_{2000} content needs to be included at optimal brush densities, at 6.5 mol%. Presentation of the ligand in a PEG-brush fashion reveals the targeting molecules increasingly to plasma proteins; hence, their faster clearance from circulation could be a result of Fc-mediated RES uptake into liver and spleen-associated macrophages (Elbayoumi and Torchillin 2006).

Concentration of surface ligands is another parameter that affects the ligands' targeting effect (Gantert et al. 2009). Higher ligand densities are envisioned as a method of increasing the probability of nanoparticle interactions with cell receptors (multivalency). However, the presence of increased non-PEG-like material on nanoparticle surfaces can be more detrimental than advantageous to delivery. A study using aptamers as a targeting ligand for prostate cancer–specific antigens (PSMA) demonstrated that higher densities of surface ligands resulted in greater accumulation of nanoparticles into the liver and the spleen (Gu et al. 2008). In that study Gu et al. used poly (D,L–lactide–coglycolyde) [PLGA] and PEG triblock copolymer–based nanoparticles.

The nanoparticles were composed by PLGA-b-PEG and PLGA-b-PEG-b-Apt (aptamer for PSMA) (Gu et al. 2008). Localization of such nanoparticles in tumors was also at lower concentrations compared to nanoparticles functionalized with lower densities of aptamers. This is suggesting that higher coverage of the PEGylated nanoparticles' surfaces with targeting ligands further shields the effect of the PEG layer, hence resulting in greater recognition by tumor and spleen-associated macrophages.

Examples of targeted nanoparticle drug delivery systems have shown faster and higher concentrations accumulation in tumors, compared to nontargeted systems. Generally, it is believed that the introduction of targeting ligands does not enhance nanoparticle accumulation into tumors, but shows higher efficacy by enhancing internalization into tumor cells.

Summary

Nanotechnology applications in medicine, termed nanomedicine, have introduced a number of nanoparticles of variable chemistry and architecture. Nanotechnology involves engineering multifunctional devices with dimensions at the nanoscale, similar dimensions as those of large biological vesicles or molecules in our body. These devices typically have features just tens to hundred nanometers across and they can carry one or two detection signals and/or therapeutic cargo(s). Being inspired by physiologically existing nanomachines, nanoparticles are designed to safely reach the target tissue and specifically release their therapeutic cargo at the site of the disease, thereby increasing the bioavailability of the drug at the site of action. Nanoparticles have the advantage of targeting certain diseased tissues (i.e., inflammation or tumors) by simply being accumulated preferentially in these tissues (passive targeting). The phenomenon is called the enhanced permeation and retention (EPR) effect, and has been proposed to explain accumulation of synthetic colloids in tumors. It is caused by leaky angiogenic vessels and poor lymphatic drainage and has been used to explain why macromolecules and nanoparticles are found at higher ratios in tumors compared to normal tissues. Nanoparticles have been suggested to treat atherosclerosis and diseases of the CNS as they can permeate membranes such as the blood-brain barrier. Several materials are suggested for the assembly of therapeutic nanoparticles. Polymers, linear and dendrimers, are associated with the drug in a covalent or noncovalent way and have been used with or without a targeting ligand (e.g., to direct the nanoparticle to a disease-related cell surface expressed receptor). Stealth liposomes are suggested to carry the drug in their aqueous core, and they are usually decorated by recognition molecules, being widely studied in the bench and applied in clinic as well. Inorganic nanoparticles such as gold and iron oxide are usually coupled chemically to the drug, PEG, and the targeting ligand. It appears that the PEG coating and ligand decoration are common constituents in most types of nanoparticles. There are several examples of successful drug delivery nanoparticles and many of them have rapidly moved to clinical trials. Nevertheless, there is still room for optimization in the area of the nanoparticle kinetics such as improving their plasma circulation and bioavailability and understanding the effect of targeting ligands on their efficiency to attach on diseased tissues. The need to develop novel and efficient ligands has never been greater, and the use of proper conjugation chemistry is mandatory.

Cross-References

- ▶ Carbon Nanotubes As Biomaterials
- ▶ Chemical Diversity of Lipids
- ▶ Dendrimers for Drug Delivery
- ▶ Glycoproteins
- ▶ Molecular Recognition: Lock-and-Key, Induced Fit, and Conformational Selection

References

Bergstrom K, Osterberg E, et al. Effects of branching and molecular weight of surface-bound poly(ethylene oxide) on protein rejection. J Biomater Sci Polym Ed. 1994;6:123–32.

Cohen BE, Bangham AD. Diffusion of small nonelectrolytes across liposome membranes. Nature. 1972;236:173–4.

Dos Santos N, Allen C, et al. Influence of poly(ethylene glycol) grafting density and polymer length on liposomes: relating plasma circulation lifetimes to protein binding. Biochim Biophys Acta. 2007;1768:1367–77.

Elbayoumi TA, Torchilin VP. Enhanced accumulation of long-circulating liposomes modified with the nucleosome-specific monoclonal antibody 2C5 in various tumours in

mice: gamma-imaging studies. Eur J Nucl Med Mol Imaging. 2006;33:1196–205.
Ferrari M. Cancer nanotechnology: opportunities and challenges. Nat Rev Cancer. 2005;5:161–71.
Gabizon A, Martin F. Polyethylene glycol-coated (pegylated) liposomal doxorubicin. Rationale for use in solid tumours. Drugs. 1997;54(Suppl 4):15–21.
Gantert M, Lewrick F, et al. Receptor-specific targeting with liposomes in vitro based on sterol-PEG(1300) anchors. Pharm Res. 2009;26:529–38.
Gu F, Zhang L, et al. Precise engineering of targeted nanoparticles by using self-assembled biointegrated block copolymers. Proc Natl Acad Sci USA. 2008;105:2586–91.
Lobatto ME, Fuster V, et al. Perspectives and opportunities for nanomedicine in the management of atherosclerosis. Nat Rev Drug Discov. 2011;10:835–52.
Matsumura Y, Maeda H. A new concept for macromolecular therapeutics in cancer chemotherapy: mechanism of tumoritropic accumulation of proteins and the antitumor agent smancs. Cancer Res. 1986;46:6387–92.
Nie S, Xing Y, et al. Nanotechnology applications in cancer. Annu Rev Biomed Eng. 2007;9:257–88.
Papahadjopoulos D, Allen TM, et al. Sterically stabilized liposomes: improvements in pharmacokinetics and antitumor therapeutic efficacy. Proc Natl Acad Sci USA. 1991;88:11460–4.
Paulos CM, Reddy JA, et al. Ligand binding and kinetics of folate receptor recycling in vivo: impact on receptor-mediated drug delivery. Mol Pharmacol. 2004;66:1406–14.
Ruoslahti E. RGD and other recognition sequences for integrins. Annu Rev Cell Dev Biol. 1996;12:697–715.
Shenoy D, Fu W, et al. Surface functionalization of gold nanoparticles using hetero-bifunctional poly(ethylene glycol) spacer for intracellular tracking and delivery. Int J Nanomedicine. 2006;1:51–7.
Tang Y, Scollard D, et al. Imaging of HER2/neu expression in BT-474 human breast cancer xenografts in athymic mice using [(99m)Tc]-HYNIC-trastuzumab (Herceptin) Fab fragments. Nucl Med Commun. 2005;26:427–32.
Wagner E, Curiel D, et al. Delivery of drugs, proteins and genes into cells using transferrin as a ligand for receptor-mediated endocytosis. Adv Drug Deliv Rev. 1994;14:113–35.
Wang AZ, Gu F, et al. Biofunctionalized targeted nanoparticles for therapeutic applications. Expert Opin Biol Ther. 2008;8:1063–70.
Wang M, Thanou M. Targeting nanoparticles to cancer. Pharmacol Res. 2010;62:90–9.
Yang H. Nanoparticle-mediated brain-specific drug delivery, imaging, and diagnosis. Pharm Res. 2010;27:1759–71.
Zolnik BS, Sadrieh N. Regulatory perspective on the importance of ADME assessment of nanoscale material containing drugs. Adv Drug Deliv Rev. 2009;61:422–7.

Nanovid Microscopy

► Single Particle Tracking

Nanowire Biosensors

Xian-En Zhang, Dong Men and Hongping Wei
State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, People's Republic of China

Synonyms

Bionanowire sensors; Nanofiber biosensors

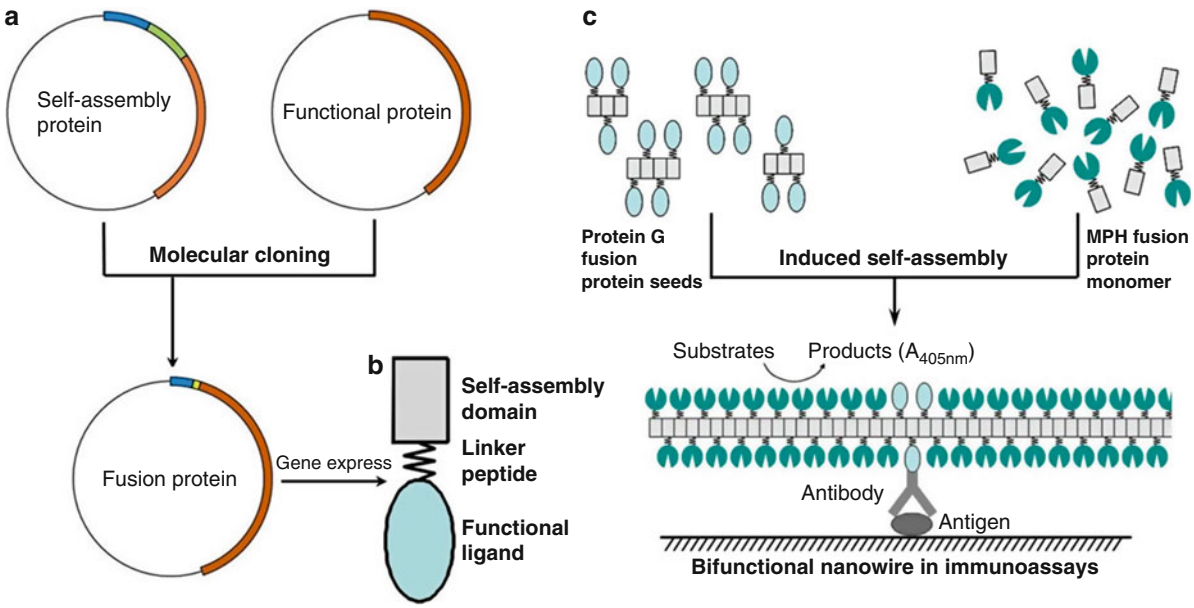
Definition

Nanowire biosensors: A class of sensors or measuring devices, of which the major sensing components are made of wires or fibers in nanometer scale formed by biological molecules, such as DNA molecules, polypeptides, fibrin proteins, and filamentous bacteriophages, etc.

Basic Characteristics

Nanowire biosensors are a class of nanosensors, of which the major sensing components are made of nanowires formed by biological molecules, referred to as bionanowires, such as DNA molecules, polypeptides, fibrin proteins, and filamentous bacteriophages, etc. A bionanowire is a one-dimensional fibril-like nanostructure, with the diameter constrained to tens of nanometers or less ($1 \sim 100 \times 10^{-9}$ m) and unconstrained length. Bionanowires offer a few significant advantages. First, biological molecules themselves are important parts of biosensors, they are capable of recognizing the targets or generating signals; second, the bionanowires can be easily modified by gene manipulation, thus providing versatile functions; third, by rational design, the self-assembly characters of bionanowires can display functional ligands on the nanowire's surface in designed orders; fourth, they possess more reaction area per volume to improve detection sensitivity and response time due to their high surface to volume ratio.

For these reasons, bionanowires are excellent candidates for sensing applications. Usually, bionanowires play a role of a mediator to integrate



Nanowire Biosensors, Fig. 1 Modular designed protein nanowire biosensor. (**a**) The construction of self-assembly unit in protein nanowires: gene fusion protocol. (**b**) The structure of fusion protein. There are three modules in fusion protein: self-assembly domain drives the fusion protein assembling into nanowire; the functional ligands generate biosensing signals; the linker peptide links the self-assembly domain and functional ligand and provide space to ensure both fusion partners remain in native activities. (**c**) The example of self-assembly bifunctional protein nanowire in biosensing. Bifunctional protein nanowires were generated by seeding-induced self-assembling of yeast amyloid protein Sup35p that genetically fused with protein G and an enzyme (methyl-parathion hydrolase, MPH), respectively. The protein nanowires possessed a high ratio of enzyme molecules to protein G, allowing a dramatic increase of the enzymatic signal when protein G was bound to an antibody target. As a result, a 100-fold enhancement of the sensitivity was obtained when applied in the detection of the *Yersinia pestis* F1 antigen (Reprinted with permission from Men et al. 2009. © 2009 American Chemical Society)

recognition (or binding) element and transduction (or signal) element in nanowire biosensors. The special dimensions of bionanowires could be explored to orient monofunctional or multifunctional molecules in high density along the longitudinal or vertical direction of the nanowires. This will help improve the detection sensitivity of a biosensor or develop multipurpose biosensors. In the process of sensing, functionalized bionanowires recognize or capture the targets, and then convert the recognition into chemical or physical signals, which are measurable.

DNA (▶ Infrared Spectroscopy of Protein Dynamics: Ultrafast Kinetics) is an ideal nanowire template, its diameter is at nanoscale (2 nm), it can be targeted uniquely at both ends, and has controllable length through its production by DNA synthesizers or polymerases. DNA nanowire biosensors are normally based on a reaction that produces electronic or optical signals by ordered DNA-protein conjugates on the DNA nanowire. Organizing functional molecules sequentially on the DNA nanowires could effectively enhance multienzymes catalysis systems and signal transduction (Carsten and Itamar 2010). One example is a sensor based on a nucleic acid-tethered ferrocene relay unit and GOx modified with a nucleic acid tether hybridized with a DNA scaffold that is covalently linked to an electrode (Piperberg et al. 2009). The programmed functionalized DNA nanowire has found efficient electrical contact between the biocatalysis center and the electrode, thus allowing bioelectrocatalytic activation of the enzyme.

Some peptides, such as the diphenylalanine peptide of Alzheimer's β-amyloid and the asparagine and glutamine-rich polypeptide of yeast prion, have the intrinsic ability to self-assemble into elongated solid nanofibrils or nanotube (an elongated nanowire with a definite inner hole). Because of their well-defined nanostructures

formed by self-assembly, peptide nanowires offer an attractive alternative for biosensor fabrication. In one example, a nonmediated amperometric biosensor was constructed based on immobilization of modified peptide nanotubes on electrode surfaces and the sensitivity for hydrogen peroxide and nicotinamide adenine dinucleotide (NADH) detection was improved (Yemini et al. 2005a, b). In another example, a label-free pathogen ensors chips (▶ Peptide Micro-array) is developed by self-assembling peptide nanowires. The antibodies-coated peptide nanowire is as a bridge between a pair of electrodes. The binding of the virus to the peptide nanotube was detected by measuring capacitance change between the electrodes (de la Rica et al. 2008).

Protein nanowires are fibril-like nanostructure formed by protein self-assembly, such as amyloid protein (▶ Amyloid Protein Biomaterials), collagen, silk protein, filamentous bacteriophage, and bacterial flagella (▶ Bacterial Flagellar Motor: Overview), etc. The surface of protein nanowires can be functionalized through gene manipulation. In a protein nanowire–based biosensor, both biological signal molecules and binding molecules are fused genetically with a self-assembly domain, respectively (▶ Protein Design for Biosensors). They are then integrated into protein nanowire through self-assembly. The resultant bifunctional protein nanowires possess a high ratio of signal molecules to biological sensing molecules, allowing a dramatic increase of the response signal when sensing to a target. Figure 1 shows a general protocol for construction of protein nanowire biosensors, which has been successfully applied to build ultrasensitive biosensors for immuneassay and detection of methyl parathion, etc. (Men et al. 2009, 2010; Leng et al. 2010).

Another kind of nanowire biosensors is biomolecule-inorganic hybrid nanowire biosensors. Hybrid nanowire biosensors are mainly made of biological sensing molecules with modified inorganic nanowires, such as silicon nanowires (▶ Carbon Nanotubes as Biomaterials), and semiconductor nanowires that have large surface areas, and show excellent field effect or electrical properties (Patolsky et al. 2006; Chris Le et al. 2007). These biosensors significantly lower the detection limit.

Cross-References

▶ Amyloid Protein Biomaterials
▶ Bacterial Flagellar Motor: Overview
▶ Carbon Nanotubes as Biomaterials
▶ Infrared Spectroscopy of Protein Dynamics: Ultrafast Kinetics
▶ Protein and Peptide Arrays
▶ Protein Design for Biosensors

References

Carsten T, Itamar W. Organizing protein-DNA hybrids as nanostructures with programmed functionalities. Trends Biotechnol. 2010;28(12):619–28.
de la Rica R, et al. Label-free pathogen detection with sensor chips assembled from peptide nanotubes. Angew Chem Int Ed Engl. 2008;47(50):9752–5.
Le Chris X, et al. Ultrasensitive assays for proteins. Analyst. 2007;132(8):724–37.
Leng Y, et al. Integration of a fluorescent molecular biosensor into self-assembled protein nanowires: a large sensitivity enhancement. Angew Chem Int Ed Engl. 2010;49(40):7243–6.
Men D, et al. Seeding-induced self-assembling protein nanowires dramatically increase the sensitivity of immunoassays. Nano Lett. 2009;9(6):2246–50.
Men D, et al. An auto-biotinylated bifunctional protein nanowire for ultra-sensitive molecular biosensing. Biosens Bioelectron. 2010;26(4):1137–41.
Patolsky F, et al. Nanowire sensors for medicine and the life sciences. Nanomedicine (Lond). 2006;1(1):51–65.
Piperberg G, et al. Control of bioelectrocatalytic transformations on DNA scaffolds. J Am Chem Soc. 2009;131(25):8724–5.
Yemini M, et al. Novel electrochemical biosensing platform using self-assembled peptide nanotubes. Nano Lett. 2005a;5(1):183–6.
Yemini M, et al. Peptide nanotube-modified electrodes for enzyme-biosensor applications. Anal Chem. 2005b;77(16):5155–9.

NCS2 (Nucleobase-Cation-Symporter-2) or NAT (Nucleobases-Ascorbate-Transporter)

▶ Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA

NDH-1

▶ NADH-Ubiquinone Oxidoreductase (Complex I)

Near UV Protein CD

Alison Rodger
Department of Chemistry, University of Warwick, Coventry, UK

Synonyms

Protein aromatic circular dichroism; Protein side chain circular dichroism

Definition

Protein circular dichroism (CD) spectroscopy is usually separated into (1) far UV or backbone CD with data collected from ~190 to 250 nm and (2) near UV or aromatic CD with data collected from 250 to 300 nm. The aromatic chromophores of protein side chains (tryptophans, tyrosines, and phenyl alanines) are planar and so have no intrinsic CD. However, their location as part of an amino acid results in an induced CD signal in their transitions. This is further enhanced when they are located in the chiral environment of a peptide or protein. The signs and intensities of the CD signals induced into the achiral aromatic side chain transitions are dependent on their environment. It is a useful fingerprint of protein identity and conformation, but seldom leads to direct structural analysis. Disulfide bonds are often counted as honorary aromatic groups since their CD may be apparent between 250 and 270 nm. Wavelengths of relevant residues are given in Table 1.

When peptides with aromatic residues assemble (such as in a membrane), bands in the 230–240 nm region of the CD spectrum appear. They arise as a result of $\pi-\pi^*$ exciton coupling of aromatic transitions, where one of the components occur at about 235 nm and the shorter wavelength component is masked by the backbone transitions (Woody 2007).

Cross-References

- Circular Dichroism Spectroscopy of Biomacromolecules
- Circular Dichroism Spectroscopy: Units
- DNA-Ligand Circular Dichroism
- Far UV Protein Circular Dichroism
- Near UV Protein CD
- Oriented Circular Dichroism Spectroscopy
- Protein Circular Dichroism Analysis

References

Woody RW. Aromatic side chain contributions to protein circular dichroism. In: Uversky VN, Permyakov EA, editors. Luminescence spectroscopy and circular dichroism: methods in protein structure and stability. Hauppauge: Nova; 2007.

Near UV Protein CD, Table 1 Spectroscopic parameters for Phe, Tyr, and Trp side chains. The assignment is using the nomenclature of Platt. The oscillator strengths for degenerate or nearly degenerate pairs (the B bands of Phe and Tyr) are given as half the observed value for the entire band (Woody 2007)

	Assignment[a]	λ_{max}/nm	ε_{max}/(mol^{-1} cm^{-1} dm^3)	Oscillator strength	Apparent vibronic structure
Phe	L_b	260	200	0.001	~500 cm^{-1}
	L_a	210	10,000	0.09	–
	B_b	180	30,000	0.45	–
	B_a	180	30,000	0.45	–
Tyr	L_b	275	1,400	0.02	~800 cm^{-1}
	L_a	230	8,800	0.13	–
	B_a	190	24,000	0.55	–
	B_b	190	24,000	0.55	–
Trp	L_b	280	3,200	0.03	~850 cm^{-1}
	L_a	270	4,500	0.13	–
	B_b	220	35,000	0.6	–
	B_a	195	20,000	0.5	–

Near-Edge X-Ray Absorption Fine Structure (NEXAFS)

► X-Ray Absorption Spectroscopy of Metals in Biology

Nearest-Neighbor Spin-Spin Correlations

► Oriented-Sample NMR of Membrane Proteins: Sensitivity Enhancement and Spectroscopic Assignment

Near-Infrared Spectroscopy in Biological Molecules and Tissues

Yukihiro Ozaki[1], Akifumi Ikehata[2] and Hideyuki Shinzawa[3]
[1]Department of Chemistry, School of Science and Technology, Kwansei Gakuin University Sanda, Sanda, Hyogo, Japan
[2]National Food Research Institute, National Agriculture and Food Research Organization, Tsukuba, Japan
[3]Advanced Institute of Science and Technology, Nagoya, Aichi, Japan

Synonyms

Chemometrics in NIR spectroscopy; NIR imaging; NIR spectroscopy

Definition

This entry describes principle, instrumentation, analysis method (chemometrics), and biological applications of near-infrared (NIR) spectroscopy. Applications discussed include the studies of protein structure, denaturation and hydration, quantitative in vivo analysis of blood glucose, mapping brain function, and imaging.

Introduction

Near-infrared (NIR) spectroscopy is spectroscopy in the region of 800–2,500 nm (12,500–4,000 cm^{-1}), being mainly concerned with absorption spectroscopy and diffuse reflectance spectroscopy (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). NIR spectroscopy is relatively new in biophysics, but its application to biological materials stretched back more than 40 years. In the 1960s, Norris et al. (Ben-Gera and Norris 1968) tried to employ NIR spectroscopy to monitor water content in grain. They paid much attention on the potential of NIR spectroscopy as a nondestructive analytical method. It was the idea of using a statistical method for extracting useful information from complicated NIR spectra that woke the "sleeping Giant" up. The first trial of biomedical application of NIR spectroscopy was in vivo monitoring of the redox behavior of cytochrome *c* oxidase by Jöbsis in 1977 (Jöbsis 1977). After these two pioneering studies, NIR spectroscopy has been spreading to various fields, including biological and biomedical sciences, food and agricultural engineering, chemical engineering and online monitoring, polymer science and engineering, and environmental sciences (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007).

NIR spectroscopy contains both electronic spectroscopy and vibrational spectroscopy (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). As vibrational spectroscopy, it deals with overtones and combination modes. Since both the overtones and combination modes are forbidden transitions according to harmonic oscillator approximation, the intensities bands due to them are very weak. The intensities of bands due to electronic transitions in the NIR region are also weak. Thus, the NIR region is, in general, very superior in permeability compared to the visible (VIS) and mid-IR region, enabling nondestructive and noninvasive analyses.

The NIR region may be divided into three regions, the 800–1,100 nm (12,500–9,900 cm^{-1}) region, the 1,100–2,000 nm (9,900–5,000 cm^{-1}) region, and the 2,000–2,500 nm (5,000–4,000 cm^{-1}) region. The first region, named short wavelength NIR (SWNIR) region or near NIR region, is concerned mainly with electronic transitions, higher-order overtones, and combination transitions, and thus it shows particularly high permeability. This region provides so-called "a window for biological bodies" and is used practically for medical applications and agricultural applications. The second region contains many bands arising from the first and second overtones and combination modes.

This region is very useful for general quantitative and qualitative analyses. The third region mostly deals with combination modes. The permeability of this region is relatively low.

NIR spectroscopy has the following advantages for studies in biophysics:

1. It is a nondestructive and noninvasive method, and thus very suitable for in vivo studies of biological materials and tissues. Neuroimaging is a good example of noninvasive analysis of NIR spectroscopy.
2. It allows for noncontact analysis and analysis using an optical fiber.
3. It is much easier to study and analyze an aqueous solution, than with IR spectroscopy.
4. It can be applied to samples in various physical states, shapes, and thickness.

An NIR spectrum consists of a number of bands arising from overtones and combination modes that overlap heavily with each other, and thus so-called multicollinearity is very strong in the NIR region (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). Therefore, to extract useful information from the NIR spectrum is not always easy. Besides conventional spectral analysis methods, chemometrics has most often been used to extract information from NIR spectra.

Instruments

The most important elements for NIR instruments are light sources, spectrometers, and detectors.

Light Sources

To date, tungsten halogen lamps are unsurpassed as economical sources of radiation for NIR spectroscopy. The lamps show broad irradiation from 350 to 2,600 nm with a maximum at around 900 nm by the cleaning reaction of tungsten. The tungsten halogen lamps are ideally stable and incoherent light source; however, the heat generation may cause damage to the sample. In that case, an IR cut filter (a cold filter) is effective to eliminate the heat. Recent cold light sources, LEDs, will significantly grow over the next decades. NIR LEDs have relatively broad emission properties (FWHMs are about 50 nm). A long-life continuum source can be attained with a combination of several LEDs having different emission properties. High-power and narrowband laser diodes (LDs) are also useful if the specific wavelength is decided. Very broad continuum coherent sources, in other words, "white lasers," based on supercontinuum generation from a nonlinear device, such as an optical fiber, are now commercially available. The emission band ranges from 500 to 2,400 nm depending on the pump laser. The supercontinuum laser is a pulse source but as small as a laptop PC.

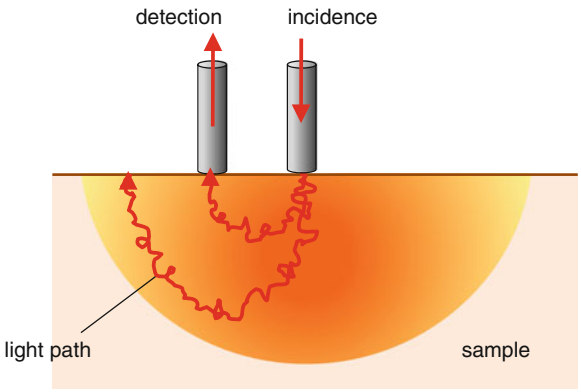
If one can use a tunable Ti:Sapphire laser, spectral scans between 700 and 1,000 nm can be performed without a monochromator. Although Ti:Sapphire laser has been known as a q-switched pulse laser, nowadays, continuous wave (CW) Ti:Sapphire laser is also available. In case one needs coherent light in a longer wavelength region (up to 2,600 nm), an optical parametric oscillator (OPO) with an Nd:YAG laser may be available. Digital Micromirror Device (DMD), which can be used instead of a monochromator, realizes a high-speed programmable spectral illumination. Since a DMD chip is a dispersive element having several hundred thousand microscopic mirrors on its surface, a DMD source can generate not only a single wavelength but also multiwavelengths (a spectral pattern) at once. Although a classic tungsten-halogen lamp is used as the original source, the ability of rapid wavelength change and special resolution will open a new avenue of NIR illumination.

Spectrometers

Each NIR spectrometer is not different from those used in the visible (VIS) or IR region. However, due to the wide NIR wavelength range, dispersion type and Fourier transform (FT) type spectrometers coexist in the NIR region (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). The dispersion spectrometers equipped with optical gratings cover a whole range of an NIR spectrum with good stability of the light intensity. On the other hand, FT spectrometers are well suited for mid- and long-wavelength regions of the NIR region and excellent at wavenumber (wavelength) accuracy.

Detectors and Others

PbS, InAs, and InGaAs detectors are frequently used for NIR spectrometers (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). For the SWNIR region, Si photodiodes (generally used for VIS) are also available (Siesler et al. 2002). Moreover, the recent development of array detectors is driving the expansion of NIR imaging instruments as will be discussed later.



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 1 Diffuse reflected light of interactance measurement

The use of fiber optics is a unique characteristic of NIR spectroscopy. Contact fiber probes, called interactance probes, realize nondestructive measurements of opaque samples. As shown in Fig. 1, incident light has to diffuse into a sample with minimal surface reflectance, and a part of the light arrives at a detection probe. An interactance method is not applicable to a transparent sample because it is based on diffuse reflectance of particles. The total path length of an interactance method is about a few cm in the SW NIR region. Although the penetration depth of the light and the distance between the incidence and detection probes are roughly related, for more precise experiments, time-resolved measurements must be performed.

Chemometrics

Chemometrics is a chemical discipline that uses mathematical and statistical methods to design or select optimal measurement procedures and experiments, and to provide maximum chemical information by analyzing chemical systems (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007), (Næs et al. 2002; Jiang and Ozaki 2002). It has been used extensively in NIR spectroscopy as a powerful tool to yield a mechanistic interpretation of complex physicochemical phenomena represented by a spectral data set. Since an NIR spectrum is often obtained as a highly convoluted form of overtone or combination bands arising from multiple chemical components in the system, key information in the NIR spectum is distributed broadly throughout data. It is possible to condense the information into a very compact form by chemometrics. Chemometrics mostly used in NIR spectroscopy are those for (1) classification, (2) multivariate calibration, and (3) self-modeling curve resolution (SMCR) (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007), (Næs et al. 2002; Jiang and Ozaki 2002).

In the practice of chemometrics for NIR spectra, a factor analytic technique called Principal Component Analysis (PCA) often plays an important role to elucidate distinct and useful information for detailed analysis of complex alternation of NIR spectra. PCA enables to condense the essence of the information present in the NIR spectral data into a very compact matrix representation referred to as scores and loadings. The basic hypothesis of such factor analysis techniques is that the improved proxy of the original data matrix can be reconstructed from only a limited number of significant factors. Thus, while the score and loading matrices contain only a small number of factors, it effectively carries all the necessary information about spectral features and leading to sorting out the convoluted information content of highly complex chemical systems.

The spectra can be represented as a matrix $\mathbf{X}$ of m by n dimension,

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \quad (1)$$

where m is the number of spectra traces in this data set and n is the number of data points per spectrum along the wavenumber (or any other spectral variable) axis. PCA may be viewed as mathematical decomposition of spectral data in terms of the orthogonal set of dominant factors, i.e., eigenvectors. PCA results in yielding two matrices called scores and loadings which complementarily represent the entire features broadly distributed in NIR spectral data.

In short, it becomes

$$\mathbf{X} = \mathbf{T}\mathbf{P}^t + \mathbf{E} \quad (2)$$

where $\mathbf{T}$ is an m by r PCA score matrix and $\mathbf{P}$ is an n by r PCA loading matrix, respectively. The rank r corresponds to the number of principal components

representing the significant portion of the information contained within the data matrix **X**. The selection of *r* is somewhat arbitrary. It is usually set to be a number, as small as possible but sufficiently large enough such that there are no obvious spectral features found in the residual matrix **E**. The residual matrix **E** is the portion of the original data, which is not accounted for by the first *r* principal components used for the data representation. Importantly, the two matrices **T** and **P** complementally represent the entire features broadly distributed in **X**. Namely, **T** holds abstract information concerning the relationship among the samples and **P** contains summary of variable, e.g., wave number which provides chemically or physically meaningful interpretation to the pattern observed in **T**. The detailed analysis of such matrices brings useful insight into building a mechanistic model for understanding complex phenomena studied by NIR spectroscopy.

Multivariate calibration is also a major chemometrics technique widely used in NIR spectroscopic study (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007), (Næs et al. 2002; Jiang and Ozaki 2002). The main purpose of calibration in NIR spectroscopy is to estimate the actual quantity of specific chemical components in the object by using NIR light. Because of the well-known collinearity problem of NIR spectra with so many data points in the wavenumber direction being highly correlated, the latent relationship among spectral variables cannot be obtained from straightforward general inverse technique like Multiple Linear Regression (MLR). Instead, a chemometrics technique called Partial Lest Squares (PLS) regression often provides a solid solution for the collinearity problem of NIR spectra. In PLS, one of the most important points lies in the fact that spectral data matrix **X** is decomposed into a linear combination of scores **T** and loadings **P** matrices. For example, the decomposition of **X** is carried out by repeatedly regressing the score along the given concentration vector **Y** containing the known concentrations of a specific analyte for a series of spectra. PLS modeling essentially ends up with a regression vector **b** which is a linear combination of the absorbance values at selected wavenumbers. The regression vector or matrix carries the information mostly relevant to the determination of the concentration of the analyte and less sensitiveness to the overall effect of interference. Once one obtained the regression vector **b**, it can later be used to estimate the concentration of the analytes even in unknown samples as long as their NIR spectra are measured.

Another important chemometric method, SMCR will be outlined in NIR Imaging section. It is important to point out that the application of chemometrics is not limited to NIR spectroscopy. It is generally applicable to various kinds of spectroscopy, e.g., IR and Raman. Similar techniques are also extensively used in Quantitative Structure-Activity Relationship (QSAR) studies which are often referred to as *Chemoinformatics*.

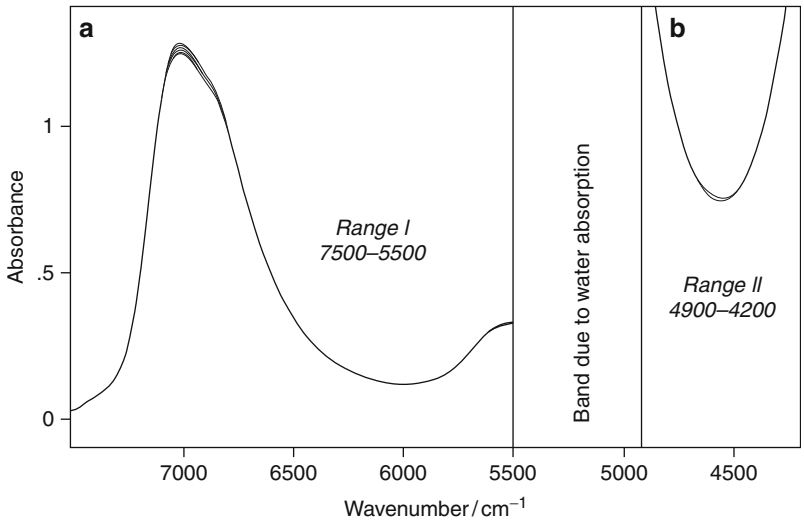
Applications

Biological Molecules

NIR spectroscopy has been extensively used to investigate structures and functions of a variety of biological molecules from water molecules, carbohydrates, to proteins (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). One of unique points of NIR spectroscopy in the studies of biological molecules is that it is powerful to explore hydrogen bondings or inter- and intramolecular interactions. This characteristic occurs because of the following two reasons. One is that the NIR region is dominated by bands ascribed to functional groups that contain a hydrogen atom (e.g., O-H, C-H, N-H). Another is that a band shift due to the formation of a hydrogen bonding is much larger in the NIR region compared with the IR region.

NIR spectroscopy has several advantages over IR spectroscopy in studying biological molecules. First, one can use a cell having thickness of 0.1–1 mm, and thus, the exact concentration of a biomolecule solution can be estimated. In the case of IR spectroscopy, one must use a very thin cell (on the order of micrometers) or an attenuated total reflection (ATR) prism for the IR measurement of the biomolecule solution. Therefore, there always remains a problem of adsorption of biological molecules on the cell. Another advantage of NIR spectroscopy is that the effects of bands due to water vapor are much weaker in the NIR region. Yet another advantage is that one can probe hydration of biomolecules.

In NIR studies of biological molecules, generalized two-dimensional (2D) correlation spectroscopy has often been used to unravel NIR spectra (Noda and Ozaki 2004). Wu et al. (2000) used 2D correlation NIR spectroscopy to explore temperature-induced



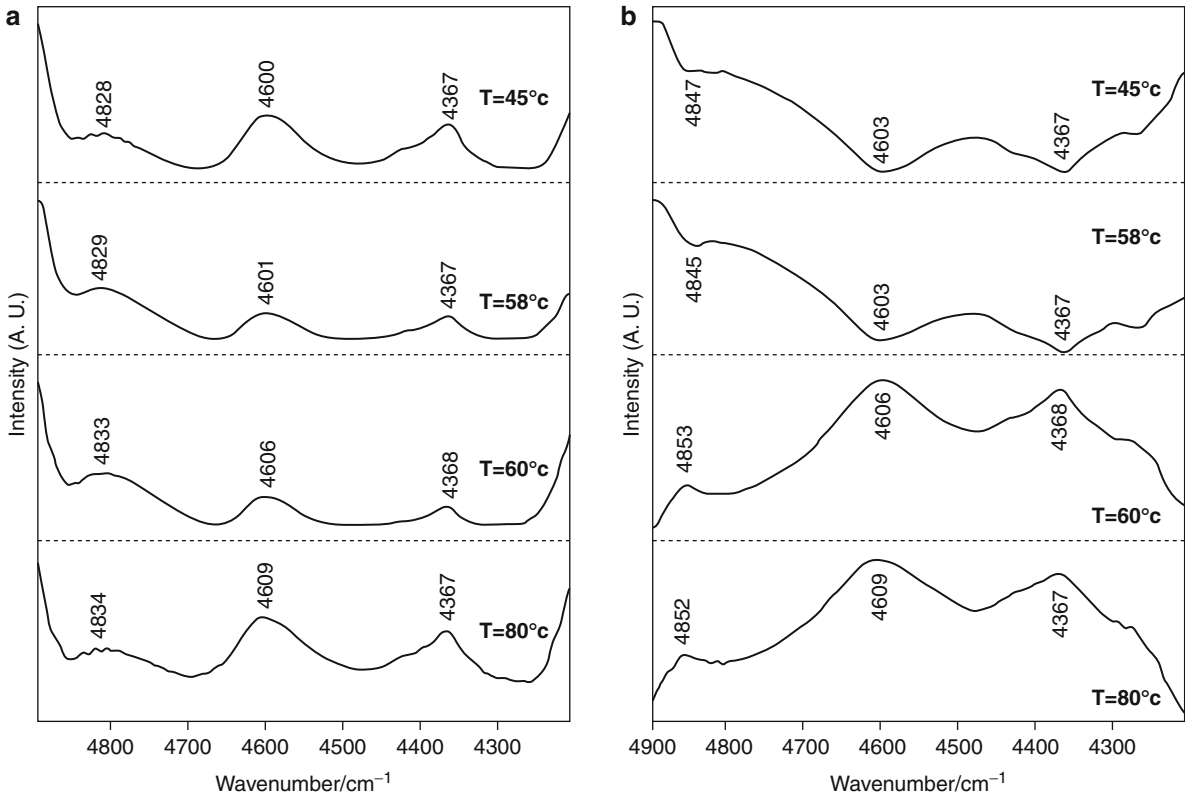
Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 2 FT-NIR spectra of HSA in aqueous solutions with concentration of 1.0, 2.0, 3.0, 4.0, and 5.0 wt.% over a temperature range of 45–80 °C (Reproduced from ref. Wu et al. 2000. With permission. Copyright (2000) American Chemical Society)

variation in the secondary structure and hydration of human serum albumin (HSA). Figure 2 shows FT-NIR spectra of HSA in aqueous solutions with concentrations of 1.0, 2.0, 3.0, 4.0, and 5.0 wt.% over a temperature range of 45–80°C (Wu et al. 2000). Concentration-perturbed 2D correlation spectra were calculated for the spectra in the 7,500–5,500 cm^{-1} and 4,900–4,200 cm^{-1} regions at different temperatures. To investigate temperature-induced changes in the secondary structure and hydration, power spectra and slice spectra (Noda and Ozaki 2004) were calculated from the synchronous and asynchronous spectra, respectively. Figure 3a, b depicts them, respectively (Wu et al. 2000). In the power spectra, a band near 4,600 cm^{-1} due to the combination mode of amide B and amide II (amide B/II) of HSA shows an abrupt shift by 5 cm^{-1} between 58°C and 60°C. Both the power and slice spectra in the 7,500–5,500 cm^{-1} region provide an explicit evidence that the hydration (water-protein interaction) changes markedly near 60°C. A comparison of temperature-dependent frequency shifts between the band near 4,600 cm^{-1} due to amide B/II of HSA and that near 7,000 cm^{-1} due to the combination mode of water reveals that the protein unfolding occurs almost in parallel with the change in the hydration. This study has demonstrated that the overtone and combination modes are very useful in monitoring subtle changes in protein dynamics (Wu et al. 2000).

Another interesting example of NIR studies of biological molecules is an investigation of C-H. . .O = C hydrogen bonding in a bacteria-synthesized biopolymer, poly (3-hydroxybutyrate) (PHB). PHB is one of the most extensively studied biodegradable thermoplastic polymers.

NIR spectroscopy has certain advantages in studying polymers and biopolymers (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007). In general, sample preparation and handling are much easier for NIR spectroscopy than for IR spectroscopy. It can be applied to thick films, sheets, granulates, powders, fibers, and textiles. Furthermore, in conjunction with light fiber technique, it enables online monitoring of polymer reactions and productions. NIR spectroscopy is also powerful in studying hydrogen bondings, hydration, and inter- and intramolecular interactions in polymers and biopolymers. By combining IR and NIR data, one can explore the anharmonicity of a band which is very sensitive to the formation and strength of a hydrogen bond.

Hu et al. (2006) investigated the C-H. . .O = C hydrogen bonding of PHB by using NIR spectroscopy. An NIR spectrum of a PHB film changes remarkably with the evolution of melt-crystallization process. Accordingly, NIR bands due to the amorphous and crystalline states can easily be assigned. Figure 4a shows temporal changes of NIR spectra in the range of 6,050–5,650 cm^{-1} of the PHB film during the melt-crystallization process at 125°C. Figure 4b depicts the second derivatives of the spectra measured at 0 and 180 min. Particular attention was paid



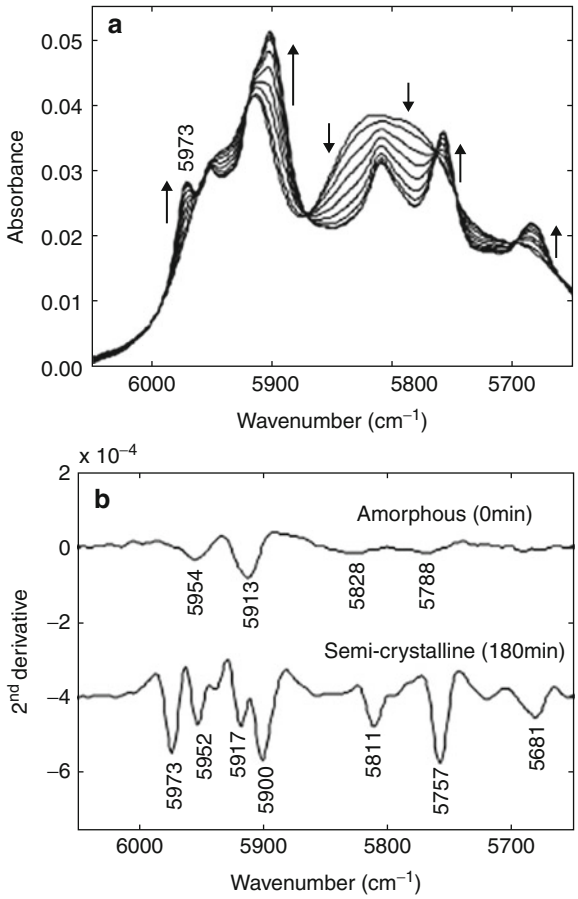
Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 3 (**a**) Power and (**b**) slice spectra of the synchronous and asynchronous 2D correlation spectra calculated from the spectra shown in Fig. 2 (Reproduced from ref. Wu et al. 2000. With permission. Copyright (2000) American Chemical Society)

to an NIR band at 5,973 cm^{-1} due to the first overtone of the C-H stretching mode and an IR band at 3,435 cm^{-1} due to the first overtone of the C = O stretching mode of the C-H. . .O = C hydrogen bonding. The anharmonicities of the C-H (5,973 cm^{-1}) and C = O (3,435 cm^{-1}) bands arising from the crystalline parts and those of the amorphous C-H (5,954 cm^{-1}) and C = O (3,457 cm^{-1}) bands were estimated by using both IR and NIR data. The anharmonicities of the crystalline C-H and C = O bands are significantly different from those of the amorphous C-H and C = O bands, supporting the idea that both C-H and C = O bonds in the crystalline state are involved in the C-H. . .O = C hydrogen bonding.

Cells and Tissues

Recently, noninvasive NIR measurements have achieved as time-domain and frequency-domain spectroscopy methods (Tromberg et al. 2000; Hawrysz and Sevick-Muraca 2000). The measurements of time-domain spectroscopy are performed by repetitive pulse and recording the phase delay at a fixed distance from the incidence site. For frequency domain spectroscopy, intensity-modulated NIR lights in MHz-GHz frequency range are used. (The continuous wave techniques may give the average distribution of path length in turbid samples from the spatial properties of scattered light shown in Fig. 1, but have no means of detecting the real path lengths of each photon.) These methods are extensively applied to detection of cancer cells as imaging and depth profiling techniques (Tromberg et al. 2000; Hawrysz and Sevick-Muraca 2000).

Metal nanoparticles are receiving particular attention as new molecular probes. Gold nanorods, whose size is about 10 nm in diameter and over 50 nm in length, show corresponding transversal surface plasma oscillation peak at 520 nm and longitudinal one at around 800 nm. The longer the length of rods grows, the longer the peak wavelength of the longitudinal mode shifts (Kim et al. 2002). The strongly absorbed



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 4 (**a**) Temporal changes of NIR spectra in the range of 6050–5650 cm^{-1} of the PHB film during the melt-crystallization process at 125 °C. (**b**) Second derivatives of the spectra measured at 0 and 180 min (Reproduced from ref. Hu et al. 2006. With permission. Copyright (2006) American Chemical Society)

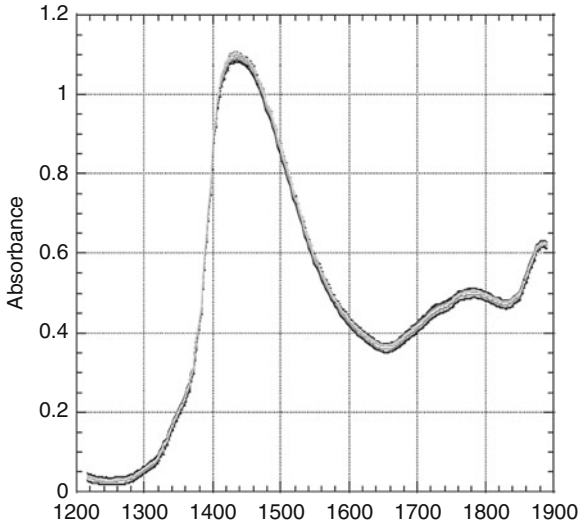
NIR light is converted efficiently into heat on a picosecond time domain due to electron–phonon and phonon–phonon processes. This heat generation is used for effective photothermal cancer therapy with low-power NIR laser irradiation (Huang et al. 2006). The excitation of surface plasma oscillations and the generation of strongly enhanced electric field are also used for surface-enhanced Raman scattering (SERS) (Guo et al. 2007).

Nondestructive and Noninvasive Analysis

Since NIR spectroscopy is a nondestructive analytical technique with no or little pretreatment, applications of NIR spectroscopy to in vivo biomedical analysis and diagnosis have received keen interest (Siesler et al. 2002; Ozaki et al. 2007; Burns and Ciurczak 2007; Workman and Weyer 2007), (Heise 2002; Du et al. 2007; Ozaki et al. 2009). Particularly, noninvasive brain function analysis (Heise 2002; Du et al. 2007) and noninvasive blood glucose measurement (Ozaki et al. 2009) have been matters of extensive NIR studies.

The in vivo blood glucose measurement with NIR spectroscopy is a very challenging project because it deals with very weak signals of glucose directly from human skin, and the physiological conditions of skin tissue such as body temperature change easily with time (Ozaki et al. 2009). One critical difficulty associated with in vivo blood assay is an extremely low signal-to-noise ratio of glucose existing in human issues. When NIR spectra are measured in vivo, background noise easily hides the glucose signal. To overcome this problem, Maruo et al. (2003, 2006) developed a new NIR spectrometer system with a set of two optical fibers to obtain the dermis spectra selectively. One set of optical fibers is attached to the skin surface vertically. The skin surface is illuminated by the measuring light through the inlet optical fiber, and the scattered light is collected by the detecting optical fiber (Fig. 1). The fiber distance controls the penetration depth of the measuring light as confirmed by computer simulations based on a Monte Carlo method. This property has been used to restrict the measure absorption to the dermis layer of skin. Figure 5 displays 50 NIR spectra of the forearm human skin of one subject measured during one oral glucose intake experiment (Maruo et al. 2006). A strong band at 1,450 nm is assigned mainly to the combination of the antisymmeric and symmetric O-H stretching modes of water. The sharp absorbance peak at 1,450 nm suggests that our system using the novel fiber probe can reduce interference from absorption by the stratum corneum.

An in vitro study using bovine serum by Maruo et al. (Maruo and Oka 1999) showed the important role of the glucose absorption peak at 1,580 nm for a glucose assay. The characteristics of the peak observed in the in vivo study described above are in accord with that of the in vitro study. This result means that the small absorbance change of glucose in vivo can be detected by the newly developed NIR instrument for blood glucose measurement (Maruo et al. 2003).



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 5 Fifty NIR spectra of the forearm human skin of one subject (Subject 6) measured during one oral glucose intake experiment (Reproduced from ref. Maruo et al. 2006. With permission. Copyright (2006) Society for Applied Spectroscopy)

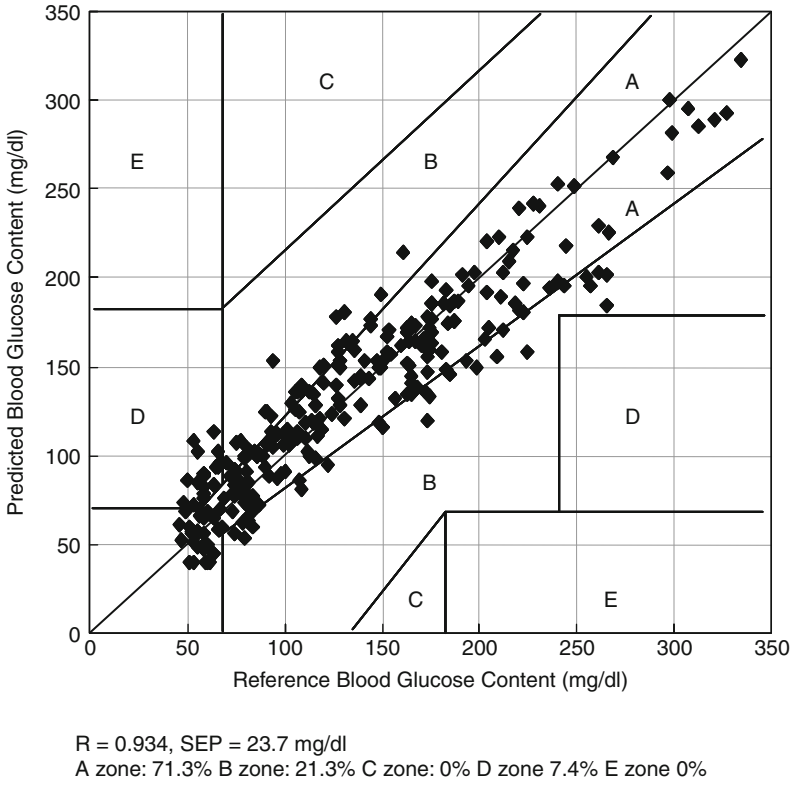
Six subjects were subjected to oral glucose intake experiment (Maruo et al. 2006). About 50 paired data sets, i.e., the reference blood glucose contents and the NIR spectra, were obtained for each experiment. Maruo et al. (2006) obtained the correlation coefficient, *R*; 0.934 and standard error of prediction (SEP); 23.7 mg/dl for the six subjects using individual calibration models. Figure 6 presents the result for Clarke error grid analysis obtained using the same data sets of the six subjects. The data plots included are: for the A zone, 71.3%; B zone, 21.3%; C zone, 0%; D zone, 7.4%; and E zone, 0%. Almost all prediction plots exist in zones A and B. However, it is noted that there are some plots in zone D. The plots in zones A and B are clinically acceptable, while those in zones C, D, and E are potentially dangerous, and, therefore, they may have the clinically significant errors. In this study, the plots in D zone have an overestimated value below 70 mg/dl of blood glucose contents. To improve these prediction errors below 100 mg/dl, not only the reduction of measurement noise but also the preparation of analytical data sets with adequate blood glucose contents is important.

Mapping brain function of humans with diffuse optical tomography (DOT) based on NIR spectropscopy has recently received great interest because it has led to a revolution in brainscience and medicine, enabling the functional mapping of the human brain on a scale that was previously accessible only through invasive studies of animals (Dehghani et al. 2009; White et al. 2009). DOT is uniquely suited to noninvasive brain and neuroimaging as it is a mobile system utilizing a small, flexible imaging cap. DOT can measure absolute changes in oxygenated (ΔHbO_2), deoxygenated (ΔHbR), and total hemoglobin (HbT), yielding more comprehensive images of the brain's hemodynamics. DOT methods are advancement beyond previous optical imaging techniques performed in topographic method, often referred to as diffuse optical imaging (DOI). DOT encapsulates a variety of technological improvements to resolution and depth sectioning. Recently, White et al. (White et al. 2009) revealed resting-state functional connectivity in the human brain with DOT. Figure 7 illustrates repeatability of functional connectivity of DOT over multiple imaging sessions (ΔHbR) (White et al. 2009). Correlation analysis from three sessions showed that the patterns were qualitatively similar over multiple days (Fig. 7). The results obtained by White et al. (2009) demonstrated the successful application of functional connectivity methods to DOT of adult human subjects.

NIR Imaging

NIR imaging is a technique to measure a spectrum for every pixel dividing an object into many spatial parts (Šašić and Ozaki 2010). A unique feature of such spectroscopic imaging technique is that it can take full advantage of spatial and molecular structural information about chemical moieties. Not only spectral variable but also spatial information is available in spectroscopic imaging. A graphical representation of NIR imaging is shown in Fig. 8. Each pixel point of an NIR image consists of a full NIR spectrum, and each microimage is a collection of spectral intensity measured at a fixed wavenumber.

Since a set of NIR imaging spectra contains information about the spatial dimension (X- and Y-axis) and spectral dimension (wave number or wavelength), the interpretation of complex variations of NIR imaging spectra becomes a challenging task. In the actual application of NIR imaging, it is often accompanied with some chemometrics techniques such as PCA and SMCR so that spatial as well as molecular structure information can be effectively extracted from the data. For example, scores represent physical or chemical distribution of main contributing sources among



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 6 Correlation between reference and predicted blood glucose contents. The lettered zones represent clinically relevant regions of performance; (**A**) clinically accurate readings; (**B**) results that would lead to benign action or inaction by the user; (**C**) results that would lead to unnecessary corrections; (**D**) results that would lead to inaction when action is necessary; (**E**) results that would lead to treatment that is opposite to what clinical accuracy would call for (Reproduced from ref. Maruo et al. 2006. With permission. Copyright (2006) Society for Applied Spectroscopy)

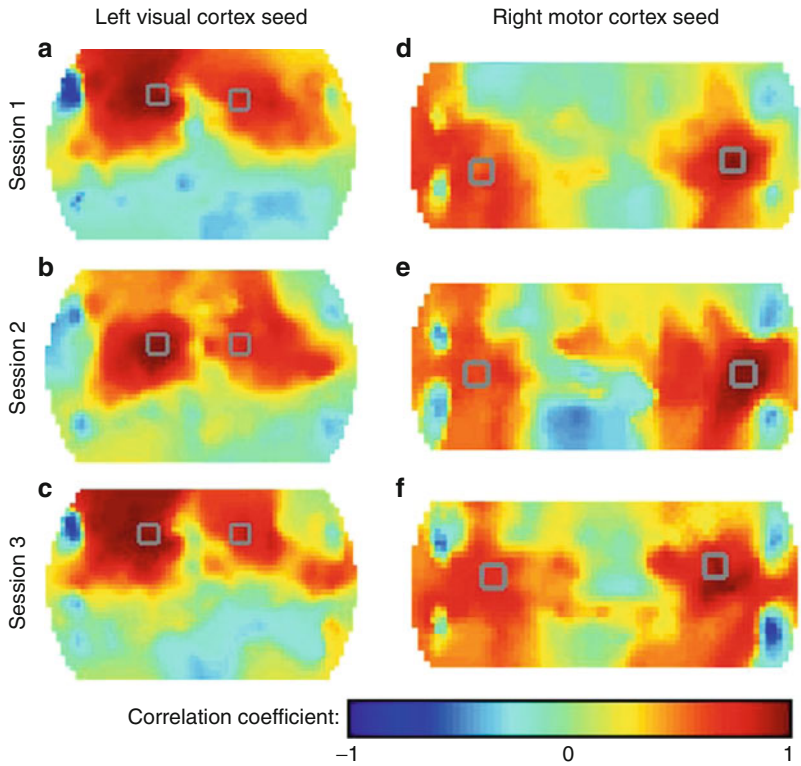
samples and loadings indicate corresponding chemical compositions. Thus, NIR imaging is useful to examine distribution of chemical components within a sample or to probe morphological feature of materials at the molecular level.

In PCA, a data matrix is decomposed in terms of the orthogonal set of dominant factors, i.e., eigenvectors. That is to say, PCA does not reveal the "true" underlying physical sources of data but provides only their linear combination fulfilling the orthogonal constrains. Thus, scores and loadings can take both positive and negative values, such that they usually do not have any direct connection to physical meaning, even though these solutions are based on solid mathematical foundation. An alternative factor analytic technique, which can provide NIR imaging data a more firm link with a physical model, is SMCR (Jiang and Ozaki 2002). In SMCR analysis, one tries to decompose the data matrix **X** into more physically interpretable forms. In a simple bilinear model, one has

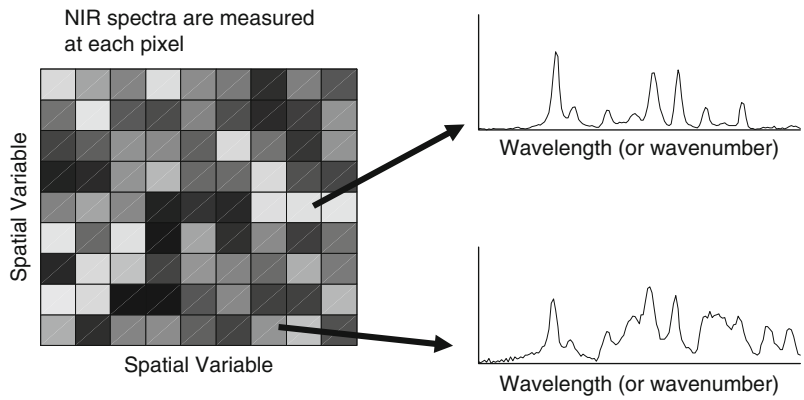
$$\mathbf{X} = \mathbf{C}\mathbf{S}^t + \mathbf{E} \tag{3}$$

where the m by r matrix **C** contains normalized concentration profiles of r individual chemical species, and the n by r matrix **S** contains pure component spectra of the same species. The basic principle of SMCR is to seek a bilinear model that gives the best fit, in the sense of least squares or weighted least squares. In SMCR analysis, **C** and **S** are usually obtained by alternating least-square (ALS) problems unlike PCA. The two matrices, **C** and **S**, complementally represent the entire features broadly distributed in **X**. Namely, concentration profile **C** holds quantitative information chemical components distributed over the samples, and pure component spectra **S** contains actual pure spectra of the corresponding chemical components.

Figure 9 represents an illustrative example of SMCR analysis based on NIR imaging of cellulose tablet (Shinzawa et al. 2009). A set of NIR spectra collected at different positions of a cellulose tablet is decomposed into concentration profiles **C** and pure component spectra **S** by SMCR. Concentration profiles offer quantitative information of each chemical component within the cellulose tablets. Namely, distribution of crystalline and amorphous components of the



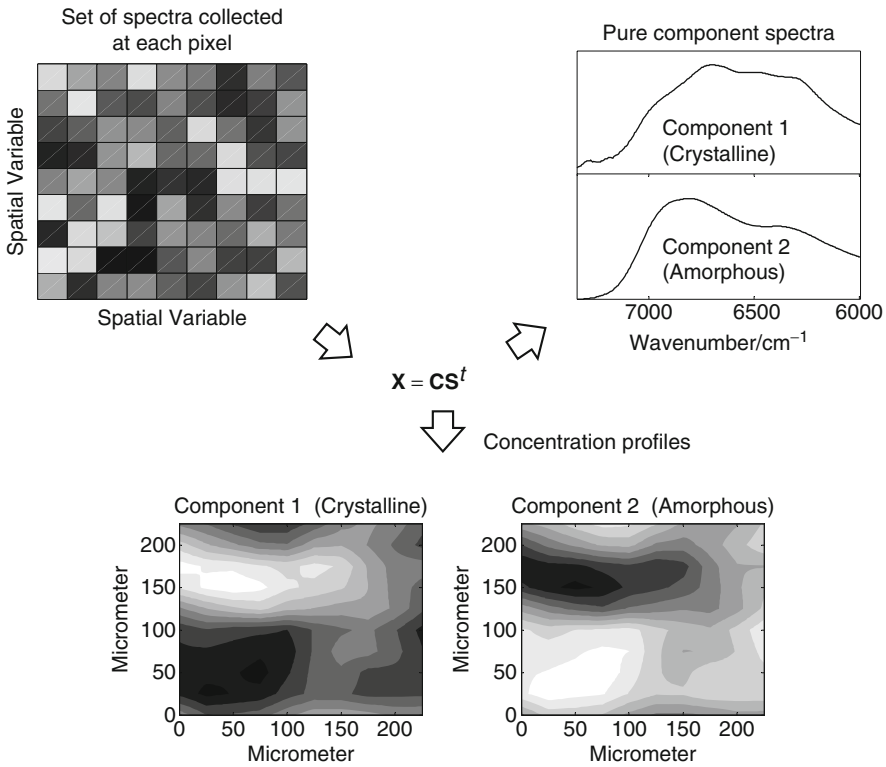
Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 7 Repeatability of DOT over multiple imaging sessions (subject 1, ΔHbR). Seed boxes are shown in gray. Images from different sessions are not co-registered. (**a–c**) Correlation maps within the visual cortex from the left visual cortex seed. (**d–f**) Correlation maps within the motor cortex from the right motor cortex seed. Note the similarity in the patterns in each session (with slight linear translations between days) (Reproduced from ref. White et al. 2009. With permission. Copyright (2009) Elsevier)



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 8 Graphical representation of NIR imaging data

cellulose. Corresponding pure component spectra represent pure NIR spectra of crystalline and amorphous components of cellulose. Thus, it becomes possible to derive key information concerning spatial distribution as well as molecular structure information of the cellulose tablet by SMCR.

The molar absorptivity of water in NIR region becomes relatively small, and NIR light can typically



Near-Infrared Spectroscopy in Biological Molecules and Tissues, Fig. 9 An illustrative example of SMCR study

penetrate much farther into a sample than IR radiation. It, therefore, can be very useful in probing the system containing some level of water content. The ability of NIR imaging offers useful information especially in the field of pharmaceuticals as well as food science since the distribution of specific chemical component within the system can often be of major interest in many product manufacturing processes.

Cross-References

- ▶ Infrared Spectroscopy of Cells, Tissues, and Biofluids
- ▶ Infrared Spectroscopy of Protein Structure
- ▶ Infrared Spectroscopy of Protein Folding, Misfolding and Aggregation
- ▶ Protein Dynamics Studied with 1D and 2D Infrared Spectroscopy
- ▶ Raman Spectroscopic Instrumentation, Experimental Considerations
- ▶ Raman Spectroscopy and Microscopy of Cells and Tissues
- ▶ Surface-Enhanced Raman Spectroscopy for Bioanalytics

References

Ben-Gera I, Norris KH. Direct spectrophotometric determination of fat and moisture in meat products. J Food Sci. 1968;33:64.

Burns DA, Ciurczak EW, editors. Handbook of near-infrared analysis. 3rd ed. CRC Press: Boca Raton; 2007.

Dehghani H, White BR, Zeff BW, Tizzard A, Culver JP. Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography. Appl Opt. 2009;48:D137.

Du Y-P, Kasemsumran S, Jiang J-H, Ozaki Y. In vivo and in vitro near infrared spectroscopic determination of blood glucose and other biomedical components with chemometrics. In: Burns DA, Ciurczak EW, editors. Handbook of near-infrared analysis. 3rd ed. Boca Raton: CRC Press; 2007. p. 1.

Guo SJ, Wang L, Wang W, Fang YX, Wang EK. Bifunctional Au@Pt hybrid nanorods. J Colloid Interface Sci. 2007;315:363.

Hawrysz DJ, Sevick-Muraca EM. Developments toward diagnostic breast cancer imaging using near-infrared optical measurements and fluorescent contrast agents. Neoplasia. 2000;2:388.

Heise HM. Applications of near-infrared spectroscopy in medical sciences. In: Siesler HW, Ozaki Y, Kawata S, Heise HM, editors. Near-infrared spectroscopy. Weinheim: Wiley-VCH; 2002. p. 259.

Hu Y, Zhang J, Sato H, Futami Y, Noda I, Ozaki Y. C-H…O hydrogen bonding and isothermal crystallization kinetics of poly(3-hydroxybutyrate) investigated by near-infrared spectroscopy. Macromol. 2006;39:3841.

Huang X, El-Sayed IH, Qian W, El-Sayed MA. Cancer cell imaging and photothermal therapy in the near-infrared region by using gold nanorods. J Am Chem Soc. 2006;128:2115.

Jiang J-H, Ozaki Y. Self-modeling curve resolution (SMCR): principles, techniques and applications. Appl Spectrosc Rev. 2002;37:321–45.
Jöbsis EF. Noninvasive infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. Science. 1977;198:1264.
Kim F, Song J-H, Yang P. Photochemical synthesis of gold nanorods. J Am Chem Soc. 2002;124(48):14316–7.
Maruo K, Oka M. U. S. Patent 5 957 841, Sept. 28; 1999
Maruo K, Tsurugi M, Chin J, Ota T, Arimoto H, Yamada Y, et al. Noninvasive blood glucose assay using a newly developed near-infrared system. IEEE J Selected Top Quant Electron. 2003; 9:322.
Maruo K, Ota T, Tsurugi M, Nakagawa T, Arimoto H, Tamura M, et al. New methodology to obtain a calibration model for noninvasive near-infrared blood glucose monitoring. Appl Spectrosc. 2006;60:441.
Næs T, Isaksson T, Fearn T, Davies T. A user-friendly guide to multivariate calibration and classification. Chichester: NIR; 2002.
Noda I, Ozaki Y. Two-dimensional correlation spectroscopy-application in vibrational spectroscopy and optical spectroscopy. Chichester: Wiley; 2004.
Ozaki Y, McClure WF, Christy AA, editors. Near-infrared spectroscopy in food science and technology. Hoboken: Wiley-Interscience; 2007.
Ozaki Y, Maruo K, Shinzawa H, Du Y-P, Kasemsumran S. In vivo nondestructive of blood glucose by near-infrared diffuse-reflectance spectroscopy. In: Tuchin VV, editor. Handbook of optical sensing of glucose in biological fluids and tissues. Boca Raton: CRC Press; 2009. p. 205.
Šašić S, Ozaki Y. Raman, infrared, and near-infrared imaging. New York: Wiley; 2010.
Shinzawa H, Awa K, Ozaki Y, Sato H. Near-infrared (NIR) imaging analysis of cellulose tablets by a band position shift. Appl Spectrosc. 2009;63:974.
Siesler HW, Ozaki Y, Kawata S, Heise HM, editors. Near-infrared spectroscopy. Weinheim: Wiley-VCH; 2002.
Tromberg BJ, Shah N, Lanning R, Cerussi A, Espinoza J, Pham T, et al. Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy. Neoplasia. 2000;2:26.
White BR, Snyder AZ, Cohen AL, Petersen SE, Raichle ME, Schlagger BL, et al. Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. Neuroimage. 2009;47:148.
Workman Jr J, Weyer L. Practical guide to near-infrared spectroscopy. Boca Raton: CRC Press; 2007.
Wu Y, Czarnik-Matusewicz B, Murayama K, Ozaki Y. Two-dimensional near-infrared spectroscopy study of human serum albumin in aqueous solutions: using overtones and combination modes to monitor temperature-dependent changes in the secondary structure. J Phys Chem. 2000;104:5840.

NEB (Nudged Elastic Band)

► Carboxypeptidase A – Computational Studies

(Negative) Lateral Tension Profile

► Lipid Bilayer Lateral Pressure Profile

Nerve Growth Factor (NGF)

► Molecular Simulations of NGF Protein

Neural Networks

► Computational Modeling of Neuronal Networks

Neuraminidase

► Influenza Neuraminidase – Computational Studies

Neutron Diffraction

► Diffraction Methods for Studying Transmembrane Pore Formation and Membrane Fusion
► Macromolecular Neutron Diffraction

Neutron Scattering of Membranes

Norbert Kučerka[1,2], Thad Harroun[3] and John Katsaras[1,4]
[1]Canadian Neutron Beam Centre, National Research Council, Chalk River, ON, Canada
[2]Department of Physical Chemistry of Drugs, Comenius University, Bratislava, Slovakia
[3]Department of Physics, Brock University, St. Catharines, ON, Canada
[4]Neutron Sciences Directorate, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Synonyms

Lipid bilayer

Definition

We discuss the use of neutron scattering as a tool for revealing the internal structure of biological membranes.

Basic Characteristics

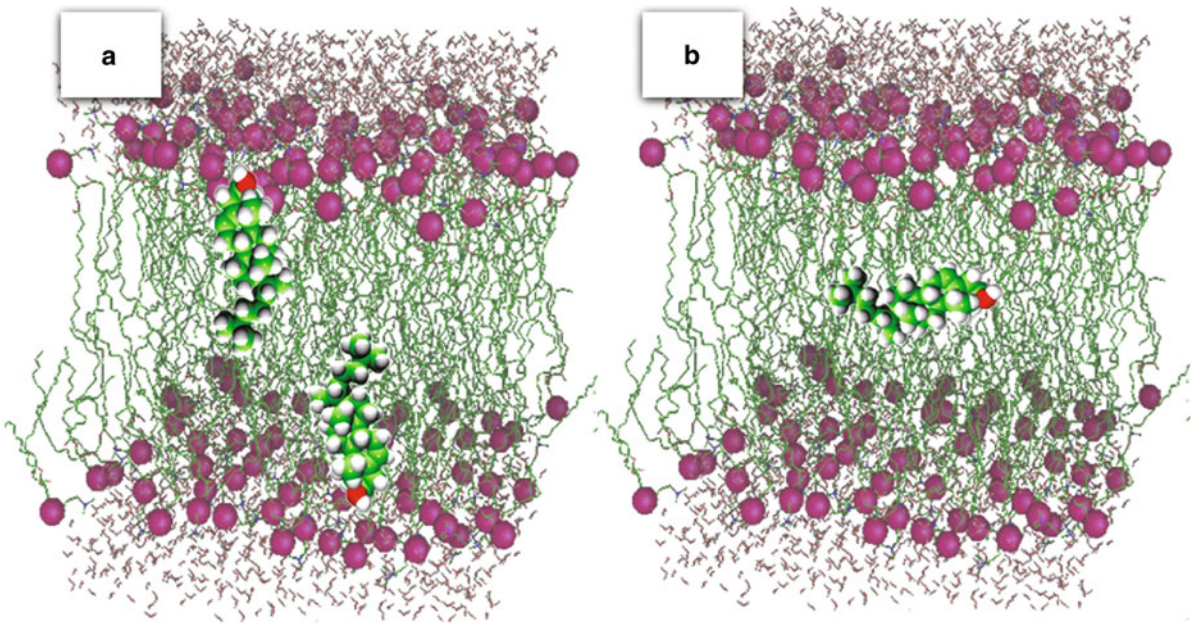
Clifford Shull and Bertram N. Brockhouse were awarded the 1994 Nobel Prize in Physics "*for pioneering contributions to the development of neutron scattering techniques*." Today, neutron scattering offers a powerful suite of techniques for determining atomic and spin structures, and dynamics (i.e., local and collective motions). These techniques are widely used in condensed matter physics, materials science, materials chemistry, polymer science, the biological sciences, and engineering. Because of their intrinsic properties, thermal and cold neutrons are ideally suited for studying a wide range of systems, including biological materials. For example, the interaction with matter of neutrons with energies less than 1 eV reveals atomic resolution detail, as their wavelengths ($\lambda > 0.03$ nm) are of the order of interatomic distances. Similarly, because their energies are comparable to chemical bond energies, and thermal motions, in crystals and liquids, neutrons are extensively used in dynamical studies. Neutron scattering techniques are also complementary to those of X-rays and electrons, especially in studies of biological materials inherently rich in hydrogen and low atomic number elements. In fact, neutron scattering takes advantage of such materials, where contrast variation methods – the selective substitution of hydrogens for deuterons – can either amplify or nullify the scattering from selected parts of a biomolecule (Fitter et al. 2006).

The technique of elastic neutron scattering has been successfully applied to four general areas of biophysics: high-resolution single crystal diffraction; a) low-resolution diffraction from samples with long-range order in one or two dimensions; b) reflectometry from thin films; c) and d) small-angle scattering from randomly oriented structures. The detailed analysis of protein interactions mediated by hydrogen atoms is usually non trivial to determine by X-ray crystallographic studies, as it is difficult to locate hydrogen atoms even when high-resolution X-ray data are available – i.e., X-ray scattering increases with increasing atomic number. On the other hand, high-resolution neutron diffraction studies have accurately revealed the positions of hydrogen atoms, and distinguished nitrogen atoms from similar atomic number elements, such as carbon and oxygen (Schoenborn and Knott 1996).

Neutron Diffraction

With regard to structural biology, the various scattering techniques complement crystallographic studies that, in many cases, require hard-to-obtain, high-quality crystals of macromolecules. Due to the intrinsic disorder present in biomimetic systems – disorder is considered important for the proper function of biological systems – the vast majority of membrane samples do not form perfect, or even near perfect crystals, that are needed to solve structure to atomic resolution. The limited amount of attainable data from such samples is then best described by broad statistical distributions and a membrane's overall characteristics. For example, in the case of model membrane systems (i.e., positionally correlated structures), the position and amplitude of Bragg reflections reveal the membrane's lamellar periodicity and one-dimensional scattering density profile. Further, by changing the systems "contrast" through the exchange of H_2O for D_2O, it is possible to determine the effect of membrane additives, such as cholesterol, on the lipid bilayer and its associated water layer, and the extent that water penetrates into the bilayer. For example, bilayers made of short- and long-chain phospholipids thickened in the presence of cholesterol. This manifestation by cholesterol on bilayer thickness was interpreted by Kučerka et al. as a disorder-order transition induced by cholesterol on the lipid's acyl chains. This effect dominates over the other option of rectifying the hydrophobic mismatch (Kučerka et al. 2009).

Another set of experiments that take advantage of the H/D substitution is selective labeling. The use of bulky labeling probes is ubiquitous in fluorescence light microscopy – and other techniques (e.g., electron spin resonance) – but they can affect a membrane's structural and dynamical properties. For the most part, substituting H for D does not alter the system's chemicophysical properties but does profoundly change the system's contrast. Importantly, neutron data from a system with and without the label differ only in the region of the label. A Fourier difference profile can then directly localize the labeled portion of a molecule (Schoenborn and Knott 1996). Over the years, this technique of isotopic substitution has been extensively used to study the structure of model membranes, proteins, DNA/RNA, and viruses



Neutron Scattering of Membranes, Fig. 1 Schematic of cholesterol's orientation in (**a**) commonly studied lipid bilayers and (**b**) PUFA lipid bilayers (Harroun et al. 2006)

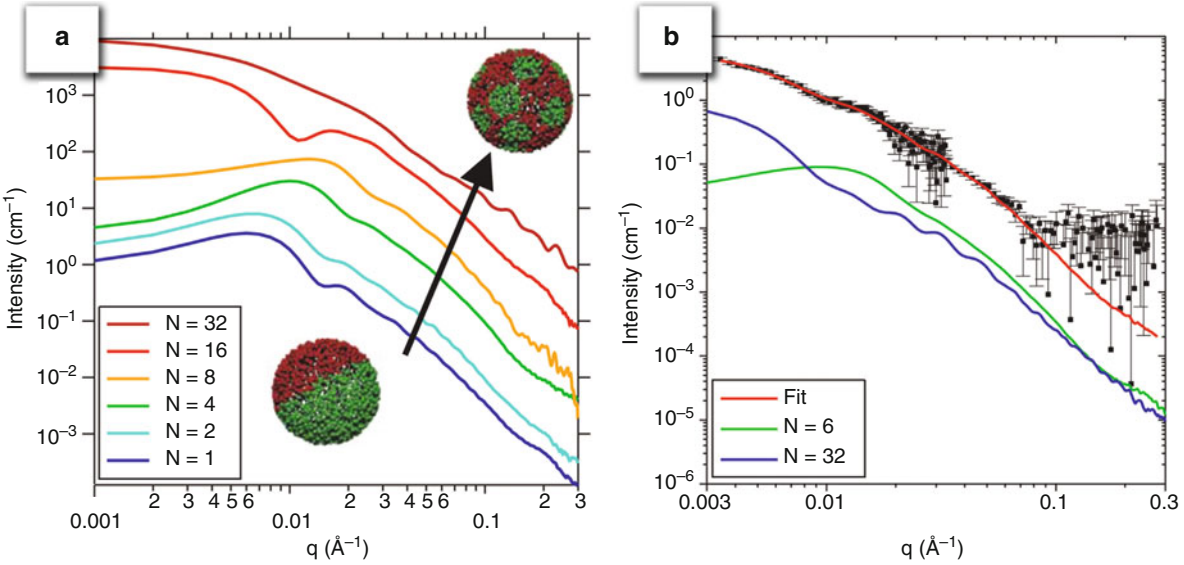
(Pabst et al. 2010). Using this technique, the location of cholesterol in lipids with polyunsaturated fatty acid (PUFA) chains was determined.

Cholesterol is understood to orient parallel to the lipid hydrocarbon chains. In 2006 using deuterated cholesterol, Harroun et al. demonstrated this to be the case in a number of different lipid bilayers. In contrast, however, cholesterol was found to sequester in the center of PUFA lipid bilayers (Fig. 1). Follow-up studies demonstrated the modulated preference of cholesterol for different lipids, clearly highlighting the need for the great diversity of lipids found in biology (Kučerka et al. 2010). For example, in plasma membranes sphingolipids are primarily located in the outer monolayer, whereas unsaturated phospholipids are more abundant in the inner leaflet. It is thus possible that the presence of PUFA in the inner bilayer leaflet may serve to enhance the transfer of cholesterol to the outer layer, potentially modifying membrane function. This mechanism then invokes the necessity for rafts – (i.e., functionalized domains) that could facilitate the biosynthetic pathways of cholesterol and its transport to and from cells.

In order to reconstruct the real-space structure of a membrane, both amplitude and phase are needed. However, in the majority of scattering experiments, only the amplitude of the structure factor can be obtained by experiment – the phase information is lost. In order to recover the phase component of the structure factor, various methods have been developed. The cholesterol example presented above used contrast variation to obtain the phase component of each individual Bragg reflection. In reflectometry and small-angle scattering experiments, real-space solutions rely on fitting the data to a realistic model. In doing so, one must keep in mind the possibility that the data can be fitted equally well by more than one model. In the case of neutron reflectometry studies, Majkrzak et al. (2000) have made good use of polarized neutrons and reference substrates made of Cu, Ni, and Mo to obtain both the amplitude (real part) and phase (imaginary) components of the complex reflection, allowing for the data (reciprocal space) to be directly inverted into a unique scattering density profile (real space).

Small-Angle Neutron Scattering

Small-angle neutron scattering (SANS) is probably the technique most commonly applied to biological materials as it can probe length scales ranging from nanometers to fractions of a micron when ultra SANS is used. In the case of unilamellar vesicles (ULVs)



Neutron Scattering of Membranes, Fig. 2 (**a**) Predicted scattering from polydisperse ULV containing multiple domains (N = 1–32). The number of domains increases on going from the bottom to the top curves. *Curves* are shifted on the vertical axis to facilitate viewing. (**b**) Fits to a 1:1:1 DOPC to DPPC to cholesterol ULV system corresponding to the superposition of signals from ULVs with single and multiple domains (Adapted from Pencer et al. 2005)

(hollow spherical structures consisting of single lipid bilayer), SANS can evaluate both the ULV's size, and the thickness of the bilayer that forms the ULV. This technique thus allows for the accurate structural determination of ULVs that have been fabricated for delivering drugs and imaging diseased tissue, introducing genetic material to living cells, and enhancing the efficacy of various medical imaging techniques (Nejat 2009). Although ULVs exist in solution, and their isotropic structural dimensions are typically characterized by low-resolution data, much insight into a system can be achieved through the previously discussed contrast variation techniques. For example, Pencer et al. (2005), using coarse-grained models of heterogeneous ULVs, demonstrated SANS' potential to detect and distinguish between lateral segregation in membranes. This was done by exploiting the unique sensitivity of neutron scattering to differences between H and D atoms (Fig. 2), as was discussed previously. Most recently, a new and innovative advance in scattering theory and experiment enables neutron data to be analyzed in such a manner as to provide not only structural information in the direction perpendicular to the plane of the bilayer, but also information regarding a membrane's in-plane organization, e.g., lateral lipid segregation (Iglič 2010).

Cross-References

- Chemical Diversity of Lipids
- Fluorescence and FRET in Membranes
- Lipid Bilayer Asymmetry
- Lipid Trafficking in Cells
- Membrane Proteins: Structure and Organization
- Protein Fluorescent Dye Labeling
- Small Angle Neutron Scattering (SANS) Software
- Supported Lipid Bilayers
- Vibrational Spectroscopy with Neutrons

References

Fitter J, Gutberlet T, Katsaras J. Neutron scattering in biology: techniques and applications. Berlin: Springer; 2006.

Harroun TA, Katsaras J, Wassall SR. Cholesterol hydroxyl group is found to reside in the center of a polyunsaturated lipid membrane. Biochemistry. 2006;45:1227–33.

Iglič A. Advances in planar lipid bilayers and liposomes, vol. 12. Burlington: Elsevier; 2010.

Kučerka N, Nieh MP, Pencer J, Sachs JN, Katsaras J. What determines the thickness of a biological membrane. Gen Physiol Biophys. 2009;28:117–25.

Kučerka N, Marquardt D, Harroun TA, Nieh MP, Wassall SR, de Jong D, Schafer L, Marrink SJ, Katsaras J. Cholesterol in bilayers with PUFA chains: doping with DMPC or POPC

results in sterol reorientation and membrane-domain formation. Biochemistry. 2010;49:7485–93.
Majkrzak CF, Berk NF, Silin V, Meuse CW. Experimental demonstration of phase determination in neutron reflectometry by variation of the surrounding media. Physica B. 2000;283:248–52.
Nejat D. Methods in enzymology: liposomes, Part G, vol. 465. Burlington: Elsevier; 2009.
Pabst G, Kučerka N, Nieh MP, Rheinstädter MC, Katsaras J. Applications of neutron and X-ray scattering to the study of biologically relevant model membranes. Chem Phys Lipids. 2010;163:460–79.
Pencer J, Mills T, Anghel V, Krueger S, Epand RM, Katsaras J. Detection of submicron-sized raft-like domains in membranes by small-angle neutron scattering. Eur Phys J E. 2005;18:447–58.
Schoenborn BP, Knott RB. Basic life sciences: neutrons in biology, vol. 64. New York: Plenum; 1996.

New Protein Folds

▶ Protein Structure: Potential Folds

N-Glycan Analysis

▶ Mass Spectrometry of *N*-Linked Carbohydrates and Glycoproteins

Nicotinamide Adenine Dinucleotide (NAD)

Mathias Ziegler
Department of Molecular Biology, University of Bergen, Bergen, Norway

Synonym

Diphosphopyridine nucleotide (DPN)

Definition

NAD is a dinucleotide composed of two mononucleotides (adenosine 5′-monophosphate, AMP, and nicotinamide mononucleotide, NMN), which are joined through their 5′ phosphate groups (Fig. 1). The molecule is present in all living cells and reversibly converted between its oxidized (NAD^+) and reduced (NADH) forms acting as a coenzyme in many metabolic hydride transfer (redox) reactions (Fig. 2). NAD^+ serves also important functions in signal transduction pathways. A phosphorylated derivative of NAD, NADP, is also present in all organisms, exhibits similar physicochemical properties, but apparently does not have a prominent role in signaling.

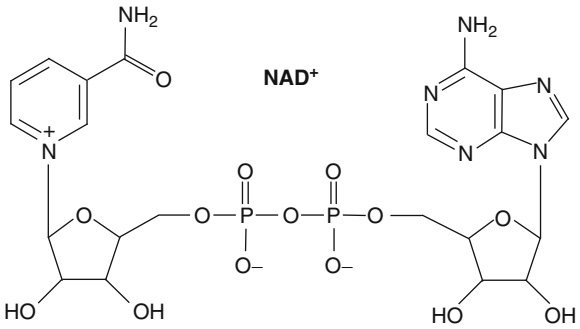
Physicochemical Characteristics of NAD

In its naturally occurring, biologically active form, the nicotinamide moiety of NAD is attached to the anomeric carbon of the ADP-ribosyl moiety in the β-conformation (Fig. 1). Both NAD^+ and NADH exhibit an absorption maximum at 259 nm owing to the presence of the adenine base ($\varepsilon = \sim 16{,}200\ M^{-1}cm^{-1}$ for NAD^+). Due to the reduced nicotinamide ring, NADH has an additional absorption maximum at 340 nm ($\varepsilon = \sim 6{,}200\ M^{-1}cm^{-1}$). This property is widely used to measure enzymatic reactions which require NAD as coenzyme. NADH is a strong reducing agent, the midpoint ▶ redox potential of the NAD+/NADH redox pair being –0.32 V.

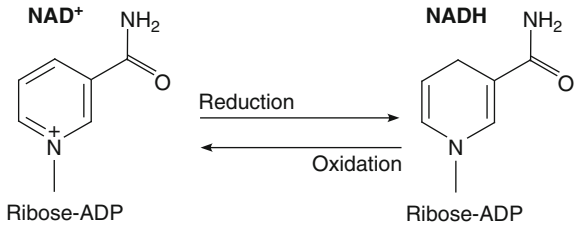
Unlike NAD^+, NADH is fluorescent with an emission maximum at ~460 nm (excitation at 340 nm). The fluorescence lifetime (~0.4 ns in aqueous solution) increases significantly, when the molecule binds to proteins. This property has been exploited to estimate concentrations of free vs. protein-bound NADH or the redox state (NADH/NAD^+ ratio) in living cells by means of fluorescence microscopy, but can also be used to determine binding constants for isolated proteins.

Biological Functions of NAD

NAD plays two distinct biological roles. Similar to ATP, it is a key molecule in both energy transduction and signaling pathways. In bioenergetics, NAD is a coenzyme which serves as reversible electron carrier in a multitude of redox reactions. In signaling processes, however, only NAD^+ is involved. NAD is regarded as an intracellular molecule. However, in higher organisms low amounts are present in extracellular fluids (e.g., blood plasma), thereby providing substrate for extracellular NAD^+-mediated signaling events.



Nicotinamide Adenine Dinucleotide (NAD), Fig. 1 Chemical structure of NAD^+



Nicotinamide Adenine Dinucleotide (NAD), Fig. 2 Function of NAD as redox carrier. *Top*: In metabolic reactions, NAD^+ can take up two electrons (and a proton), which are added to the nicotinamide moiety. *Bottom*: The reversible oxidation of an alcohol to an aldehyde illustrates the typical redox function of NAD in metabolism

Redox Functions of NAD

Most enzymes that use NAD as a coenzyme are oxidoreductases. That is, they oxidize a substrate (extract electrons) and at the same time reduce NAD^+ to NADH (Fig. 2; note that the designations oxidation and reduction in the Figure refer to the changes in NAD, not the substrate). Generally, these reactions are reversible. In the major catabolic routes, electrons from nutrients (such as glucose and fatty acids) are transferred onto NAD^+ to yield NADH. Eventually, the electrons are donated to the respiratory chain (to drive ATP synthesis) and thereby NADH is reoxidized to NAD^+ via ▶ NADH:ubiquinone oxidoreductase (complex I). The redox functions of NAD and the metabolic enzymes catalyzing the corresponding reactions are comprehensively covered in all standard textbooks of biochemistry.

Signaling Functions of NAD^+

Over many years, the biological role of NAD had been thought to be limited to metabolic redox reactions. It has now become clear that NAD^+ is also a key element of various regulatory pathways (Berger et al. 2004). Three distinct types of reactions are known in which NAD^+ is cleaved to yield nicotinamide, while the remaining ADP-ribosyl moiety is converted into signaling molecules or attached to proteins. First, NAD^+ glycohydrolases (the best characterized representative in humans being CD38) produce ADP-ribose and cyclic ADP-ribose (cADPR). These molecules activate calcium channels in the plasma membrane or the endoplasmic reticulum, respectively, thereby leading to an increase of the cytosolic calcium concentration (Malavasi et al. 2008). Second, ADP-ribosyl transferases catalyze the modification of intra- and extracellular proteins causing changes in the biological activities of these proteins. Cell-surface ADP-ribosyl transferases are predominantly involved in immunological functions (Seman et al. 2004). The intracellular enzymes have a broad range of functions including regulation of metabolic enzymes and control of nuclear processes. Poly-ADP-ribose polymerases (PARPs) generate polymers of ADP-ribose and attach them to target proteins. This modification has a critical impact on important nuclear events and has been demonstrated only in higher eukaryotes. The major mammalian PARP enzyme, PARP-1, is catalytically activated by DNA strand breaks. Therefore, extensive DNA damage can cause depletion of cellular NAD and cell death owing to rapid degradation of NAD^+ by PARP-1 in response to genotoxic assaults (Krishnakumar and Kraus 2010). Third, a class of protein deacetylases, the sirtuins, require NAD^+ as a co-substrate. The acetyl group of the acetylated protein is transferred onto ADP-ribose, thereby yielding unmodified protein and O-acetyl ADP-ribose, another potential signaling derivative of NAD^+. This mechanism has been identified as a major link between metabolism (by sensing the NAD concentration and redox state) and cellular/organismal fate in higher eukaryotes. The deacetylation of nuclear proteins (such as histones and transcription factors) by sirtuins is highly selective and regulates, for example, key metabolic processes, the biological clock and life span (Imai and Guarente 2010). Generally, NAD^+-dependent protein deacetylation is ancient and also found in bacteria. In higher eukaryotes, it also controls the activity of key enzymes of metabolic pathways compartmentalized in mitochondria.

Subcellular Compartmentation and Biosynthesis of NAD

In eukaryotes, the majority of NAD resides in the mitochondria and the nucleo-cytoplasm (the nucleus and cytosol are thought to be freely exchangeable), as well as in chloroplasts in plants. This is in agreement with the known metabolic and signaling pathways that require NAD.

Cellular NAD synthesis (Berger et al. 2004; Magni et al. 2008) relies on different precursors. Most cells produce their NAD from the vitamin B3 precursors nicotinamide and nicotinic acid, collectively termed "niacin." In mammals, vitamin B3 deficiency causes the disease pellagra. In some bacteria, NAD can be generated from aspartate, whereas in higher organisms a pathway of tryptophan degradation generates an NAD precursor, quinolinic acid. The ribosides of nicotinamide and nicotinic acid can also be utilized. Once they have entered cells, the precursors are converted to the corresponding mononucleotides (e.g., NMN from nicotinamide). The mononucleotide is then converted to the dinucleotide by adding the AMP moiety from ATP. The corresponding enzymes, NMN/NAMN adenylyltransferases are essential as they are required in all routes of NAD synthesis, regardless of the extracellular precursor used. When NAD is synthesized via a nicotinic acid derivative, the nicotinic acid dinucleotide is eventually amidated by NAD synthetase.

Cross-References

- ▶ Electron Transfer Cofactors
- ▶ NADH-Ubiquinone Oxidoreductase (Complex I)
- ▶ Redox Potential

References

Berger F, Ramírez-Hernández MH, Ziegler M. The new life of a centenarian: signalling functions of NAD(P). Trends Biochem Sci. 2004;29:111–8.

Imai S, Guarente L. Ten years of NAD-dependent SIR2 family deacetylases: implications for metabolic diseases. Trends Pharmacol Sci. 2010;31:212–20.

Krishnakumar R, Kraus WL. The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. Mol Cell. 2010;39:8–24.

Magni G, Orsomando G, Raffelli N, Ruggieri S. Enzymology of mammalian NAD metabolism in health and disease. Front Biosci. 2008;13:6135–54.

Malavasi F, Deaglio S, Funaro A, Ferrero E, Horenstein AL, Ortolan E, Vaisitti T, Aydin S. Evolution and function of the ADP ribosyl cyclase/CD38 gene family in physiology and pathology. Physiol Rev. 2008;88:841–886.

Seman M, Adriouch S, Haag F, Koch-Nolte F. Ecto-ADP-ribosyltransferases (ARTs): emerging actors in cell communication and signaling. Curr Med Chem. 2004;11:857–72.

NIR Imaging

▶ Near-Infrared Spectroscopy in Biological Molecules and Tissues

NIR Spectroscopy

▶ Near-Infrared Spectroscopy in Biological Molecules and Tissues

Nitric Oxide Synthase - Computational Studies

Marc Willem van der Kamp
Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, UK

Synonyms

NOS

Definition

A heme-containing enzyme that catalyzes the production of nitric oxide from L-arginine.

Basic Characteristics

Nitric oxide (NO) is a cellular signaling molecule that has an important role in many biological processes in a wide range of organisms. Production of NO by nitric

oxide synthases (NOSs) is achieved by catalysis of the two-step oxidation of L-arginine to L-citrulline. As in ▶ cytochrome P450 enzymes, the sulfur of a cysteine residue coordinates to the heme iron. In addition, NOSs feature a conserved tryptophan that forms a hydrogen bond with the sulfur atom. It has been proposed that the tryptophan modulates the electronic properties of the heme and the heme-bound ligands. DFT QM/MM calculations have been used to investigate the role of the tryptophan in the product state, i.e., NO bound to the heme iron (Fernandez et al. 2005). Several studies on the reaction mechanism of NOS have also been performed, with small "cluster" models with quantum mechanical (QM) electronic structure calculations and with ▶ QM/MM methods on larger molecular models incorporating the whole enzyme environment (de Visser 2009). These studies focus on the initial reaction step, the oxidation of L-arginine to *N*-hydroxo-arginine. It was suggested that due to the unique and highly polar L-arginine substrate, the catalytic cycle for the first step in NOS diverges from cytochrome P450 enzymes after formation of Compound I, and Compound I is not the active oxidant, in contrast to cytochrome P450 enzymes. Subsequent QM/MM calculations considered three different possible reaction mechanisms for the first step in NOS. Only the pathway that featured a singly protonated ferric-peroxy complex, combined with formation of a cation-radical species shared by the tetrahydrobiopterin cofactor and the deprotonated arginine substrate, was deemed energetically reasonable and could account for experimental data on the reaction (Cho et al. 2009). The detailed catalytic mechanism of this enzyme is not yet entirely resolved, but it is clear that computational molecular simulation and modeling provides a crucial means of proposing and testing proposed reaction pathways, because unambiguous experimental observation of the species involved is nigh impossible. A study into the nature of the tetrahydrobiopterin cofactor bound to NOS is an example of how experimental (e.g., EPR) and computational studies can be combined to help elucidate further details (Stoll et al. 2010).

Cross-References

▶ Cytochrome P450 – Computational Studies
▶ QM/MM Methods

References

Cho K-B, Carvajal MA, et al. First half-reaction mechanism of nitric oxide synthase: the role of proton and oxygen coupled electron transfer in the reaction by quantum mechanics/molecular mechanics. J Phys Chem B. 2009;113(1):336–46.
de Visser SP. Density functional theory (DFT) and combined quantum mechanical/molecular mechanics (QM/MM) studies on the oxygen activation step in nitric oxide synthase enzymes. Biochem Soc Trans. 2009;37:373–7.
Fernandez ML, Marti MA, et al. Proximal effects in the modulation of nitric oxide synthase reactivity: a QM-MM study. J Biol Inorg Chem. 2005;10(6):595–604.
Stoll S, Nejatyjahromy Y, et al. Nitric oxide synthase stabilizes the tetrahydrobiopterin cofactor radical by controlling its protonation state. J Am Chem Soc. 2010;132(33):11812–23.

NMR

▶ Assignment of ^{19}F Resonances in Protein Solution State NMR Studies
▶ Fragment Screen
▶ In-Cell NMR
▶ Multidimensional NMR Spectroscopy
▶ NMR in Drug Discovery – Introduction
▶ NMR Studies of Macromolecular Interactions – Introduction
▶ Nuclear Overhauser Effect
▶ Protein NMR Resonance Assignment
▶ Pulsed Field Gradient NMR
▶ SAR by NMR
▶ Stereo-Array Isotope Labeling (SAIL) Method
▶ Structure Determination by NMR: Overview
▶ Triple Resonance NMR

NMR in Drug Discovery – Introduction

Steven Halouska and Robert Powers
Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE, USA

Synonyms

Ligand-binding interactions; NMR; Target and hit validation techniques

Definition

Processes by which drugs are discovered and designed with analytical techniques such as NMR.

Introduction

Drug discovery is an ongoing and challenging process that is fraught with failure, but the successes have had profound impacts on human health. Fundamentally, drug discovery is a multidisciplinary endeavor requiring, among others, bioinformatics and computational chemistry, cell biology, medicinal chemistry, enzymology, high-throughput screening (HTS), molecular biology, protein chemistry, genomics and other "omics" technologies, and structural biology. Drug discovery is also an iterative procedure comprising four stages: target identification and validation, lead discovery, lead optimization, and clinical trials (Betz et al. 2006). In general, the drug discovery process starts by identifying a protein target associated with drug resistance, the progression or pathology of a disease, or the virulence of an organism. Usually, the protein of interest is chosen because it is perturbed when the system is in a disease state, or the target is a critical or essential protein, for instance, required for microbial or tumor survivability. Part of the target identification process also requires understanding the protein's biological or cellular function and potentially obtaining its three-dimensional (3D) structure. The process of finding an initial set of antagonists or agonists is typically accomplished using HTS and a library composed of hundreds of thousands to millions of compounds. Commonly, three to five chemical classes from the HTS results are selected as chemical leads for further optimization. The selection of chemical leads is based on activity, chemical novelty and patentability, drug-like characteristics, synthetic accessibility, and diversity. Different analogs of the chemical leads are synthesized to increase affinity, selectivity, and potency, while reducing toxicity issues. Effectively, the goal is to evolve the chemical leads into drug candidates. Overall, the drug discovery process is extremely time consuming and expensive, requiring approximately 12 years and costing >$800 million dollars. This is due, in part, to the complexity of the biological system, limited knowledge about the protein targets, and the inherent difficulty of converting an inhibitor to a drug, where only 1 in 5,000 chemical leads becomes a drug (Light and Warburton 2011). Therefore, technological advancements and new methodologies that improve the success rate are desperately needed.

Nuclear magnetic resonance (NMR) is a versatile analytical tool that is used in all phases of the drug discovery process prior to initiating clinical trials (Betz et al. 2006). NMR is routinely used to characterize the structures of both small molecules and large biomolecules. More importantly, NMR is often used to address an initial and important question in the drug discovery process: does the ligand selectively and specifically bind to the protein target in a biologically relevant manner? This is an essential step in evaluating the results obtained from HTS, in silico screens, or fragment-based screens, and for evolving chemical leads into drug candidates. NMR is uniquely suited for detecting protein-ligand interactions, for identifying the ligand-binding site, for calculating dissociation constants (K_D), for determining a 3D structure of the protein-ligand complex, and for monitoring in vivo activity. Furthermore, this information can be obtained using HTS-NMR, where hundreds to thousands of potential drugs can be screened. In addition, cryoprobe technology, high-field magnets (>800 MHz), and enhancement in NMR pulse sequences and HTS technology have substantially increased throughput, signal-to-noise, and the detection of low levels of proteins or weak binding ligands. Thus, HTS-NMR has been applied to metabolomics, and extended to larger chemical libraries and larger protein targets (>100 kDa), while simultaneously minimizing resources (experimental time, material) (Pellecchia et al. 2008).

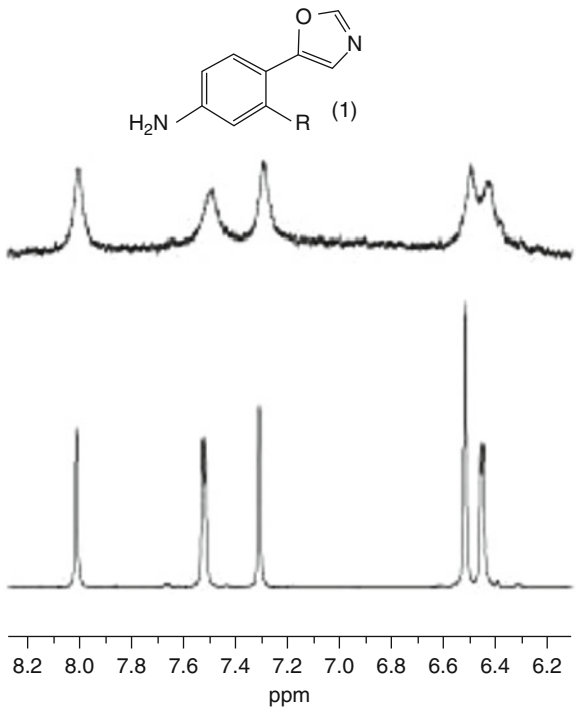
An inherent value of HTS-NMR, with respect to drug discovery, is its nearly universal application (Widmer and Jahnke 2004). HTS-NMR does not require any target-specific setup because the screen is independent of the protein's function. Thus, proteins recently identified by genomics can be screened immediately, as long as an adequate amount of labeled or unlabeled material is available; and the protein falls within the allowable molecular weight (MW) range for the specific NMR experiment. In general, the protein is overexpressed and isotopically labeled with ^{15}N and/or ^{13}C for structure determination or HTS affinity screens. *Escherichia coli* and *Bacillus subtillis* are the most commonly used expression systems because their robustness permit high cell density and expression rates (Heller and

Kessler 2001). The 3D protein structure is critical to the drug discovery process because it enables the evaluation of the biological relevance of the potential drug candidates. This is accomplished by determining if the ligand-binding site corresponds to the protein's active site or functional epitope. Similarly, ligand selectivity and specificity is determined based on the ligand binding to other binding sites or proteins. Also, the structure is routinely used to annotate functionally uncharacterized proteins based on structural homology to proteins with an assigned function. The NMR assignments and 3D protein structures are typically determined using a standard set of 3D triple resonance and NOESY experiments (Kanelis et al. 2001).

HTS-NMR is used to identify chemical leads, and validate HTS and in silico screening results. SAR by NMR (structure-activity relationship by NMR) was the first illustration of HTS-NMR. Since then, there have been numerous NMR experiments designed for the sole purpose of being used in an NMR ligand affinity screen (Pellecchia et al. 2002). Some popular one-dimensional NMR experiments (or ligand-focused screens), include saturation transfer difference (STD), WaterLOGSY (water ligand observation by gradient spectroscopy), NOE pumping, and diffusion edited NMR (Widmer and Jahnke 2004). The primary goal of these experiments is to use an observable change in the NMR spectrum of the chemical lead (peak intensity, peak width, chemical shift, etc.) to identify a productive binding interaction. Advantages of these techniques include rapid data collection (<5 min.), minimal sample requirements (<10 μM), and no need for isotopically labeled protein. But in general, these techniques do not provide any information on the ligand-binding site. Alternatively, protein-focused screens, SAR by NMR, along with Multi-Step NMR, MS/NMR, NMR-SOLVE, and SHAPES, among others, are used to identify ligand-binding sites in addition to identify binders. Importantly, these methods are typically used as part of a fragment-based screen that requires determining the binding proximity of two or more ligands. The goal is to chemically link multiple fragments to achieve a corresponding enhancement in binding affinity. But, these screens do require isotopically labeled proteins and significantly longer data acquisition times. HTS-NMR can also be used to evaluate in vivo drug toxicity and efficacy by using NMR metabolomics protocols. NMR is used to compare metabolic profiles before and after drug treatment to determine the biological impact of the drug candidate.

Ligand-Focused HTS-NMR

The relatively low sensitivity of NMR is an important issue in its application to HTS. Correspondingly, HTS-NMR requires longer data acquisition times and sample requirements compared to standard high-throughput screens. Importantly, screening hundreds of thousands to millions of compounds, while routine for HTS, is completely impractical for NMR. Thus, designing NMR experiments to minimize both instrument time and sample requirements is critical to improve throughput; and the primary goal of ligand-focused 1D NMR screens. Simply, a ligand-focused screen monitors changes in the 1D ^{1}H NMR spectrum of the ligand resulting from the addition of the protein target. The ligand is typically in large excess (>10-fold) of the protein. In general, the 1D ^{1}H-NMR methods exploit large differences in physical properties between small molecules and large biomolecules; and correspondingly between the free and bound states of the ligand. For example, a schematic representation in Fig. 1 illustrates the effects of differing T_2 relaxation times on ligand binding (Fejzo et al. 1999). Specifically, small molecules undergo rapid Brownian motion and have correspondingly slow T_2 relaxation times. Conversely, proteins undergo slow Brownian motion and have fast T_2 relaxation times. Therefore, the NMR spectrum of a small molecule consists of narrower peaks compared to a protein NMR spectrum. As a small molecule binds a protein, it will inherit the physical properties of the protein. As a result, the NMR peaks of the small molecule will broaden proportionally to the difference in line widths and fraction of bound ligand. Ligand-focused 1D NMR screens have some distinct advantages. The NMR experiments are relatively fast (<5 min). Only small quantities, ~1–10 μM of unlabeled protein are required per NMR experiment. Additionally, the ligands can be screened as mixtures of ~4 to >20 different compounds with a concentration of ~100 μM per ligand. Assembling large mixtures can be challenging since it may be difficult to find compatible sets of compounds that maintain solubility and stability while avoiding cross-reactivity.



NMR in Drug Discovery – Introduction, Fig. 1 1D ^{1}H spectra of 1 mM free ligand (*bottom*) in the presence of 100 μM IMP dehydrogenase (*top*). The significant line broadening indicates a ligand-protein interaction. (Reprinted with permission from (Fejzo et al. 1999). ©1999 by Elsevier)

The appearance of the 1D ^{1}H NMR spectrum also reflects on how rapidly the ligand exchanges between the bound and free state. Additionally, the appearance of the 1D ^{1}H NMR spectrum is influenced by the ligand's affinity to the protein target (Pellecchia et al. 2002):

$$K_D = \frac{K_{off}}{K_{on}} = \frac{[L]_F[P]_F}{[PL]} \tag{1}$$

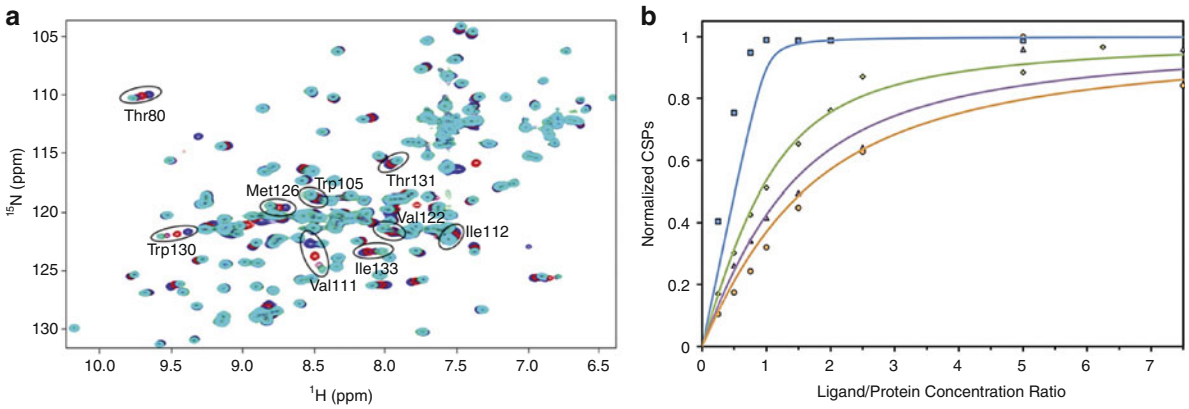
where K_D is the dissociation constant, K_{off} is the rate constant for dissociation, K_{on} is the rate constant for formation, $[L]_F$ is the concentration of free ligand, $[P]_F$ is the concentration of the free target protein, $[PL]$ is the concentration of the protein-ligand complex. Tight binders usually have slow exchange rates and a K_D of less than 10 nM. In this case, a separate signal will be observed for both the bound and free ligands, where the bound signal is usually broadened beyond detections. Correspondingly, a binding event is indicated by a decrease in the intensity of the ligand NMR spectrum. If the ligand binds weakly ($K_D > 10$ μM), the exchange between the bound and free state is fast and the observed spectrum consists of a single signal, whose chemical shift and line width is the weighted average between the bound and free state. For intermediate exchange, the ligand peak broadening is typically enhanced. Thus, in the case of fast exchange, binding affinities can be estimated from HTS-NMR using 1D ^{1}H NMR spectra (Shortridge et al. 2008). The K_D can be calculated by measuring the peak height and line width using the following equations:

$$K_D = \left[\left(\frac{c[P]_T}{B_{\text{single}}} - c[P]_T\right) - [L]_T\right]$$
$$\text{where } c = \frac{\upsilon_B}{\upsilon_F} - 1 \text{ and } B_{\text{single}} = 1 - \frac{I_B}{I_F} \tag{2}$$

where $[P]_T$ and $[L]_T$ are the total protein and ligand concentrations, respectively, v_B and v_F are the line width for the bound and free state, respectively, and I_B and I_F are the peak height for the bound and free state, respectively. In addition to fast exchange, the approach also assumes that there is only a single binding site. These assumptions are reasonable in the early stages of drug discovery because initial chemical leads will generally have weak binding affinities. Obtaining binding affinities from HTS-NMR have also been described for STD experiments, for the displacement of known low-affinity inhibitors, or for chemical shift changes using ^{19}F NMR with ^{19}F-containing compounds, among other NMR experiments.

Target-Focused HTS-NMR

One important advantage of HTS-NMR is the ability to detect weak binding ligands, which is a critical factor in fragment-based screens. Most target-focused approaches to HTS-NMR rely on chemical shift perturbations in the protein NMR spectrum, where the protein is either ^{15}N or ^{13}C labeled. The ligand is unlabeled and undetected by the target-based NMR experiments. Therefore, the ligand is still typically in excess (≥5-fold) to maximize the observation of weak binding ligands ($K_D > 1$ mM). Two-dimensional (2D) ^{1}H-^{15}N-heteronuclear single quantum coherence (^{1}H-^{15}N-HSQC) experiments are commonly used for HTS-NMR. Other common HTS-NMR experiments include the 2D ^{1}H-^{15}N-TROSY, 2D ^{1}H-^{13}C HSQC, and recent derivatives that enable very



NMR in Drug Discovery – Introduction, Fig. 2 (**a**) Overlay of the 2D ^{1}H-^{15}N-HSQC spectra of *B. subtilis* protein YndB with an increasing concentration of chalcone ranging from 0 μM (*blue*) to 160 μM (*cyan*). (**b**) NMR titration data for *trans*-chalcone (*blue*), flavanone (*green*), flavone (*purple*), and flavonol (*orange*). The normalized chemical shift perturbations (CSP) for the nine most perturbed residues are plotted versus the protein-ligand concentration ratios. The theoretical curve displayed for *trans*-chalcone corresponds to a K_D of 1 μM and represents the upper limit for the K_D. The measure K_D values are ≤1 μM (*trans*-chalcone), 32 ± 3 μM (flavanone), 62 ± 9 μM (flavone), and 86 ± 16 μM (flavonol). (Reprinted with permission from (Stark et al. 2010). © 2010 by John Wiley and Sons)

rapid data collection (e.g., SOFAST). A peak in the 2D ^{1}H-^{15}N-HSQC spectrum is correlated to each amino acid residue (except proline) present in the protein sequence. Each peak is also associated with both a ^{1}H and ^{15}N chemical shift resulting from the bonded NH (amide) group. The relative location (chemical shifts) of each peak in the 2D ^{1}H-^{15}N-HSQC spectrum is strongly dependent on the local environment for each amino acid residue in the protein structure. Thus, the addition of a ligand known to bind the protein results in local environmental changes for amino acids in direct contact with the bound ligand, and for residues that undergo an induced structural change. Correspondingly, chemical shift changes are observed for a select number of peaks that can be mapped onto the protein structure to identify the ligand-binding site. A unique spatial clustering of residues that incur a chemical shift perturbation (CSP) indicates a specific binding interaction and the location of the ligand-binding site. Conversely, a random distribution of residues with CSPs on the proteins surface indicates a nonspecific interaction. Alternatively, an observation that a majority of residues incur a CSP or experience a decrease in peak intensity probably infers a ligand-induced aggregation, precipitation, or denaturation, which typically eliminates the compound's consideration as a chemical lead. Importantly, the observation that two or more ligands (fragments) share proximal, but not identical binding sites, based on CSP mapping, presents an opportunity to chemically link the fragments. The linked fragments are expected to have an enhanced affinity and represent a starting point for further optimization.

The addition of an increasing concentration of unlabeled ligand will cause a progressive change in the protein's chemical shifts if the ligand-binding affinity is weak and in fast exchange (Roberts 2000). Thus, a binding affinity can be determined by following CSPs as a function of ligand concentration:

$$CSP_{obs} = \frac{(K_D + [L] + [P]) - \sqrt{(K_D + [L] + [P])^2 - (4[L][P])}}{2[P]} \quad (3)$$

where CSP_{obs} is the 2D ^{1}H-^{15}N HSQC chemical shift perturbations, $[P]$ is the protein concentrations, $[L]$ is the ligand concentration, and K_D is the dissociation constant. Conversely, a tightly bound ligand in slow exchange will result in two sets of NMR peaks in the HSQC spectrum for the bound and free state. The ratio of intensities for the bound and free states can be used to determine the fraction bound and measure a K_D from a standard binding isotherm. Figure 2 illustrates the application of a 2D ^{1}H-^{15}N HSQC titration experiment to measure chemical shift perturbations and calculate a K_D (Stark et al. 2010).

HTS-NMR and Lead Optimization

It has been estimated that the number of possible chemical compounds is on the order of 10^{60} (Bohacek et al. 1996). Clearly, obtaining and experimentally screening such an immense compound library is impossible. Instead, a screening library for HTS-NMR usually consists of only a few hundred to several thousand compounds. Commonly, the HTS-NMR screening library is a set of chemical leads from HTS. Obviously, the quality of the chemical leads that emerge from HTS is dependent on the quality of the original HTS library (Lipinski 2004). A major source of populating a screening library is from prior HTS lead optimization. Unfortunately, this often leads to higher MW and more lipophilic compounds due to exploiting hydrophobic interactions that increase potency. The result is a compound library with reduced pharmacokinetics or pharmacodynamics properties. In essence, HTS-NMR provides an invaluable function by further validating the HTS chemical leads and experimentally eliminating compounds with poor drug-like properties. HTS-NMR can experimentally verify acceptable physical behavior, such as good solubility, stability, purity, and a lack of aggregation or micelle-like behavior. HTS-NMR can also confirm that the chemical lead specifically binds the protein target in a biologically relevant manner. Additionally, selection of compounds can also abide by the Lipinski's rule of 5, which is based on predicted molecular properties important for drug-like behavior. Of course, strictly following Lipinski's rule of 5 may eliminate potentially interesting and novel leads. Instead, a combination of HTS-NMR and Lipinski's rule of 5 provides the most flexibility. Additionally, ligand efficiency (LE) provides a better approach to prioritize chemical leads instead of traditional measures of relative activity (IC_{50}, or K_i). LE is defined as:

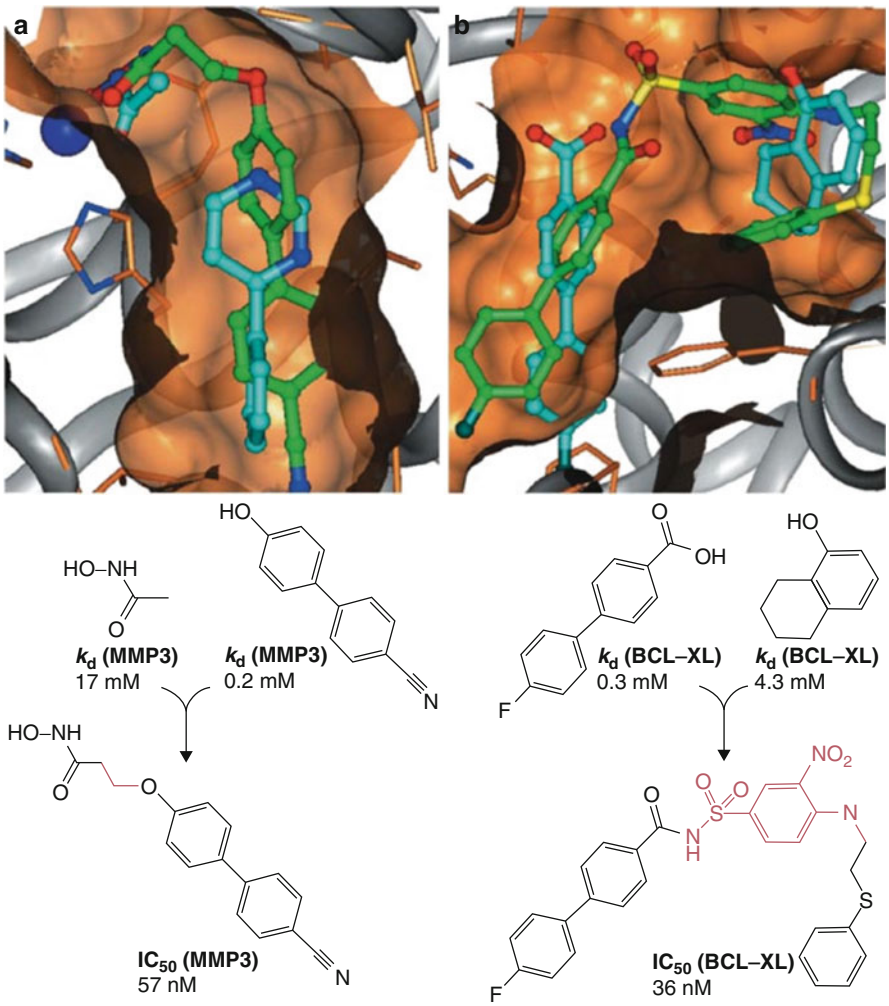
$$LE = \frac{\Delta G}{N} = \frac{-RT \ln K_i}{N} \tag{4}$$

where ΔG is the Gibbs free energy, K_i is the dissociation constant for inhibitor binding, and N is the number of non-hydrogen atoms.

Based on LE, lower MW (<300 Da) chemical leads provide a better opportunity to successfully evolve a compound into a drug (Zartler and Shapiro 2005). Similarly, a library of low MW compounds provides a more efficient coverage of structural space because the number of low MW compounds is significantly reduced. The efficiency of low MW compounds for drug discovery is the basis of fragment-based screening. Correspondingly, HTS-NMR using a fragment-based library has become an integral part of the drug discovery process. A fragment-based screen is complementary to standard HTS because the two chemical libraries typically cover different regions of structural space. As a result, chemical leads from a fragment-based screen tend to be novel compounds and distinct from HTS leads. Also, chemical leads from a fragment-based screen tend to be more drug-like, where the lower MW provides more opportunities for optimization and maintaining high ligand efficiency. Again, the primary goal behind a fragment-based screen is to identify two or more fragments that bind in proximal sites on the target protein. Chemically linking these fragments in a manner that maintains the original binding orientations will result in a new lead molecule with enhanced affinity. The process is illustrated in Fig. 3. HTS-NMR is well-suited to screening a fragment-based library because the low MW drug-like fragments are expected to have weak binding affinities (K_D > μM – mM) that are readily detectable by NMR. Patentability may be a potential challenge with fragment-based libraries. Multiple pharmaceutical and biotechnology companies will use very similar screening libraries since chemical fragments of known drugs are a well-defined set of compounds. Thus, it will not be surprising to obtain similar chemical leads when fragment-based libraries are screened against an identical protein target.

NMR Functional Annotation

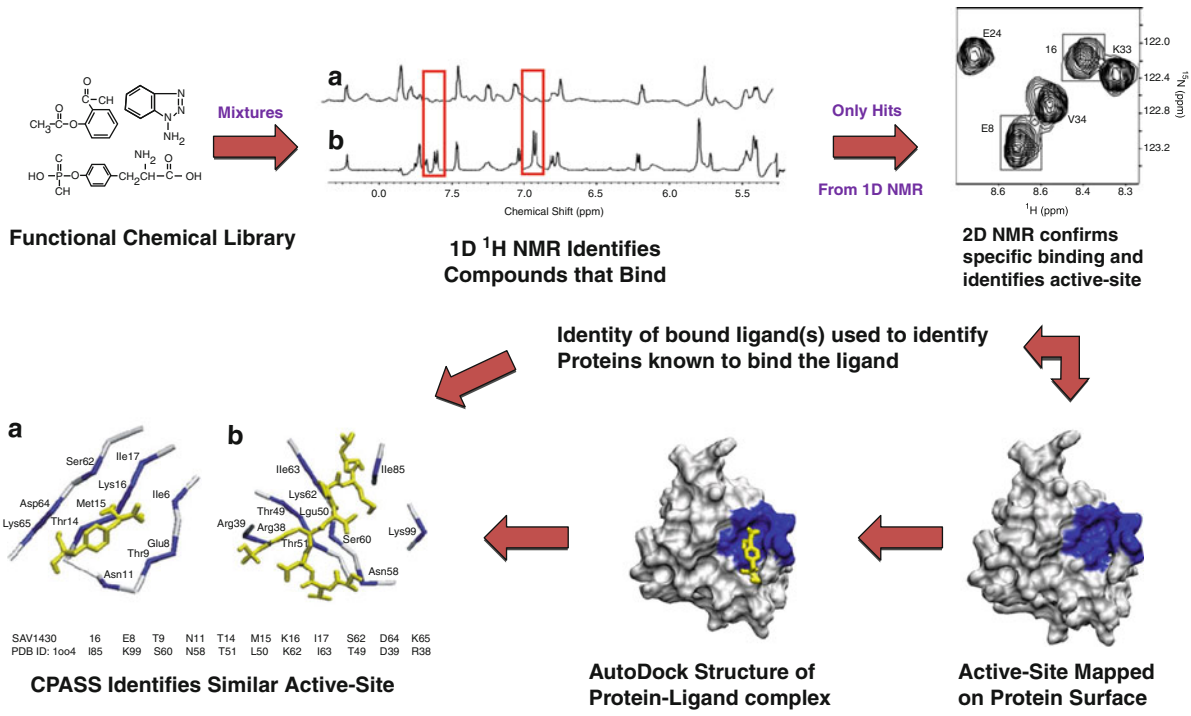
Functional annotation of unknown protein targets plays an important role in the drug discovery process because it allows for identification of novel targets that may be involved in human disease. Functional annotation is primarily obtained by sequence or structure homology. The majority of the ~19 million proteins identified to date (UniPortKB; http://www.uniprot.org) are functionally annotated through sequence or structure homology. The transfer of functional information is vastly more common than the direct acquisition of experimental data. But at least 40% of prokaryotic and eukaryotic genomes are explicitly annotated as "hypothetical" or "uncharacterized" proteins because of a lack of



NMR in Drug Discovery – Introduction, Fig. 3 Applications of SAR by NMR method for fragment design showing ligand binding for (**a**) matrix metalloproteinase 3 (**b**) Bcl-xL. In each case at the *top*, the identified fragment leads are shown with *cyan* carbons, whereas the linked compounds are denoted with *green* carbon atoms. (Reprinted with permission from (Hajduk and Greer 2007). © 2007 by Nature Publishing Group)

homology to functionally annotated proteins. HTS-NMR can be used to assist in the functional classification of unannotated proteins when sequence and structure homology fails.

A fundamental component to understanding a protein's function is derived from its interaction partners. The identity of ligands and the location of binding sites can be leveraged to infer a function in a manner similar to sequence and structure homology. This is based on the observation that amino acids in an active site or functional epitope are evolutionarily more stable than the rest of the protein. Effectively, these residues are required to maintain function, where proteins that share similar ligand-binding sites are predicted to be functional homologs. FAST-NMR (**F**unctional **A**nnotation **S**creening **T**echnology using NMR) uses HTS-NMR to identify ligands and binding sites for functionally uncharacterized proteins. FAST-NMR then exploits a structure and sequence similarity to a ligand-binding site from an annotated protein to infer a function for an uncharacterized protein (Powers et al. 2008). A schematic representation of the FAST-NMR protocol is illustrated in Fig. 4. Unlike traditional HTS-NMR, the screening library contains only biologically active molecules that are active against a defined protein target. In essence, the library is used as chemical probes to identify structural homologs of the natural ligand and the location of ligand-binding sites. FAST-NMR uses a tiered approach to HTS-NMR that includes a 1D ^{1}H NMR ligand-focused screen followed by a 2D ^{1}H-^{15}N-HSQC target-focused CSP screen. The ligand-focused screen is used to quickly identify binders, where only the hits are further screened in the target-focused CSP screen. The 2D ^{1}H-^{15}N-HSQC CSP screen is used to verify binders, identify the



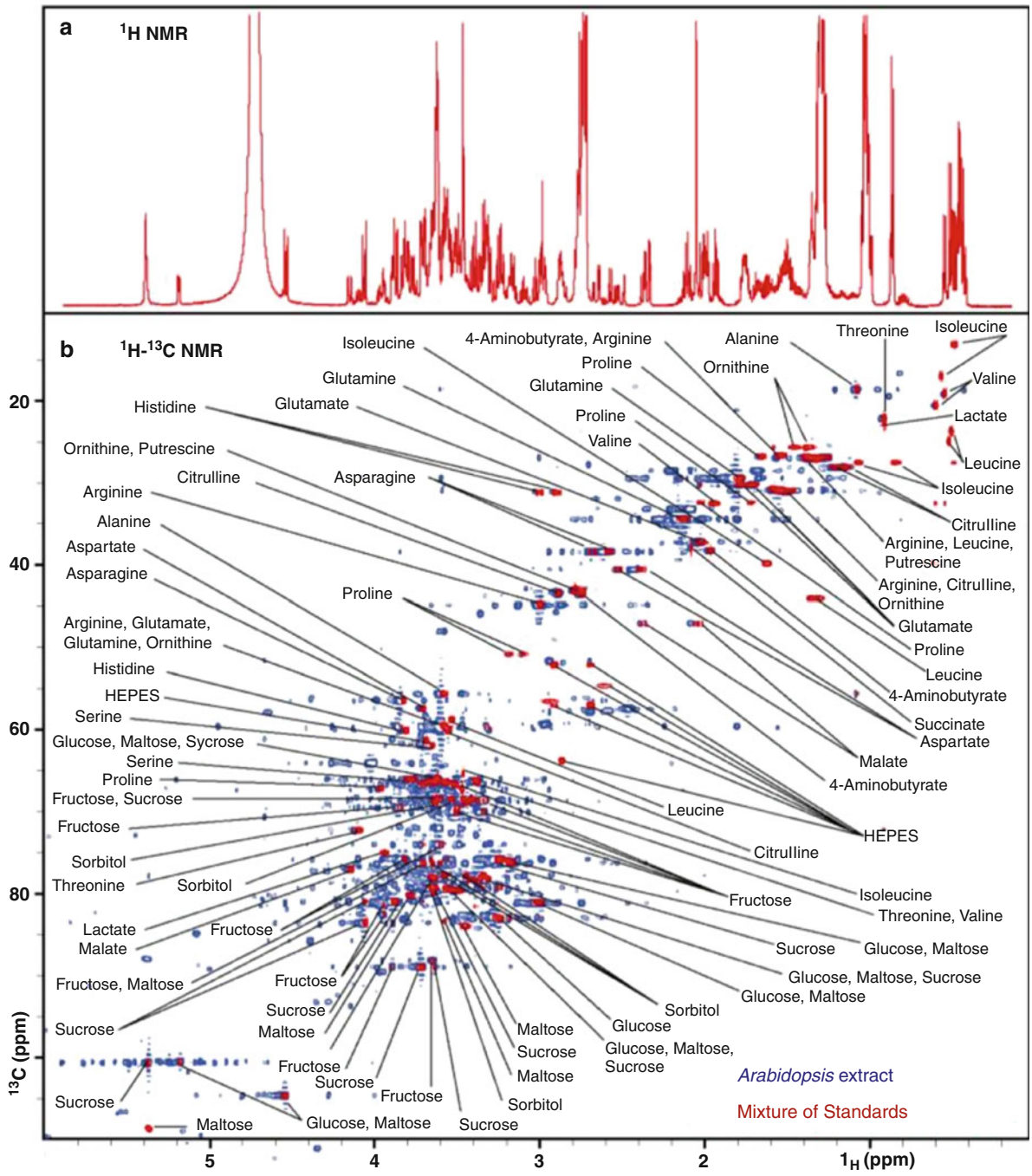
NMR in Drug Discovery – Introduction, Fig. 4 Schematic diagram for FAST-NMR. (Reprinted with permission from (Powers et al. 2008). © 2008 by Elsevier)

ligand-binding site, and generate a protein-ligand co-structure. This structure is then used as an input for CPASS (**C**omparison of **P**rotein **A**ctive-**S**ite **S**tructure) to identify homologous ligand-binding sites from functionally annotated proteins. The CPASS database contains all the unique ligand-binding sites present in the RCSB PBD.

NMR Metabolomics and In Vivo Drug Activity

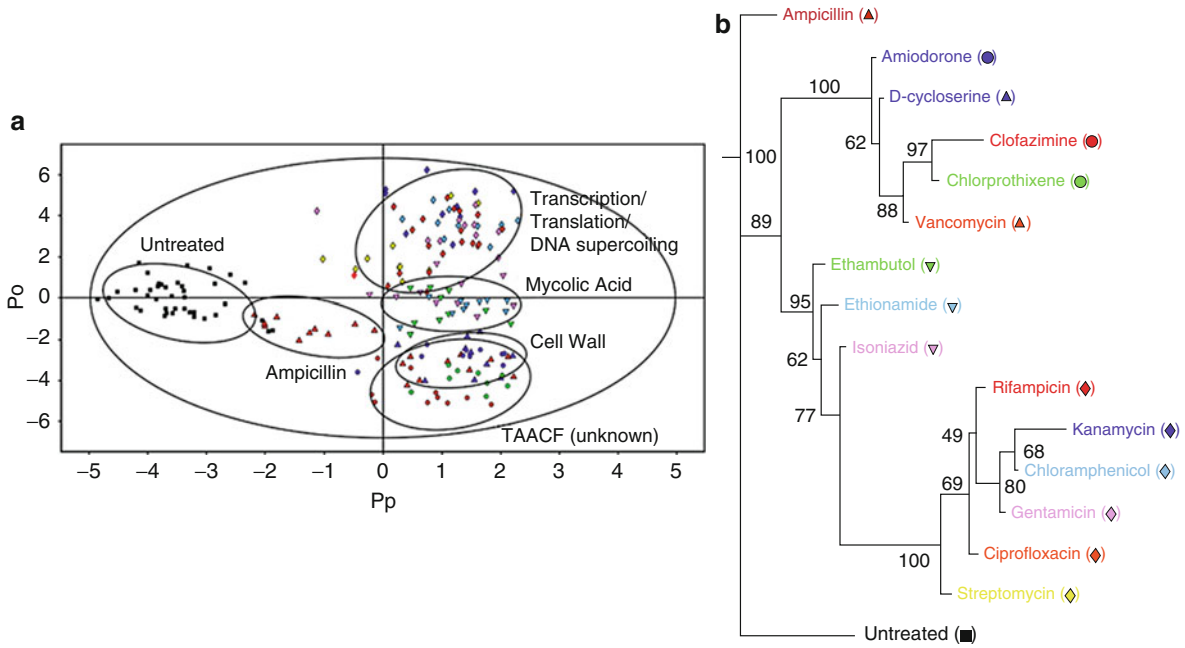
One means of analyzing the state of a biological system is achieved by monitoring the metabolome – all the metabolites present in a cell, tissue, organ, or organism. Correspondingly, metabolomics is the study of the changes in the concentration and identity of these metabolites that results from environmental or genetic stress from a disease state or drug treatment. In essence, the metabolome provides a chemical fingerprint or signature that uniquely defines the state of the system. Metabolomics has an intrinsic advantage over genomics and proteomics analysis since observed changes in the metabolome are directly coupled with changes in protein activity and cell function. Metabolites are the end product of all cellular processes, and are a direct result of enzymatic and protein activity. NMR metabolomics can assist drug discovery by characterizing drug efficacy, selectivity, and toxicity (Powers 2009). Importantly, NMR metabolomics may identify and eliminate troublesome compounds prior to initiating a clinical trial. Thus, HTS-NMR can also be used to follow the impact of chemical leads on metabolites from cellular extracts (bacteria, human cell lines, tumors, etc.) or biofluids (urine, serum, saliva, etc.). Common NMR experiments used for metabolomics includes: 1D ^{1}H NMR, 2D ^{1}H-^{1}H TOCSY, and 2D ^{1}H-^{13}C HSQC.

Metabolites are typically detected using 1D ^{1}H-NMR, where peak heights relative to an internal standard are used to determine metabolite concentrations. Because of the low sensitivity of NMR, only the most abundant ($\sim$ >1 μM) metabolites are observed. Also, a metabolomic sample is a complex heterogeneous mixture containing numerous metabolites. The limited spectral resolution and severe peak overlap in a 1D ^{1}H-NMR spectrum makes unambiguous metabolite



NMR in Drug Discovery – Introduction, Fig. 5 (**a**) One-dimensional ^{1}H-NMR spectrum of an equimolar mixture of the 26 small-molecule standards. (**b**) Two-dimensional ^{1}H-^{13}C-HSQC NMR spectra of the same synthetic mixture (*red*) overlaid onto a spectrum of aqueous whole-plant extract from *Arabidopsis thaliana* (*blue*). (Reprinted with permission from (Lewis et al. 2007). © 2007 by American Chemical Society)

assignment particularly challenging. Furthermore, since the metabolome has not been completely characterized, erroneous assignments may result. Alternatively, 2D ^{1}H-^{13}C HSQC spectra are commonly used to identify metabolites. The higher spectral resolution, the larger ^{13}C chemical shift dispersion, and the observation of



NMR in Drug Discovery – Introduction, Fig. 6 (**a**) 2D OPLS-DA scores plot demonstrating the clustering pattern for 12 antibiotics with known biological targets and three compounds of unknown in vivo activity: untreated *Mycobacterium smegmatis* cells, chloramphenicol, ciprofloxacin, gentamicin, kanamycin, rifampicin, streptomycin, ethambutol, ethionamide, isoniazid, ampicillin, D-cycloserine, vancomycin, amiodorone, chlorprothixene, and clofazimine treated *M. smegmatis* cells. The symbols correspond with the coloring scheme and labeled symbols indicated on the tree diagram in (**b**). The ellipses correspond to the 95% confidence limits from a normal distribution for each cluster. The untreated *M. smegmatis* cells (*black square*) was designated the control class, and the remainder of the cells were designated as treated. The OPLS-DA used one predictive component and six orthogonal components to yield a R^2X of 0.715, R^2Y of 0.803, and Q^2 of 0.671. (**b**) Metabolomics tree diagram determined from the OPLS-DA scores plot. The coloring scheme and associated symbol for each compound in the tree diagram correlates with colored symbols in the OPLS-DA scores plot. The bootstrap numbers for each node are indicated on the tree diagram. (Reprinted with permission from (Halouska et al. 2012). © 2011 by American Chemical Society)

chemical shifts for each C-H pair significantly improve the accuracy of metabolite assignment. The metabolomic sample does need to be supplemented with a ^{13}C-labeled metabolite to avoid long acquisition times because of the low natural abundance of ^{13}C (1.1%). Correspondingly, the only metabolites observed in the 2D ^{1}H-^{13}C HSQC spectrum are derived from the ^{13}C-labeled metabolite. Also, because of variations in J-coupling constants, and T_1 and T_2 relaxation times, direct quantitation of a standard HSQC spectrum to calculate metabolite concentrations is unreliable. Nevertheless, modification to the HSQC NMR pulse sequence has produced quantitative HSQC experiments (e.g., $HSQC_0$, QQ-HSQC, etc.). Metabolites can also be quantified by using the fast quantification by NMR method (FMQ), where a series of standard metabolites over a range of concentrations are screened (Fig. 5). A biological extract can then be compared against the standards to determine the metabolite concentrations (Lewis et al. 2007).

Several databases have been developed that enable metabolite identification from experimental chemical shift data. The databases contain NMR and mass spectroscopy data for thousands of metabolites from a broad range of organisms. Therefore, metabolite identification needs to be verified with KEGG and MetaCyc metabolic pathway databases to verify the metabolite is actually present in a specific organism (Powers 2009). The NMR metabolomic databases include: Human Metabalome Database (HMDB), Madison Metabolomics Database (MMCD), and Platform for Riken Metabolomics (PRIMe).

A common application of NMR metabolomics data is to monitor global differences between metabolic samples instead of following changes in specific

metabolites. This approach involves the use of multivariate statistical analysis and pattern recognition. Multivariate statistical techniques such as principal component analysis (PCA) and orthogonal partial least-squares discriminant analysis (OPLS-DA) are routinely employed to capture global perturbations in the metabolome (Stoyanova and Brown 2002; Bylesjo et al. 2006). PCA is an unsupervised method that reduces multivariate dataset to a single point and projects the major variations in the dataset into a few axes called principal components. In this way, spectral variations are captured in a model that can be easily visualized. As a result, similar metabolomic data will cluster closely together and separately from other distinct metabolomes. Conversely, OPLS-DA is a supervised method used to determine the variations within the data set that is correlated to the classification label (treatment, control, disease state). If the variations within the data set (noise) are not correlated with the classification label, then the variations are filtered out, resulting in a single latent vector. This is analogous to PCA, where PCA captures all the intrinsic variations within the data set. Applying PCA and OPLS-DA to NMR metabolomics data sets is extremely valuable for classifying drug activity and toxicity in a living system based on relative clustering patterns (Fig. 6). This also allows for determining the in vivo mechanism of action for a new chemical lead or drug (Halouska et al. 2012).

Summary

Nuclear magnetic resonance (NMR) is a versatile analytical tool with a wide range of applications that may improve the success rate of drug discovery. NMR uses ligand-focused or target-focused screening techniques to discover chemical leads with a large range of affinities. Correspondingly, NMR is ideally suited for fragment-based screening, which complements standard high-throughput screens and expands the coverage of compound structural space. NMR can also be employed to identify the location of ligand-binding sites and determine a protein-ligand co-structure. This is essential information for linking fragments to enhance binding affinity and to evolve chemical leads into drug candidates. NMR ligand affinity screens can also be applied to assist in the functional annotation of uncharacterized proteins to identify new therapeutic targets. Finally, NMR metabolomics strategies can identify the in vivo mechanism of action, and determine the in vivo efficacy and toxicity for chemical leads and help alleviate the many failures encountered in clinical trials.

Cross-References

- BMRB
- Flow NMR
- Fragment Screen
- Mass Spectrometry: Application to Protein-Ligand Interactions
- Metabolomics
- Multidimensional NMR Spectroscopy
- NMR
- NMR-based Structural Proteomics
- Protein NMR – Introduction
- Protein Secondary Structure Prediction in 2012
- SAR by NMR
- SOFAST HMQC
- Total Correlation Spectroscopy (TOCSY) in NMR Protein Structure Determination
- Triple Resonance NMR
- TROSY

References

Betz M, Saxena K, et al. Biomolecular NMR: a chaperone to drug discovery. Curr Opin Chem Biol. 2006;10(3):219–25.

Bohacek RS, McMartin C, et al. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996;16(1):3–50.

Bylesjo M, Rantalainen M, et al. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. J Chemom. 2006;20(8–10):341–51.

Fejzo J, Lepre CA, et al. The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. Chem Biol. 1999;6(10):755–69.

Hadjduk J, Greer J. A decade of fragment-based drug design: stratigeic advances and leassons learned. Nat Rev Drug Discov. 2007;6:211–19.

Halouska S, Fenton RJ, et al. Predicting the in vivo mechanism of action for drug leads using NMR metabolomics. ACS Chem Biol. 2012;7(1):166–71.

Heller M, Kessler H. NMR spectroscopy in drug design. Pure Appl Chem. 2001;73(9):1429–36.

Kanelis V, Forman-Kay JD, et al. Multidimensional NMR methods for protein structure determination. IUBMB Life. 2001;52(6):291–302.

Lewis IA, Schommer SC, et al. Method for determining molar concentrations of metabolites in complex solutions from

two-dimensional 1H-13C NMR spectra. Anal Chem. 2007;79(24):9385–90.
Light DW, Warburton R. Demythologizing the high costs of pharmaceutical research. Biosocieties. 2011;6(1):34–50.
Lipinski C. Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol. 2004;1(4):337–41.
Pellecchia M, Sem DS, et al. NMR in drug discovery. Nat Rev Drug Discov. 2002;1(3):211–9.
Pellecchia M, Bertini I, et al. Perspectives on NMR in drug discovery: a technique comes of age. Nat Rev Drug Discov. 2008;7(9):738–45.
Powers R. NMR metabolomics and drug discovery. Magn Reson Chem MRC. 2009;47(Suppl 1):S2–11.
Powers R, Mercier KA, et al. The application of FAST-NMR for the identification of novel drug discovery targets. Drug Discov Today. 2008;13(3–4):172–9.
Roberts GCK. Applications of NMR in drug discovery. Drug Discov Today. 2000;5(6):230–40.
Shortridge MD, Hage DS, et al. Estimating protein-ligand binding affinity using high-throughput screening by NMR. J Comb Chem. 2008;10(6):948–58.
Stark JL, Mercier KA, et al. Solution structure and function of YndB, an AHSA1 protein from *Bacillus subtilis*. Proteins. 2010;78(16):3328–40.
Stoyanova R, Brown TR. NMR spectral quantitation by principal component analysis. III. A generalized procedure for determination of lineshape variations. J Magn Reson. 2002; 154(2):163–75.
Widmer H, Jahnke W. Protein NMR in biomedical research. Cell Mol Life Sci CMLS. 2004;61(5):580–99.
Zartler ER, Shapiro MJ. Fragonomics: fragment-based drug discovery. Curr Opin Chem Biol. 2005;9(4):366–70.

NMR Ligand Affinity Screen

► Fragment Screen

NMR Methods for Kinetic Analysis

Igor L. Barsukov
School of Biological Sciences,
University of Liverpool, Liverpool, UK

Definition

NMR kinetic analysis aims at deriving reaction rates from the characteristics of the NMR spectra.

Basic Characteristics

The unique ability of NMR to resolve and monitor signals of individual nuclei in different chemical environment makes it a powerful method for kinetic analysis. NMR spectra are highly sensitive to the environment, and no modification of the molecules is normally required to detect the changes due to the interaction and kinetic effects. Its application, however, is limited by the low sensitivity and relatively long time of data acquisition. These restrictions are particularly severe for monitoring nonequilibrium processes where reaction course is followed after system perturbation or mixing, making these experiments generally suitable only for reaction times longer than seconds. Faster rates can be obtained by the NMR analysis of dynamic equilibrium where no net changes occur, although the reaction is proceeding at a detectable rate.

Nonequilibrium Analysis

The fastest way to monitor a reaction time course is through a sequential collection of 1D NMR spectra. A single scan acquisition takes a fraction of a second to complete and, even on high-field spectrometers, sufficient signal-to-noise ratio requires at least tens of scans at sub-millimolar concentrations. This restricts the delay between each point of the time course to seconds or longer. Additionally, there is a dead time before the experiment caused by sample insertion, temperature equilibration, and magnetic field adjustment. In a normal NMR setup, the dead time is measured in minutes, although can be dramatically reduced with the use of flow probes. Due to these considerations, even in an optimized setup at high concentrations, the measurable rate constants are less than 0.1 s^{-1}. For enzymatic reactions, the rate of substrate conversion can be easily adjusted by reducing enzyme concentration, making these reactions amenable to the nonequilibrium NMR analysis.

With a slow enough reaction rates, NMR provides a wealth of information for characterizing reaction pathways and intermediates. Signals of small molecules are usually well resolved in a 1D spectrum, and individual reaction products can be identified and followed selectively by monitoring corresponding resonances. If 1D resolution is not sufficient, a time series of 2D COSY or TOCSY spectra can be collected to improve resonance separation and help with product identification. Each of the 2D spectra normally takes tens of minutes to acquire, reducing the reaction rates available to the analysis. For ^{13}C or ^{15}N enriched molecules, corresponding HSQC spectra are used to enhance selectivity and resolution. Routinely, a sequential series

of [^{1}H,^{15}N]-HSQC spectra are used to determine hydrogen exchange rates in proteins (Cavanagh et al. 2007).

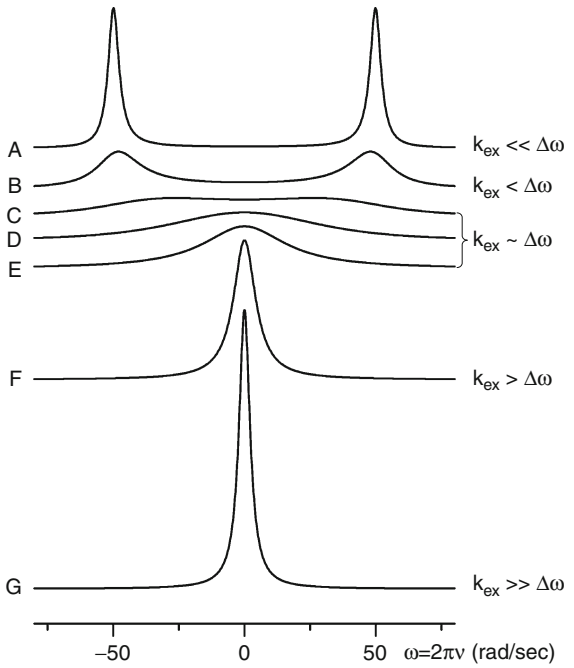
Reactions that can be stopped by changing reaction conditions make a special case where faster rates can be monitored by NMR. The reaction timescale is then determined by the characteristics of the **quenched flow** rather than NMR limitations. Once the reaction is stopped, the intermediate concentrations and compositions of the products can be analyzed by 1D or 2D NMR spectroscopy, providing detailed characterization at each time point. Such analysis is often used for study of hydrogen exchange in protein folding, where exchange is quenched by a pH jump.

Chemical Exchange

In a dynamic equilibrium, the rate of spin exchange between different environments has a strong effect on the NMR spectra. The exchange process is modeled using McConnell modifications of the Bloch equations, described in details for two spins in (Rule and Hitchens 2006; Levitt 2001). Matrix-based multi-spin exchanged equations are presented in (Cavanagh et al. 2007). A general case of multi-spin high-order reaction is analyzed in (Ernst et al. 1987). Briefly, in a two site exchange A ↔ B with the rate constants k_1 and k_{-1} for the forward and reverse reactions, the spectral changes caused by the exchange depend on the relationship between the exchange rate $k_{ex} = k_1 + k_{-1}$ and the frequency difference $\Delta\omega = \omega_A - \omega_B$ for the spin in the two states. The frequency difference can vary for different spins of the same molecule, resulting in different exchange contributions into specific resonances. Five exchange regimes are normally identified that lead to specific effects on the NMR spectra (Rule and Hitchens 2006; Levitt 2001) illustrated in Fig. 1:

- Very slow, $k_{ex} \ll \Delta\omega$, two resonances are observed for a single spin with negligible effect on the line width
- Slow, $k_{ex} < \Delta\omega$, two resonances are observed broadened by the exchange
- Intermediate, $k_{ex} \sim \Delta\omega$, single very broad resonance is observed with a complicated lineshape
- Fast, $k_{ex} > \Delta\omega$, a single broadened resonance is observed
- Very fast, $k_{ex} \gg \Delta\omega$, a single resonance is observed

For two separately observed resonances in the slow exchange regime, their relative amplitudes are proportional to the populations of the two states, while for a single resonance of the fast exchange regime, the chemical shift is linearly dependent on the relative population. These can be used to evaluate equilibrium constant.



NMR Methods for Kinetic Analysis, Fig. 1 Effect of chemical exchange on NMR spectra simulated for two states of equal population. The frequency difference $\Delta\omega = 2\pi\Delta\nu = 100$ rad s^{-1}, exchange rate constants k_{ex} from top to bottom are A: 5, B: 20, C: 60, D: 100, E: 150, F: 500, and G: 1,000 s^{-1}. Three *curves* are shown for the intermediate exchange to illustrate resonance coalescence and maximum broadening. Note that the exchange regime is determined by the relationship between k_{ex} and $\Delta\omega$ (rad s^{-1}) rather than $\Delta\nu$ (Hz) normally used for resonance frequency measurements

$$K_{eq} = k_1/k_{-1} = p_A/p_B,$$

with p_A and p_B corresponding to the respective populations on the two states. Note that in the fast exchange regime, chemical shifts of the individual states are required to determine the populations. The exchange broadening is maximal when the populations are equal. Ligand binding makes a special case of the chemical exchange because populations of the states can be easily controlled through the total concentrations of the components. This allows evaluation of the binding constant and chemical shifts of the free and bound form for the fast exchange regime by analyzing concentration dependence of the NMR spectra.

Dynamic Equilibrium Methods

Depending on the exchange regime, different methods are used to evaluate the exchange rate. In the very slow range, the exchange contribution into the linewidth is negligible. However, if the exchange rate is faster than spin-lattice relaxation rate R_1, net magnetization can be transferred between the exchanging sites. This transfer is detected using exchange spectroscopy in experiments identical to those used for the detection of Nuclear Overhauser Effect (NOE) (Ernst et al. 1987; Rule and Hitchens 2006). In a 1D version, the signal of the spin in one of the states is selectively perturbed, and the change in the intensity of the resonance corresponding to the exchanged state is monitored. The exchange rate is determined from the dependence of intensity change on the transfer time. In a 2D version of the experiment, cross-peaks are detected between the resonances corresponding to the exchanging states and the exchange rate is determined from their intensities. Care has to be taken to avoid contributions from other magnetization transfer mechanisms. The R_1 values normally range between 1 s^{-1} for proteins to <0.1 s^{-1} for small molecules, placing the rate constants measurable by the exchange spectroscopy at fastest end of the rates accessible by nonequilibrium measurements.

In the slow exchange regime, exchange contribution into the linewidth becomes significant and can be used to evaluate the exchange rate. If two well-separated lines are observed, the increase in the linewidths of the two sites is (Rule and Hitchens 2006):

$$\pi\Delta\nu^{1/2}{}_{A}=R_{2A}+k_1;\ \pi\Delta\nu^{1/2}{}_{B}=R_{2B}+k_{-1}$$

where R_{2A} and R_{2B} are the intrinsic relaxation rates in the absence of exchange. If the intrinsic relaxation rates are known or negligible relative to the exchange contribution, the exchange rate constants can be calculated directly from the linewidths. When lines start coalescing, no analytical approximation for the lineshapes can be derived and the exchange rate has to be evaluated through numerical data fitting. The maximum line broadening and complete coalescence of the resonances occurs when k_{ex} is close to $\Delta\omega$ (Fig. 1). As the exchange rate increases further, the linewidth decreases and the lineshape simplifies with the exchange broadening in the fast exchange regime equal to (Rule and Hitchens 2006):

$$R_2=p_AR_{2A}+p_BR_{2B}+p_Ap_B\Delta\omega^2/k_{ex}.$$

The exchange rate can only be evaluated if the populations, chemical shift differences, and the relaxation parameters of the individual states are known, which is usually possible only for ligand binding.

Under fast to very fast conditions, the exchange contribution can be derived independently from other parameters using relaxation dispersion experiments that measure the dependence of the transverse relaxation rate on delay between refocusing pulses in CPMG-based experiment or on the intensity of the effective magnetic field strength in the spin-lock variant (Lipchock and Loria 2009; Rule and Hitchens 2006). The exchange contribution into transverse relaxation during CPMG sequence with the delay between refocusing pulses τ depends on the multiple τ^*k_{ex} and disappears completely if $\tau^{-1} \gg k_{ex}$. The dispersion curve is obtained by measuring relaxation rates at different τ values, and the exchange rate constant is derived by fitting the experimental curve into the model equation. Fast transverse relaxation in proteins limits the maximum τ value to ~10 ms and the sample heating prevents using values significantly shorter than 1 ms, limiting the accessible exchange rate constants to $\sim10^2$–10^4 s^{-1}. Similar rate constants can be obtained in the on-resonance spin-lock relaxation dispersion experiment where the relaxation rate is measured as a function of the spin-lock field applied on-resonance, with the fastest rate constant limited by the intensity of the radio-frequency field. Faster rate constants up to $\sim10^5$ s^{-1} are available from the off-resonance spin-lock experiments, where the strength of the effective magnetic field is increased through the contribution of the main magnetic field.

Overall, NMR offers methods for detailed kinetic analysis of processes on the time scale of days to fractions of milliseconds. Simultaneous monitoring of multiple sites and identification of intermediate states and interaction sites provide means to correlate structural and kinetic information.

Cross-References

▶ Quenched-Flow Methods

References

Cavanagh J, Fairbrother WJ, Palmer III AG, Rance M, Skelton NJ. Protein NMR spectroscopy. Principles and practice. 2nd ed. Amsterdam: Elsevier; 2007.

Ernst RR, Bodenhausen G, Wokaun A. Principles of nuclear magnetic resonance in one and two dimensions. Oxford: Clarendon; 1987.
Levitt MH. Spin dynamics. Basics of nuclear magnetic resonance. Chichester: Wiley; 2001.
Lipchock JM, Loria JP. Monitoring molecular interactions by NMR. In: Shriver JW, editor. Methods in molecular biology, vol. 490. New Haven: Humana press; 2009. p. 115–34.
Rule GS, Hitchens TK. Fundamentals of protein NMR spectroscopy. Dordrecht: Springer; 2006.

NMR of Carbohydrates

► Carbohydrate NMR Spectroscopy

NMR of Lipids

Anthony Watts
Biomembrane Structure Unit, Department of Biochemistry, University of Oxford, Oxford, UK

Synonyms

Membrane dynamics; Membrane lipids; Membrane order; Membrane structure; NMR of membranes

Definition

NMR provides a way of probing the molecular arrangements that lipids can adopt, their molecular order and dynamics, usually in hydrated solution and either with or without other components, such as membrane proteins or ions, under a range of conditions of temperature, pressure, and pH. The information can usually be obtained with minimal perturbation chemically, although the inherent low sensitivity of NMR can present practical limitations.

Introduction

Lipids may indigenously contain NMR visible nuclei, such as protons (^{1}H), nitrogen (^{14}N), carbon (^{13}C at its natural abundances of 1.1%), or phosphorous (^{31}P), but can also be labeled with other nuclei, such as deuterium (^{2}H), fluorine (^{19}F), oxygen (^{17}O), or enriched with carbon (^{13}C), to enable detection with more or less perturbation of the lipid properties depending on the nuclear substitution.

Chemical Analysis of Lipids by NMR

High-resolution, solution-state (^{1}H, ^{13}C, or ^{31}P) NMR is a routine method for the chemical analysis of lipids, either after synthesis or of a naturally occurring, and usually purified, lipid.

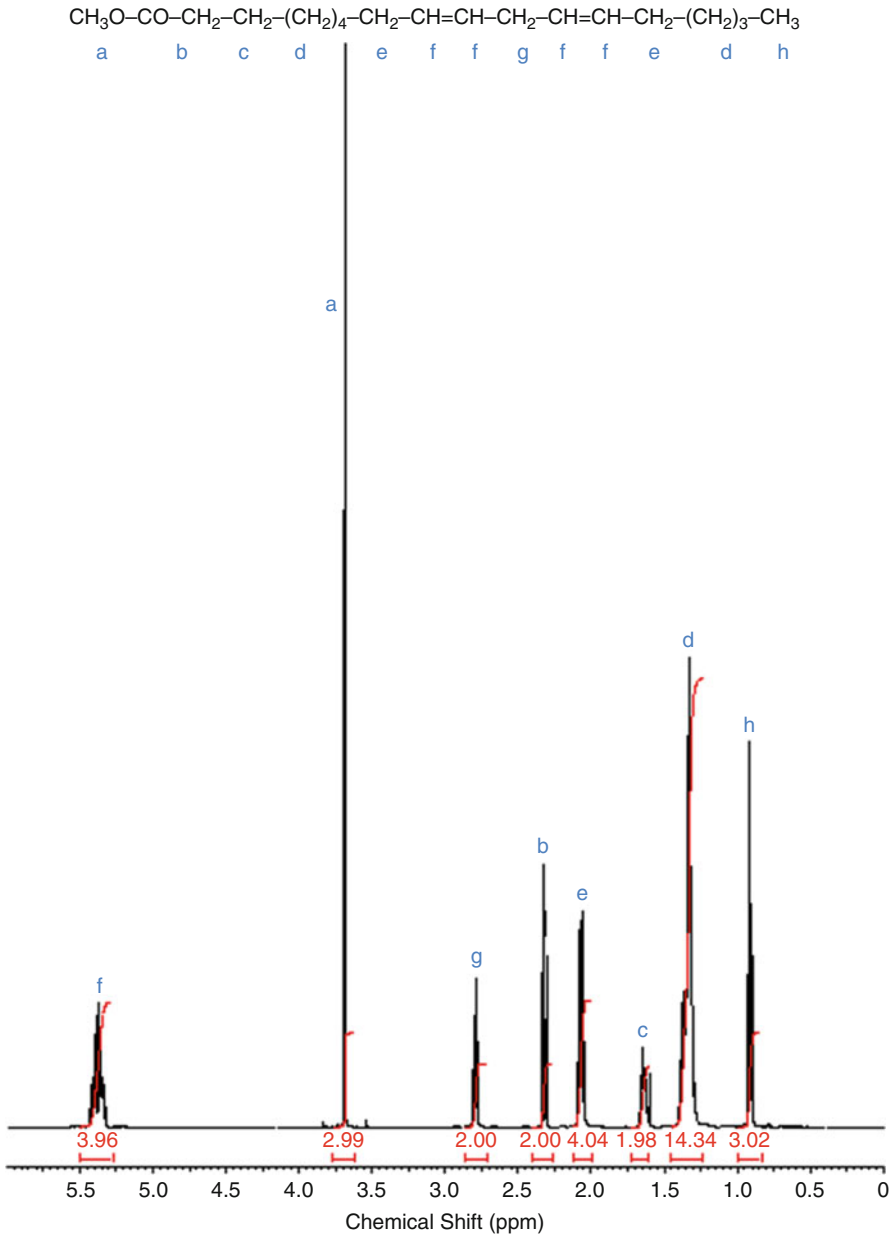
The solvent used for NMR study will depend on the lipid solubility, but chloroform/methanol and dichloromethane/methanol are universally used since in these, lipids are often monomeric and readily soluble. Carbon tetrachloride is suitable for nonpolar lipids. For lipids extracted from natural membranes, the solvent mix and temperatures used for extraction can change the efficiency of extraction. Around 0.4–0.7 mL of sample is required in a 5 mm standard solution-state NMR tube, at a concentration of 1–20 mM, but this can vary depending on the lipid and solvent.

For natural abundance proton (^{1}H) NMR, the solvents are deuterated to reduce the solvent proton signal and provide a lock-signal, but for ^{13}C NMR, the solvent natural abundance is not so much of a problem and can be reduced digitally, or the sample volume decreased thereby increasing the lipid concentration. Spectral dispersion is also much larger in ^{13}C NMR (~200 ppm) than in ^{1}H NMR (~10 ppm), so spectral overlap is significantly reduced.

The usual assignment strategies used in small molecule NMR analysis also apply to lipids, using a routine standard for reference, such as tetramethylsilane (TMS). Because of the complex electronic distribution induced by chain unsaturation, the introduction of one or more double bonds gives rise to several extra resonances in the spectrum compared to the saturated chain homologues. An extensive range of proton and ^{13}C NMR assignments is available from the AOCS Lipid Library http://lipidlibrary.aocs.org/index.html.

Proton NMR of Lipids

In proton NMR spectra of lipids in isotropic solvents (Fig. 1), from lower to high chemical shift (ppm with respect to TMS), it is the lipid chain terminal methyl

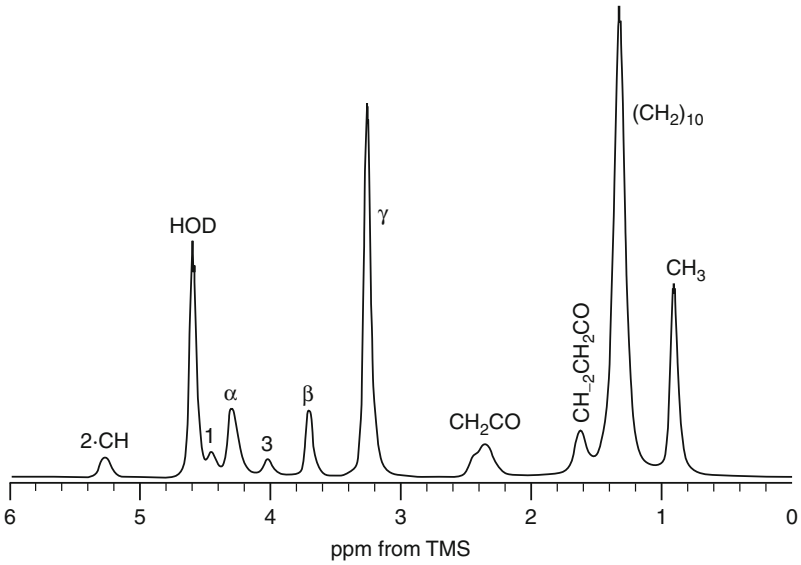


NMR of Lipids, Fig. 1 Proton NMR spectra of lipids. Methyl linolenate in deuterochloroform solution at 300 MHz (tetramethylsilane as internal standard) shows features in the spectrum corresponding to the protons in the fatty acid labeled "a to h" (Christie 1982)

protons that are often the narrowest lines (~0.8–1.0 ppm), with an envelope (1.6–1.3 ppm) of resonances for those protons at carbons in the center of the hydrocarbon chain (~3 to ~8), followed by protons at methylene groups either side of the C=C group (2.0–2.2 ppm). Protons adjacent to the C=O group are observed next (2.2–2.3 ppm), and methylene group protons between the C=C are at ~2.7–2.8 ppm. Characteristic resonances are observed for olefinic or other unsaturated groups (~5–6 ppm), and if the acid is esterified, (as here in Fig. 1), methoxy protons are usually sharp and at ~3.5–3.7 ppm. Similar strategies can be used for phospholipids and sterols.

Quantification of resonances to aid molecular analysis is more straightforward in ^{1}H NMR spectra than ^{13}C spectra, since the relaxation (and hence integrated signal intensity) of ^{13}C is complicated by the proximity of protons.

Although much of the NMR for chemical analysis of lipids is now superseded by mass spectrometry, for



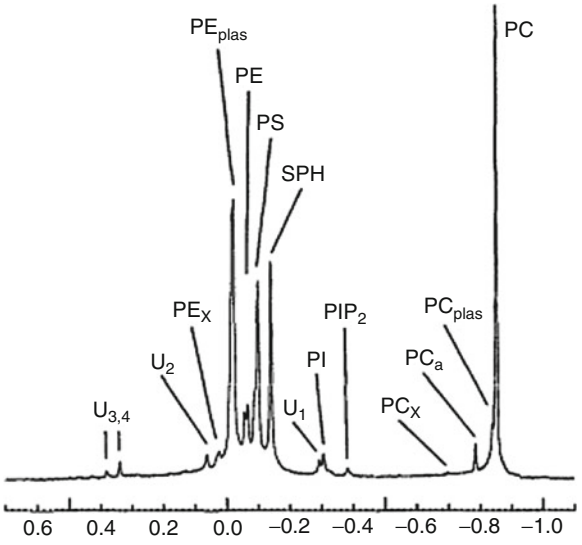
NMR of Lipids, Fig. 2 500 MHz proton 1D Fourier-transform MAS (2.7 kHz) NMR spectrum of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC)-^2H_2O (50 wt%) at 36°C. Assignments are as shown for the glycerol (1, 2-CH, 3), head group (a, b, g), terminal methyl (CH_3), and acyl chain ($(CH_2)_{10}$) protons (Adapted from Forbes et al. 1988)

more complex lipid bilayer and membrane NMR, especially solid state, magic-angle spinning NMR of hydrated lipid bilayers (Fig. 2), these assignments in solution are then vital in the interpretation of the membrane or bilayer spectra to give molecular details.

Phosphorous-31 NMR of Phospholipids

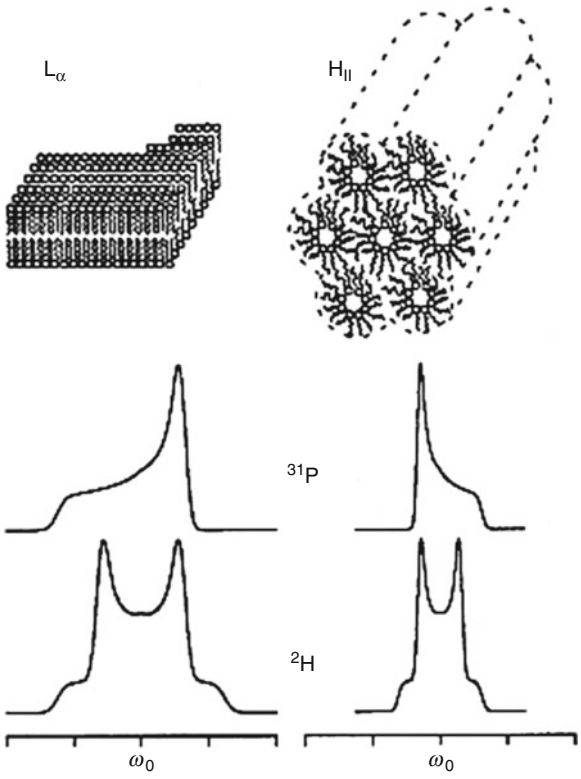
Phosphorous-31 has highly anisotropic, axially symmetric magnetic interactions (maximum anisotropy $\sim \Delta\sigma = \sigma_{11},\sigma_{33} - \sigma_{22} \sim 70$ ppm; ~13 kHz at 500 MHz; given that $\sigma_{11} \sim \sigma_{33}$ (~85 ppm) $\gg \sigma_{22}$ (14 ppm) for the phosphoesters of phospholipids). This anisotropy is averaged to ~40 ppm (~6 kHz at 500 MHz) by the fast motion of the phospholipid head group around the molecular long axis in a lipid bilayer or other polymorphic form, with an order parameter (S_{mol}) of ~0.1–0.2 and within a cone of angle ~10–20° (angle Θ, where $S_{mol} = ½(<3\cos^2\Theta - 1>)$, and the resultant NMR spectra can be diagnostic of the lipid polymorphic form.

Phosphorous-31 in phospholipids shows chemical shift sensitivity (1–3 ppm) to the ester moiety at the head group, and can therefore be used for chemical identification of phospholipid classes in isotropic solvent lipid extracts from natural membranes (Fig. 3) to give high-resolution-like spectra. The different phospholipid types in mixed bilayers, however, are not usually distinguished in a statistically recorded, wide-line NMR powder pattern (1–3 ppm is not resolved over ~40 ppm wide powder spectra) (Fig. 4).



NMR of Lipids, Fig. 3 Typical ^{31}P NMR profile of solvent-extracted phospholipids from bovine white matter. The assignments of the resonances are PC, phosphatidylcholine; PCplas, PC-plasmalogen; PCa, alkylacyl-PC; PS, phosphatidylserine; SPH, sphingomyelin; PE, phosphatidylethanolamine; PEplas, PE-plasmalogen; PI, phosphatidylinositol; PIP2, PI-bisphosphate; PEx, PCx, unidentified PE and PC derivatives; Um, unidentified phospholipids (Adapted from Edzes et al. 1992)

Magic-angle spinning (MAS) solid-state NMR can be used to average out the chemical shifts spectral



NMR of Lipids, Fig. 4 Comparison of the L_α (on the *left*) and hexagonal H_{II} (on the *right*) phases of phospholipid water dispersions. A schematic representation of their geometry is shown at the *top*. Due to the rapid diffusion around the cylinder axes in the H_{II} phase, both the ^{31}P NMR (*middle*) and the ^{2}H NMR (*bottom*) spectra are scaled by a factor of $-1/2$. The ^{31}P NMR spectra can yield both the absolute value of S and its sign. Because of their symmetry, the ^{2}H NMR spectra do not provide the information about the sign of S. For illustrative purposes, the simulated powder patterns are shown on arbitrary horizontal and vertical scales, with ω_0 indicating the Larmor frequency for each nucleus

anisotropy (spinning rate, $\omega_r > {}^{31}P$ chemical shift anisotropy, ~6 kHz for phospholipids in bilayers) and therefore different lipids types in bilayers can be resolved, quantified and any specific interactions with ions or proteins quantified.

Deuterium NMR of Labeled Lipids

In deuterium NMR, specific proton-deuterium substitutions can be made at various positions in the lipid, or alternatively, uniform labeling, usually in the acyl chains, is used. Deuterons in the lipid head groups are particularly sensitive to surface interactions due to ions, pH, and protein interactions.

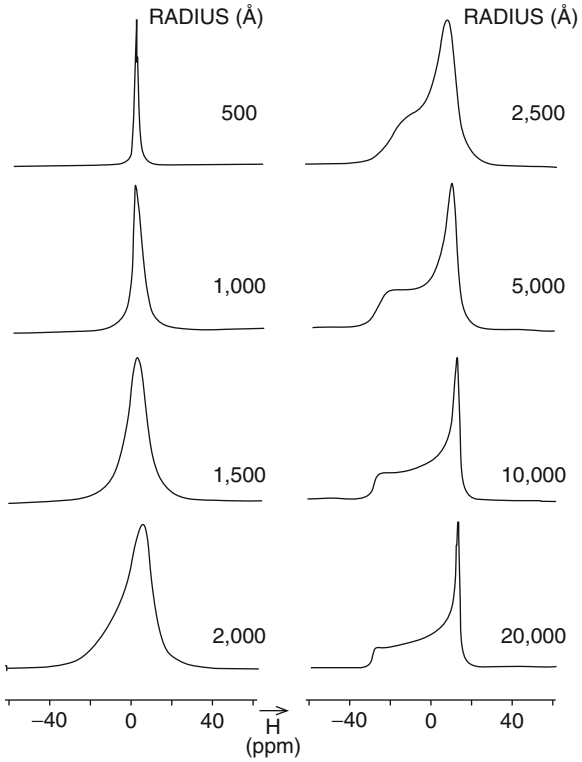
Quadrupole splittings from the spin $m_I = 1$ spectra are measured for assigned -CD, $-CD_2$, or $-CD_3$ groups, or groups of these labeled positions. More usually a multiply, or per-deuterated chain or head group is used to form bilayers. Because of the low sensitivity of ^{2}H as an NMR nucleus, and its wide spectral lines (up to 127 kHz; spin = 1; $\gamma = 6.54$ ~1/7 Hz sensitivity of ^{1}H), most of the systems studied composed of the ^{2}H-labeled lipid, or in mixture with other non-labeled lipids where the ^{2}H-lipid was still a major component. The quadrupole splittings are sensitive to lipid order and lipid dynamics (see: ▶ Lipid Organization, Aggregation, and Self-assembly).

Lipid Polymorphism by NMR

Large, extended bilayers: Proton NMR spectra of lipids in large extended bilayers are usually featureless because of the strong proton dipolar broadening (~100 kHz) which is not averaged through molecular tumbling (τ_r (extended bilayers) $\gg$~100 kHz). However, magic-angle spinning (MAS) and decoupling can relieve this broadening to give narrow spectral lines (Fig. 2), although very fast (>~15 kHz) spinning can centrifuge dispersions in the NMR sample tube.

The anisotropy of phosphorous-31 and deuterium nuclei has been found to be most useful in wide-line (non-spinning) NMR for characterizing and studying large extended phospholipid bilayers, which need to have a rotational correlation time (τ_r) $\gg$the maximum anisotropy of the nuclear interactions; ~6 kHz for ^{31}P; ~127 kHz for ^{2}H. These static bilayers, on the NMR anisotropy averaging timescale, give rise to spherically averaged powder patterns (Fig. 4), and characterize fast long axis motion of the lipids in bilayers around the bilayer normal, and reduced amplitude of motion of the head groups and acyl chains within the bilayer. The amplitude of motion is characterized by order parameters, and the rate of motion by dynamic measurements (see: ▶ Lipid Organization, Aggregation, and Self-assembly).

Small bilayer vesicles: If phospholipid bilayers are rendered into small bilayer vesicles by, for example, sonication or extrusion through a filter, the phosphorous-31



NMR of Lipids, Fig. 5 The size (radius, R) of a small vesicle can be estimated from the line width of the ^{31}P NMR spectrum, assuming vesicle tumbling is the major contribution to the spectral broadening, which is a good assumption for most liquid-crystalline bilayers. Simulated ^{31}P-NMR spectra of dioleoylphosphatidylcholine vesicles of different sizes (radius, R) at 30°C using a ^{31}P anisotropy, $(\sigma_{\parallel} - \sigma_{\perp}) = 3{,}550$ Hz, solvent viscosity, $\eta = 0.008$P, orientationally independent Lorentzian full line width at half height = 60 Hz, $D_{diff} = 6.2 \times 10^{-7}$ cm^2/s, and R as indicated (From Burnell et al. 1980)

NMR spectrum obtained is for an isotropically averaged system. The NMR line width can give an indication of the size of the vesicle for ~50–500 nm as a result of the averaging of the ^{31}P chemical shift anisotropy (Fig. 5).

The averaging arises from both lateral lipid diffusion around the vesicle and vesicle tumbling, due to Brownian rotational diffusion, described by $\tau_c^{-1} = 6/r^2\,(D_t + D_{diff})$ where $D_t = kT/8\pi r\eta$, η is the viscosity of the medium, and D_{diff} is the rate of lateral lipid diffusion in the bilayer ($D_{diff} \sim 10^{-7}$–10^{-8} s). However, other factors, such as the lipid thermotropic phase, lipid-protein interactions, and ionic interactions, can also reduce the ^{31}P anisotropy of lipids in a polymorphic phase.

Hexagonal phases: Lipids, notably phospholipids, with head groups that occupy a small surface area, and acyl chains that occupy a large bilayer volume, can form inverted lipid polymorphic phases termed hexagonal, H_{II} phases. The hexagonal packing of lipid tubes is confirmed by the long spacing ratios ($1/\sqrt{3}$; 1/2; $1/\sqrt{7}\ldots$) in x-ray scattering of H_{II} phase lipids. Phosphorous-31 and deuterium NMR of deuterated lipids give rise to inverted powder patterns of half spectral anisotropy (~20–25 ppm) compared to the same or similar lipids in a bilayer (~40–50 ppm) (Fig. 4) due to the averaging of the spectral anisotropy by lateral diffusion down the tubes and around the tube circumference.

Phospholipids can convert from a bilayer phase at lower temperatures to a hexagonal H_{II} phase at higher temperatures as the chain volume increases with increasing thermal motions.

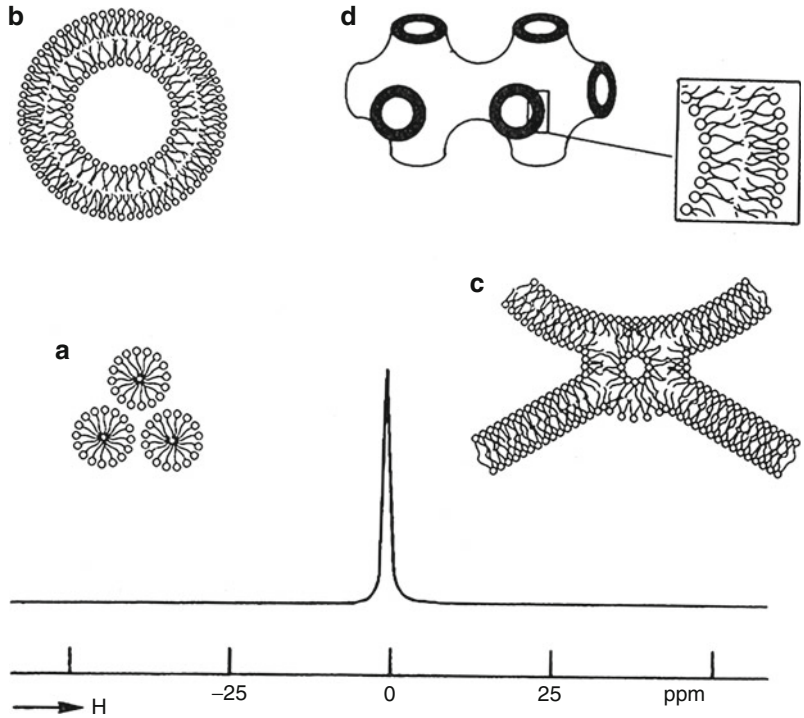
Phosphorous-31 NMR studies have been useful in describing polymorphic phase diagrams for phospholipids and helped in the concept of developing area/volume models for understanding lipid polymorphism.

Cubic, micellar, and rhombic phases: Phosphorous-31 NMR spectra of cubic, micellar, and rhombic lipid phases show an isotropic spectrum (Fig. 6) due to the fast lateral diffusion (D_{diff}) of phospholipids around the various curved structures at a rate ($1/D_{diff}$) fast enough to average the ^{31}P CSA (~3–6 kHz; ~20–40 ppm). This averaging occurs for cubic phases or bilayer vesicles (radius < ~30–50 nm), as well as for micelles, and the mechanisms are the same as for small bilayer vesicles.

Lipid Domains and Phase Separation Detected by NMR

Multicomponent lipid bilayers can give rise to NMR spectra that can be used to distinguish lipid domains or lateral phase separation. For this to be evident, lipid exchange between the various separated regions must be slow on the timescale of the NMR property being exploited. Most usually, phosphorous-31 and deuterium NMR are used since the rate of averaging of the chemical shift or quadrupolar anisotropies, respectively, falls in the 10^3–10^6 s range.

A significant number of phospholipids and lipids have been shown to induce lateral phase separation, including cholesterol, ceramides, and sphingomyelin, in mixtures with phospholipids, usually phosphatidylcholines or phosphatidylethanolamines. The size, differential chain order, flexibility, and life times of domains and their lipids can be described by NMR, which distinguishes



NMR of Lipids, Fig. 6 Isotropic ^{31}P NMR spectra are recorded for cubic, micellar, rhombic phases, and small vesicles comprised of phospholipids

liquid-ordered and liquid-crystalline components as a result of the slow, two (or more) site exchange of the lipids on the NMR anisotropy averaging timescale (~4–6 kHz for ^{31}P; ~127 kHz for ^{2}H) between the different phases (liquid-ordered, liquid-disordered, etc).

Lipid Order Parameters by NMR, S_{mol}

Often lipid bilayers are described as "fluid" or as having "flexible" chains – such definitions lack physical units or definition, and so order (in ° as an amplitude, or an order parameter, S_{mol}) and dynamics (as a correlation time, τ_c, or rate of motion, τ_c^{-1}) are required to describe fully lipid behavior (Fig. 7).

Lipid order is determined by the amplitude of acyl chain motion within a polymorphic lipid phase, most usually bilayers. Order is defined by an order parameter S_{mol}:

$$S_{mol} = 1/2(3 < \cos^2\theta - 1 >)$$

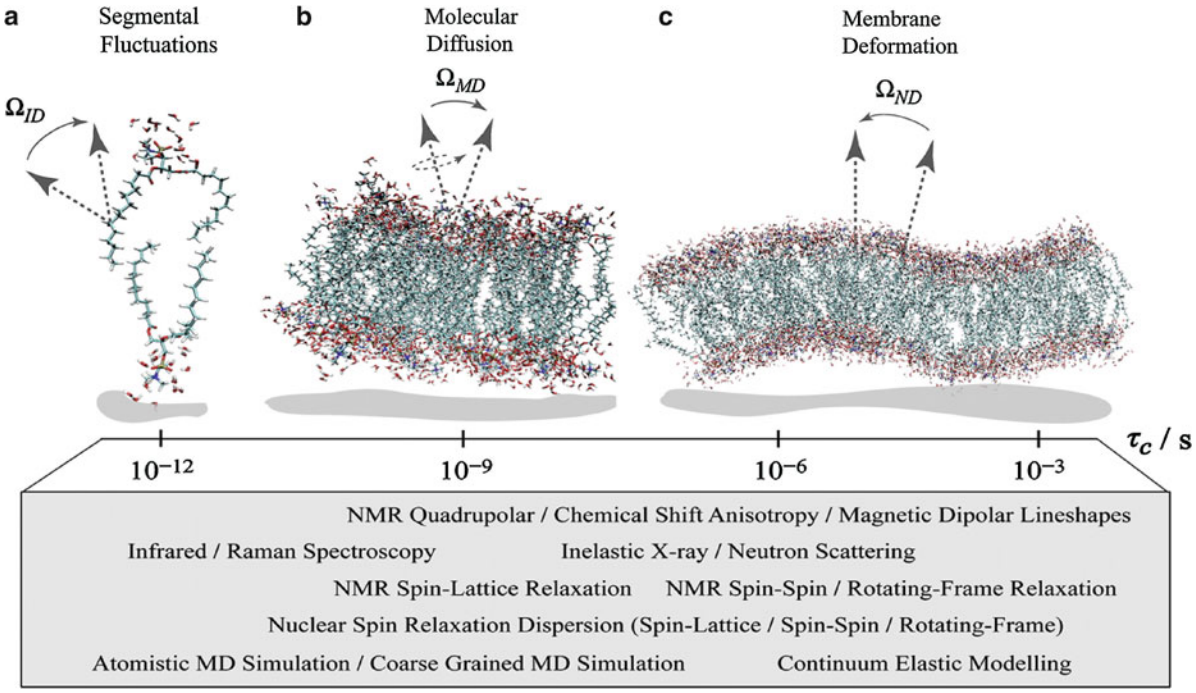
where θ is the angle of a specific moiety (often a reporter group) with a given reference such as the bilayer normal. Since $0° > \theta > 90°$, the order parameter S is defined as $0 > |S| > 1$; $S = 1$ is defined for *all-trans*, ordered chains and $S = 0$ for isotropic, disordered chains.

NMR order parameters are measured from anisotropy averaging of an NMR property, such as quadrupolar interactions ($\Delta\nu_q$) in deuterium NMR or chemical shift (Δ) in phosphorous-31 NMR. For ^{19}F-labeled lipids, the NMR spectral anisotropy arises from residual chemical shift anisotropy ($\sigma_{\parallel} - \sigma_{\perp}$); the proton-fluorine intrachain interactions (Δ_I) are modulated by motional amplitude to give ^{19}F NMR spectra from which S_{mol} can be determined.

Since the value of S_{mol} is defined as a time-averaged ($< >$) property over the amplitude (θ) of acyl chain motion, the rate of motion (or as a correlation time, τ_c^{-1}) of the reporter needs to be fast with respect to the NMR property (for ^{2}H in lipid chains, $\Delta\nu_{q(max)} \sim 127$ kHz, and so $\tau_c^{-1} < \sim 10^{-4}$ s) being measured.

The way in which S_{mol} is measured experimentally depends on the nucleus being detected. For specifically ^{2}H-labeled lipids, the individual quadrupolar splitting ($\Delta\nu_q$ in Hz) can be determined directly (Fig. 8) and S_{mol} calculated from:

$$|S_{mol}| = \Delta\nu_{q(obs)}/\Delta\nu_{q(max)}$$



NMR of Lipids, Fig. 7 NMR spectroscopy reveals phospholipid membrane dynamics and structure over a range of timescales. Phospholipid mobility is characterized by segmental fluctuations, molecular diffusion, and viscoelastic membrane deformation. Orientational fluctuations correspond to geometry of interactions via Euler angles Ω and by correlation times τ_c of the motions. (**a**) Principal axis system of ^{13}C-^{1}H or C-^{2}H bonds fluctuates due to motions of internal segmental frame (*I*) with respect to the membrane director axis (*D*). (**b**) Diffusive phospholipid motions are described by anisotropic reorientation of molecule-fixed frame (*M*) with respect to the membrane director axis (*D*). (**c**) Liquid-crystalline bilayer lends itself to propagation of thermally excited quasi-periodic fluctuations in membrane curvature expressed by motion of the local membrane normal (*N*) relative to membrane director axis (*D*). The appropriate range of timescales of various complementary biophysical methods is indicated (From Leftin and Brown 2011)

where $\Delta\nu_{q(max)} = 3/2(e^2qQ/h) \sim 127$ kHz for hydrocarbon chains.

Here, molecular averaging around the long molecular axis is assumed, which is a good approximation for acyl chains in phospholipid bilayers. However, if the axis about which averaging occurs does not coincide with the molecular axis, for example, in –C=C– kinks in chains, or in the head groups of phospholipids, then $S_{mol} \neq S_{CD}$, the order parameter of the -CD group. Now:

$$S_{CD} = S_{coll} \cdot S_{mol} \cdot S_{intra}$$

where (S_{coll}) represents the fluctuations of the bilayer normal, (S_{mol}) the reorientations of the molecular long axis around the bilayer normal, and (S_{intra}) the reorientations of the C-D bond vector with respect to the molecular axis (Fig. 8), giving:

$$S_{(obs)} = S_{bilayer} \cdot S_{CD}$$

$$S_{bilayer} = 1/2(3<\cos^2\theta - 1>)$$

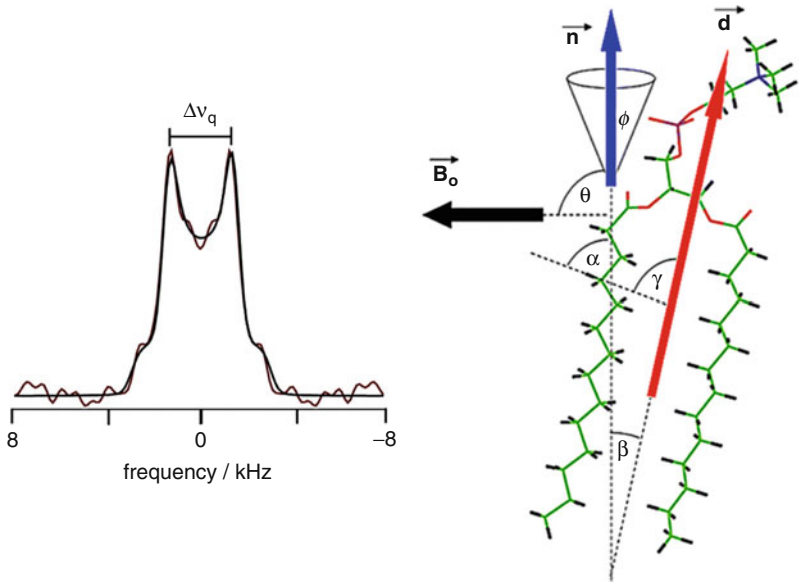
$$S_{CD} = 1/2(3<\cos^2\alpha - 1>)$$

$$S_{coll} = 1/2(3<\cos^2\phi - 1>)$$

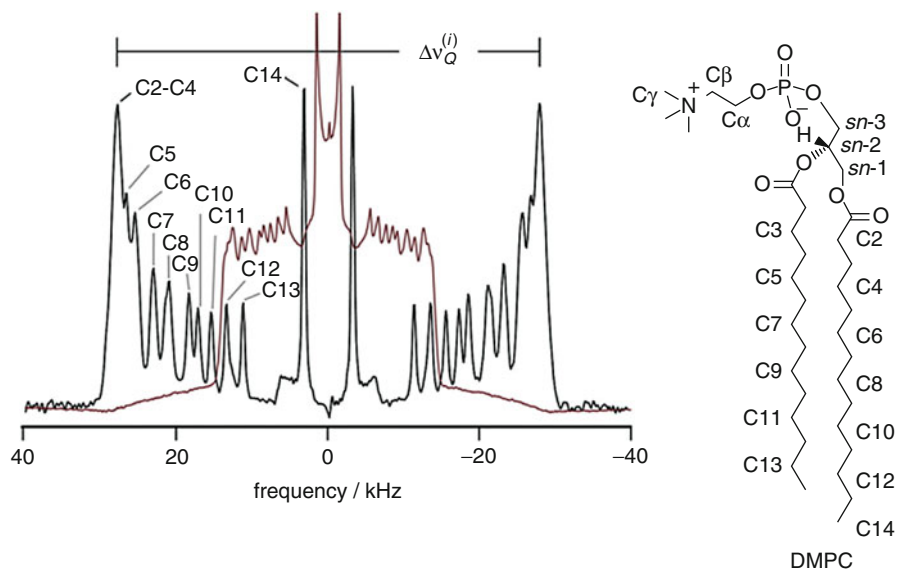
$$S_{mol} = 1/2(3<\cos^2\beta - 1>)$$

$$S_{intra} = 1/2(3<\cos^2\gamma - 1>)$$

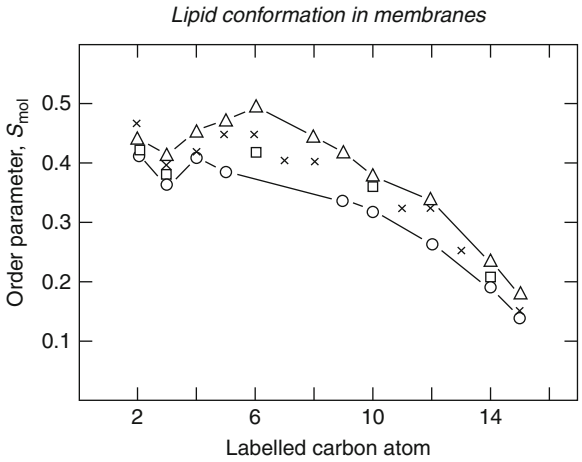
The deuterium NMR spectra for per-deuterated lipid chains comprise an overlap of all individual labeled positions and two approaches are used to extract order information. Firstly, individually assigned resonances can be identified either directly (e.g., for the lower part



NMR of Lipids, Fig. 8 Typical deuterium NMR spectrum (*left*) for a single -CD_2 on an acyl chain of a phospholipid in hydrated bilayers. The quadrupole splitting ($\Delta\nu_q$ in Hz) is measured as shown and can be used to give information about local order of the lipid chain, as an order parameter, S_{CD} (Adapted from Leftin and Brown 2011). When the axis of motional averaging does not coincide with the long molecular axis ($\vec{\mathbf{n}}$), and is at an angle (θ) with respect to the applied magnetic field ($\vec{\mathbf{B}}_o$), the observed CD vector is now tilted at an angle a with respect to the bilayer normal. If the rotation of lipids about different axes are independent, then the order parameter may be represented as the product of all individual contributions $S_{CD} = S_{coll} \cdot S_{mol} \cdot S_{intra}$. The bilayer normal is now subject to fluctuations around the vector $\vec{\mathbf{n}}$ within a cone (angle φ), whereas the molecular director $\vec{\mathbf{d}}$ is at an angle b to the bilayer normal, and the observed CD vector is at an angle γ with respect to the molecular director (Adapted from Vermeer et al. 2007)



NMR of Lipids, Fig. 9 A typical deuterium NMR spectrum of a per-deuterated acyl chain phospholipid (here d_{54}, 1,2-dimyristoylphosphatidyl choline) in bilayers (*light red line*) showing multiple overlapping quadrupole splittings that can be deconvoluted (*black lines*) into the individual spectral components for each C-atom (Adapted from Leftin and Brown 2011)



NMR of Lipids, Fig. 10 Normalized order profiles of different bilayers. Variation of the molecular order parameter, S_{mol}, with the segment position. ○, 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine. Δ, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine. □, 1,2-dipalmitoyl-*sn*-glycero-3-phosphoserine. ×, *Acholeplasma laidlawii* (Adapted from Seelig and Browning 1978)

of the chain and terminal -CD_3) or de-Paking routines that can be used to deconvolute individual values of $\Delta\nu_q$ (Fig. 9).

Secondly, the first moment (M_1) of a ^{2}H-NMR spectral envelope from per-deuterated lipids can give an averaged order change for a bilayer perturbation, induced, for example, by other lipids, peptides, or proteins. M is defined as:

$$M_1 = 1/A\ \Sigma|\omega_i I_i|$$

where $A = \Sigma\ I_i$, the total area under the spectrum and I_i is the intensity of the ^{2}H-NMR spectra at frequency ω_i. The origin of the frequency axis is usually defined as the center of symmetry of the spectrum, and the limits for the M_I integration enclose the observed spectrum.

A characteristic deuterium NMR derived order parameter profile across a bilayer shows a similar order parameter ($|S_{mol}| \sim 0.4$) for acyl chain methylene groups in the upper half (glycerol – ~C8/10) of the bilayer and then a decreasing order parameter from the center of the acyl chains (~C8/10) to very large amplitude of motion for chains in the center of the bilayer (Fig. 10). Similar order profiles are also observed in natural membranes (Fig. 10).

Since $S_{CD} = S_{coll} \cdot S_{mol} \cdot S_{intra}$ (see above), the measured quadrupole splitting for deuterated segments at the –C=C– bonds in unsaturated lipid chains, or in the head groups of lipids, is smaller than in similar saturated lipids or than expected, and to determine S_{mol}, some assumptions about the axis of motional averaging need to be made to estimate S_{coll} and S_{intra}.

Lipid Dynamics by NMR

Often lipid bilayers are described as "fluid" or have "flexible" chains – such definitions lack physical units or definition, and so order (in ° as an amplitude, or an order parameter, S_{mol}) and dynamics (as a correlation time, τ_c, or rate of motion, τ_c^{-1}) are required to describe fully lipid behavior.

The timescale sensitivity (ps–ms) and versatility of nuclei (^{1}H, ^{13}C, ^{31}P, ^{2}H) that can be used in NMR of lipids allows a detailed description of lipid dynamics to be constructed (Fig. 7) using NMR relaxation parameters (T_1; T_2) and chemical shift or quadrupolar anisotropy averaging.

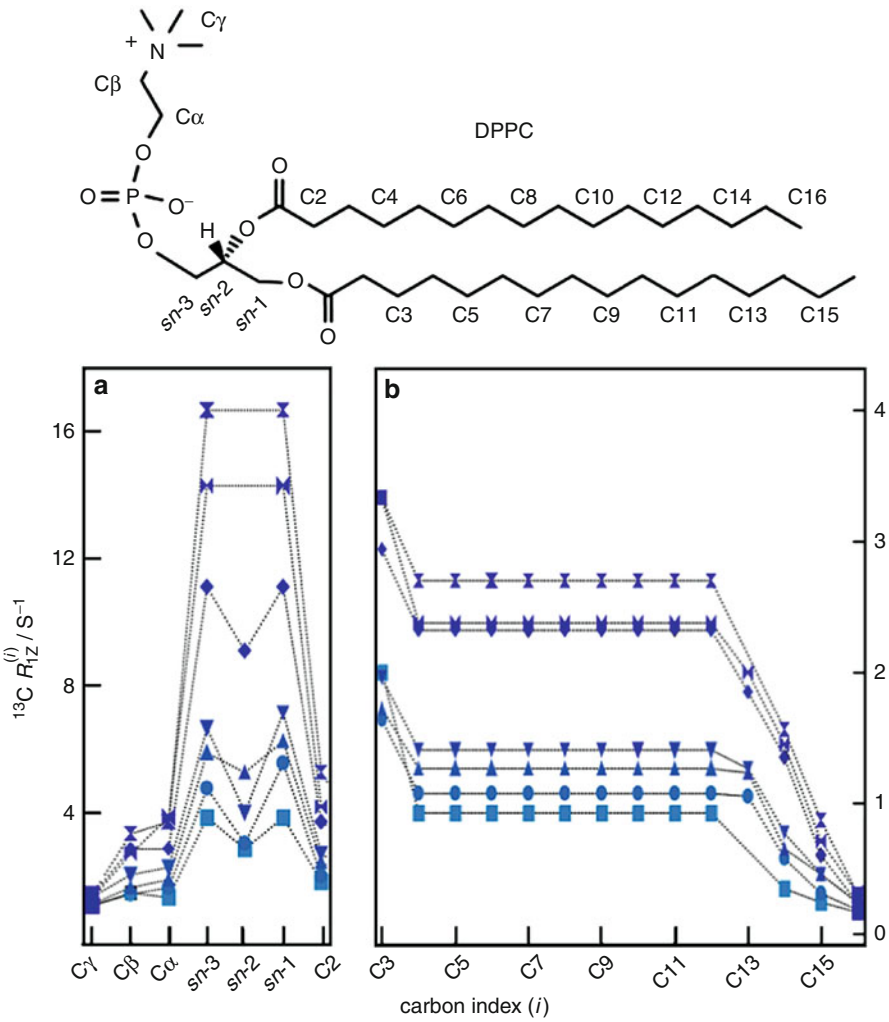
^{13}C NMR spin-lattice relaxation rates measured for lipids in bilayers show a profile (Fig. 11) with very fast chain motion at the center of the bilayer, and a gradient of motional rates from the bilayer surface to the center, with complex dynamics in the glycerol region. The frequency selection of the NMR experiment permits a more detailed quantification of dynamic timescales for individual ^{13}C positions.

Double bonds in lipid chains induce a complex dynamic behavior in which allylic groups show increased and vinylic groups reduced motional rates.

Lipid Hydration and Ionization State

Lipid hydration increases lipid mobility as a result of the hydration-induced surface area, and polar head groups undergo conformational adjustments. Deuterium and phosphorus-31 NMR are suitable methods for sensing dynamics of lipids in response to hydration and ionization state of lipids or charged lipophiles. The ^{2}H-NMR spectrum is sensitive to both order and rate of surface motion, through changes in the quadrupole splitting (Fig. 12) and relaxation rates (Fig. 13).

Ionization states for lipid head groups and for charged lipophiles solvated in bilayers can also be measured from ^{2}H-NMR spectra of head-group deuterated lipids. The bilayer surface environment is very



NMR of Lipids, Fig. 11 High-resolution ^{13}C NMR spin-lattice relaxation rate profiles for liquid-crystalline DPPC show a significant NMR detection frequency dependence of motional rates for lipids in bilayers. Spin-lattice relaxation rates $R^{(i)}_{1Z}$ are presented for various carbon positions (i) of DPPC at 50°C at NMR Larmor frequencies of (⧗) 15.04 MHz, (⋈) 20.00 MHz, (◆) 25.15 MHz, (▼) 45.29 MHz, (▲) 90.80 MHz, (●) 125.76 MHz, and (■) 150.84 MHz. (**a**) The glycerol backbone and choline head group are resolved in natural abundance ^{13}C NMR in the 50–80 ppm region and exhibit a pronounced dispersion of the relaxation rates. (**b**) Acyl chain segments are observed in the 0–40 ppm fingerprint region of the high-resolution ^{13}C NMR spectrum. The molecular structure of DPPC and assignments corresponding to the carbon index (i) are shown

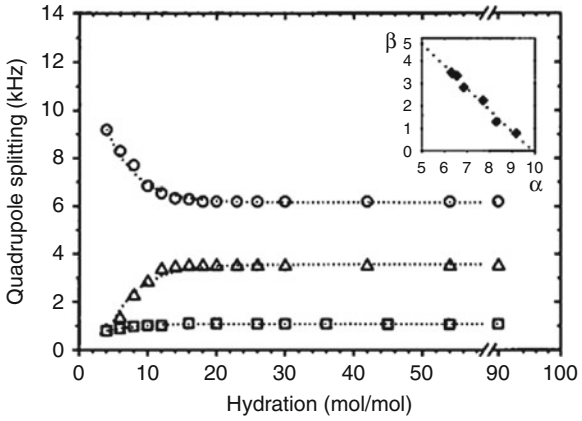
different from a homogeneous aqueous environment and consequently pK_a and electrostatic behavior is significantly different for lipid-solvated lipophiles. For example, fatty acids and charged anesthetics have shifted pK_a behavior when in bilayers compared with when in aqueous solutions.

Lipid Partitioning by NMR

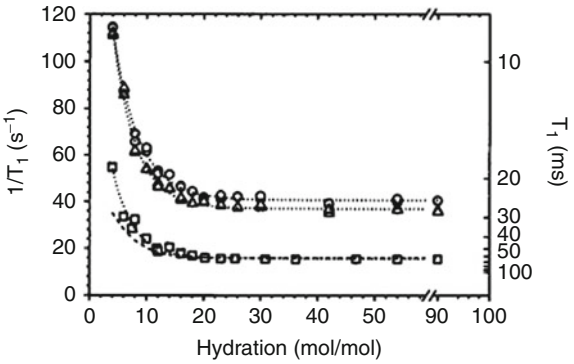
NMR is a useful tool for determining the partition behavior of lipophiles into lipid bilayers and membranes. If the lipophile can be NMR labeled, for example, with deuterium, then NMR resonances from the lipophile in the aqueous phase are resolved from the lipophile dissolved in the bilayer. If two populations are observed, then exchange of the partitioning component is slow on the NMR timescale of the nucleus and property being observed. The advantage of this approach is that equilibrium measurements can be made without the need for filtration or centrifugation used in conventional partition coefficient assays.

Lipid-Protein Interactions by NMR

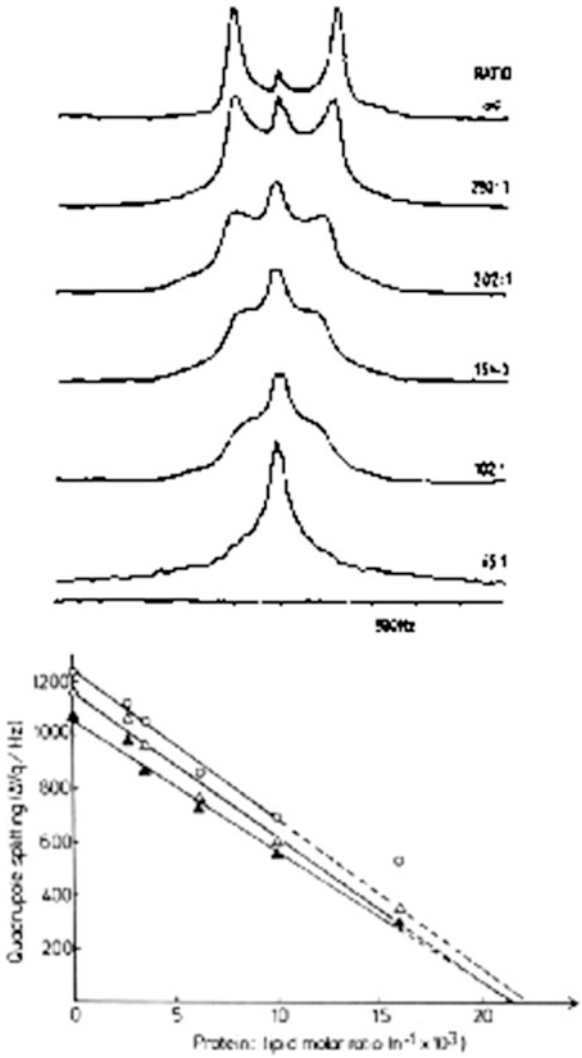
NMR of lipids can be used to report on protein-lipid interactions in bilayers since proteins perturb the lipid order and/or dynamics. Proteins may induce lipids to phase-separate within the bilayer, in which case two or more NMR spectral components are observed (see: Lipid domains and phase separation detected by NMR), or proteins may induce an averaged perturbation on the lipid order and/or dynamics giving rise to



NMR of Lipids, Fig. 12 ^{2}H-NMR quadrupole splitting $\Delta\nu_Q$ recorded as a function of the water/lipid ratio n_w of DOPC multibilayers at 30°C. The data points for the α (○), β (Δ), and γ (□) segments of the choline head group were curve-fitted using (1) (*dotted lines*). The inset shows the linear correlation between the α and β splittings, fitted by the straight line which gives a gradient of $m \sim -1$



NMR of Lipids, Fig. 13 ^{2}H-NMR spin-lattice relaxation rates, $1/T_1$, recorded as a function of the water/lipid ratio n_w of DOPC multibilayers at 30°C, for the α (*circle*), β (*triangle*), and γ (*square*) segments of the choline head group. The isothermally recorded data points were curve-fitted (*dotted lines*) using (1). The *dashed line* for the γ segment represents the hypothetical results calculated for a uniformly reduced temperature ($T_{red} = 0.18$), to take into account the effect of the phase transition on the molecular dynamics in the bilayer



NMR of Lipids, Fig. 14 The deuterium NMR spectra (*upper*) from deuterated lipid head group (phosphatidylcholine, $-N^+(CD_3)_3$) are responsive to the presence of an integral protein (from pure bilayers to 65:1, lipids: protein, molar basis) in the bilayers. The collapse of the quadrupole splitting (*lower*) with increased protein content indicates fast lipid-protein interactions on the deuterium NMR anisotropy averaging timescale (rate $< \sim 10^4$ s)

single component averaged spectra (Fig. 14). Both situations have been used to study lipid-protein interactions by wide-line and magic-angle spinning NMR of lipids.

As an approach to study the specificity of peptide and protein interactions at the structural level, solid-state NMR recoupling methods (REDOR and rotational resonance) can be used to obtain information about the distances between nuclei in lipids and membrane-associated peptides at high resolution. The specificity of lipid-peptide interactions is therefore defined and can give biological insights into antimicrobial or toxins interactions with membranes.

NMR spectra of lipid-protein interaction can give information about the specificity of intermolecular interactions, the degree of perturbation of lipid order and dynamics by proteins, and structure for a membrane-embedded protein.

Summary

NMR can be used as a tool to reveal molecular details about lipids in many different polymorphic and sample forms, and help gain an understanding of thermodynamic behavior, molecular interactions with other membrane components, ions, and pH effects on lipid properties. The intrinsic insensitivity of the method compared with some other biophysical methods is partly outweighed by the lack of chemical perturbation in observing either naturally NMR visible isotopes, or through non-perturbing labeling.

Cross-References

- ▶ ^{19}F NMR
- ▶ Lipid Domains
- ▶ Membrane Fluidity
- ▶ Membrane Lipid Electrostatics
- ▶ NMR
- ▶ Protein NMR Resonance Assignment
- ▶ Residual Dipolar Coupling

References

Burnell EE, et al. Effects of tumbling and lateral diffusion on phosphatidylcholine model membranes 31P-NMR lineshapes. Biochim Biophys Acta. 1980;603:63–9.
Christie WW. Lipid analysis. 2nd ed. Oxford: Pergamon Books; 1982.
Edzes HT, et al. Analysis of phospholipids in brain tissue by 31P NMR at different compositions of the solvent system chloroform-methanol-water. Magn Reson
Forbes J, Bowers J, Shan X, Moran L, Oldfield E, Moscarello MA. Some new developments in solid-state nuclear magnetic resonance spectroscopic studies of lipids and biological membranes, including the effects of cholesterol in model and natural systems. J Chem Soc Faraday Trans. 1988;84:3821–49.
Leftin A, Brown MF. An NMR database for simulations of membrane dynamics. Biochim Biophys Acta. 2011;1808:818.
Lindblom G, Gröbner G. NMR on lipid membranes and their proteins. Curr Opin Coll Int Sci. 2006;11:24–9.
Lindblom G, Orädd G. Lipid lateral diffusion and membrane heterogeneity. Biochim Biophys Acta. 2009;1788:234–44.
Seelig J. Deuterium magnetic resonance: theory and applications to lipid membranes. Q Rev Biophys. 1977;10(3):353–418.
Seelig J, Browning JL. General features of phospholipid conformation in membranes. FEBS Lett. 1978;92:41–4.
Seelig J, Seelig A. Lipid conformation in model membranes and biological membranes. Q Rev Biophys. 1980;13:19–61.
Veatch SL, et al. Liquid domains in vesicles investigated by NMR and fluorescence microscopy. Biophys J. 2004;86:2910–22.
Vermeer LS, et al. Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with ^{2}H NMR experiments. Eur Biophys J. 2007;36:919–31.
Watts A. Magnetic resonance studies of phospholipid-protein interactions in bilayers. In: Cevc G, editor. Phospholipids handbook. Marcel Dekker: New York; 1993. p. 687–740.
Watts A. Solid state NMR approaches for studying the interaction of peptides and proteins with membranes. Biochem Biophys Acta. 1998;1376:297–318.

NMR of Membranes

- ▶ NMR of Lipids

NMR on Solid Proteins Under Fast Sample Rotation

- ▶ Magic Angle Spinning Solid-State NMR on Proteins

NMR Studies of Macromolecular Interactions – Introduction

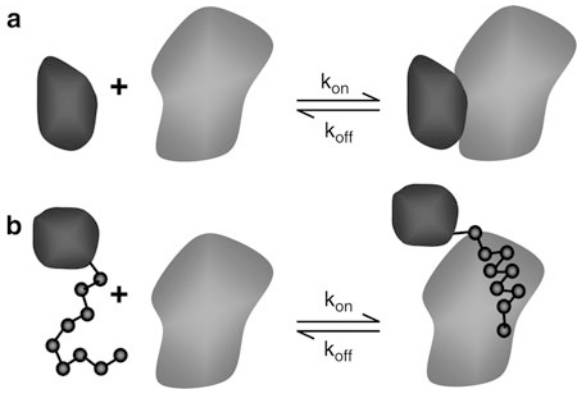
Iain D. Campbell
Department of Biochemistry, University of Oxford, Oxford, UK

Synonyms

Ligand-binding interactions; NMR

Introduction

Entire DNA sequences of hundreds of organisms (http://www.ebi.ac.uk/genomes/) and nearly 80,000 coordinate sets of macromolecules (http://www.rcsb.org/) are now available at the click of a mouse. The next step in trying to understand life is to determine how the molecules in a cell interact with each other, temporally and spatially, in coordinated networks. Work has begun to define this "interactome" and impressive and bewildering maps have been constructed using genetics, affinity purification, and bioinformatics protocols



NMR Studies of Macromolecular Interactions – Introduction, Fig. 1 An illustration of the two main types of protein-protein interaction observed. (**a**) shows a complex formed by two structured domains. (**b**) shows a disordered region of a protein binding to a structured domain. In (**b**) there is an entropic cost in forming the complex which tends to reduce the affinity without reducing the specificity

(Bader et al. 2008; http://thebiogrid.org/). In addition to these, still error-prone, network maps, detailed kinetic and structural information about the complexes formed are needed. Solution state NMR is increasingly recognized as a unique and powerful tool in this area. A search of the PDB database reveals that over 1,000 of 9,300 NMR structures are of protein-ligand complexes. NMR, however, gives not only structure but also valuable kinetic, dynamic, and thermodynamic information about complexes. It is especially useful for analyzing the numerous weak but specific interactions that occur in the cell. In this brief entry I will focus on protein-protein interaction studies. Other reviews on this topic (Zuiderweg 2002; Bonvin et al. 2005; O'Connell et al. 2009) and related topics, such as protein-drug and protein-nucleic acid (Bonvin et al. 2005) interactions are available.

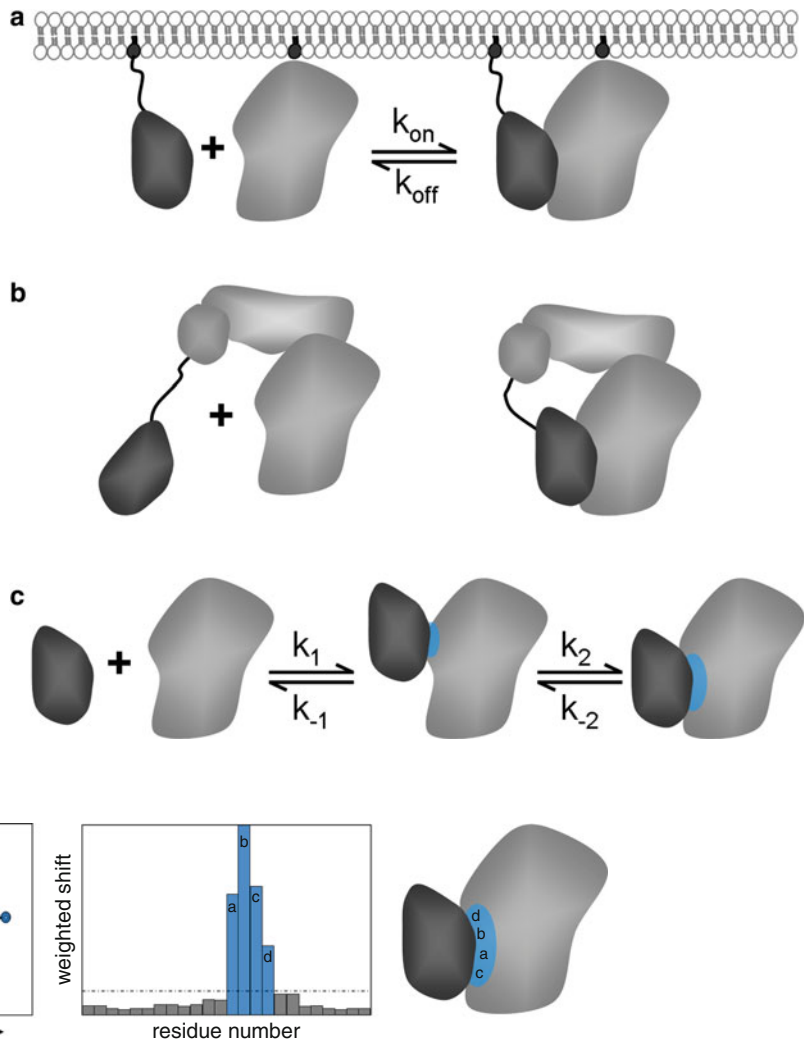
Protein-Protein Interactions

At the structural level, protein-protein interactions can be considered in two main classes (Bader et al. 2008), one where two structured domains form a complex (Fig. 1a) and the other where a complex is formed from a domain and a disordered peptide segment of another protein (Fig. 1b). A large number of proteins (~30%) in the various genomes contain disordered regions and these regions are often involved in protein interaction networks; indeed they often act as "hubs" for promiscuous interactions with several different partners (Tompa et al. 2009).

Interactions between two proteins are often mediated by other factors. One example is the reduction of dimensionality of a search (Adam and Delbrück 1968), for example, binding of two proteins to a membrane (Fig. 2a) or DNA will increase the effective on-rate k_{on}. Local concentrations can also be increased by mutual binding to a scaffold protein or by tethering two domains together (Fig. 2b) (Kuriyan and Eisenberg 2007). There has also been increased interest recently in the role of "encounter" complexes that form in a rapid binding step, followed by formation of the final complex (see below and Fig. 2c).

NMR Methods for Detecting Interactions

NMR description of the $P + L \rightleftharpoons PL$ equilibrium: The NMR parameters describing a protein in free (P) and bound (PL) states will in general be different. Key parameters defining the two states include the chemical shifts (δ_P and δ_{PL}) and the relaxation times ($T_{1,2P}$ and $T_{1,2PL}$). The NMR signals that are observed from free and bound states depend on k_{on} and k_{off} in Fig. 1. The exchange regime can be "fast," where one averaged signal is observed for free and bound states, with NMR parameters that are a weighted mean of each state. The exchange regime is said to be "slow" when two non-averaged signals are observed for the free and bound states. "Intermediate" exchange is when the lines are broadened by the exchange process; k_{off} and $(\delta_P - \delta_{PL})$ are key parameters in determining the observed exchange regime, with fast exchange being observed when $k_{off} \gg (\delta_P - \delta_{PL})$. Intermediate exchange often causes detection problems because of broad lines (careful design of experiment can help; Reibarkh et al. 2006) but this regime has the advantage that the exchange parameters, k_{off} etc., can be determined by line-shape analysis (see below). Since protein-protein exchange reactions usually have a relatively small range of diffusion-limited k_{on} values, the observed exchange regime is correlated with affinity; if $K_d = k_{off}/k_{on} > 10^{-5}$ M, the exchange regime is often "fast" while if $K_d < 10^{-6}$ M, the exchange rate is often "slow."

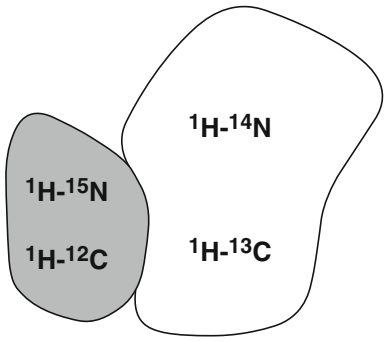


NMR Studies of Macromolecular Interactions – Introduction, Fig. 2 Illustration of protein-protein interactions mediated by (**a**) a membrane; (**b**) another protein and a tethered domain; (**c**) by the initial formation of an encounter complex

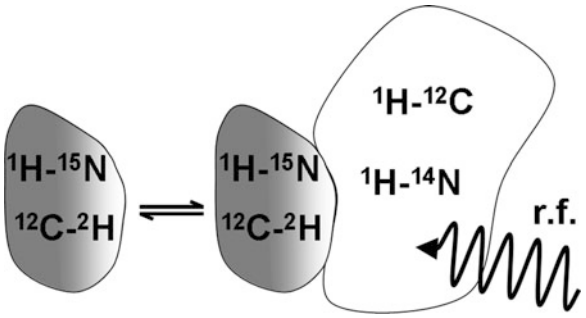
NMR Studies of Macromolecular Interactions – Introduction, Fig. 3 Illustration of the CSP experiment. ^{1}H–^{15}N HSQC spectra of the *light gray* protein give single peaks for each amide proton. On addition of the *black* protein, peaks, *a*–*d*, are perturbed more than the others. These are likely to occur at the protein-protein interface

Chemical shift perturbation (*CSP*): There are powerful two-dimensional NMR experiments that can be used to monitor the shifts of specific groups in a protein. These include the HSQC (heteronuclear single quantum coherence) and TROSY (transverse relaxation optimized spectroscopy) experiments (reviewed in Cavanagh et al. 2007). In a ^{15}N-labeled protein all amino acids, except proline, can give rise to single peaks in ^{1}H–^{15}NHSQC or TROSY spectra. These individual ^{15}N–^{1}H cross peaks can be assigned to specific backbone resonances in the protein with standard experiments on proteins labeled with ^{13}C and ^{15}N. When a ligand is added to a sample, some resonances shift (or broaden) selectively as a result of complex formation. These perturbations can be used to locate the binding sites. A weighted shift plot is often used [e.g., $\sqrt{\{(\Delta\delta_N/5)^2 + \Delta\delta_H{}^2\}}$] that takes account of the different shift ranges of different nuclei (e.g., $\Delta\delta_N$ the induced shift in the ^{15}N dimension is $\sim 5\times$ the shift induced in the ^{1}H dimension, $\Delta\delta_H$). Assigned resonances that experience significant shift perturbation can be mapped on a known structure (Fig. 3).

CSP is used extensively in NMR studies of protein-ligand interactions. In the fast exchange regime, it has the advantage that, as well as the interaction surface, the affinity of the interaction can be obtained from a plot of



NMR Studies of Macromolecular Interactions – Introduction, Fig. 4 Illustration of differential isotope labeling of two proteins. Isotope-filtered pulse sequences can then be used to detect intermolecular NOE selectively, e.g., between groups containing ^{1}H–^{13}C and ^{1}H–^{12}C or ^{1}H–^{15}N and ^{1}H–^{14}N. ^{15}N and ^{13}C have relatively low natural abundance but can be readily incorporated in recombinant proteins expressed in bacteria or yeast. Deuterium (^{2}H) labeling is also often used, especially for larger systems because it can reduce relaxation mechanisms and simplify the spectra



NMR Studies of Macromolecular Interactions – Introduction, Fig. 5 Illustration of cross saturation. An unlabeled larger system is irradiated with radio frequency to saturate the ^{1}H signals. This saturation is transferred to the smaller isotope labeled protein in the complex and residues in the interface region (shown as *white*) have reduced intensity; this reduction is observed in the "free" form of the ^{15}N labeled protein which is present in excess

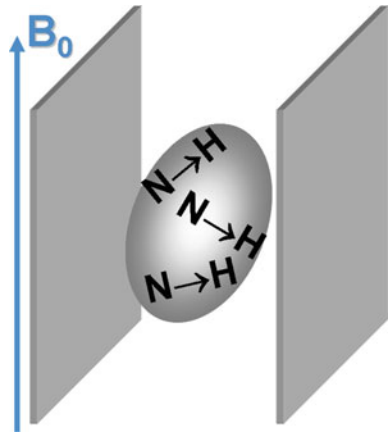
induced shift against concentration of added ligand. A limitation is that the mechanism of perturbation is not generally well understood and induced conformational changes cannot readily be distinguished from surface contacts. CSP is, however, relatively easy to apply and is very informative if assignments are available. The method can be applied to very large systems; examples include the proteasome (Sprangers et al. 2008) and GroEL (Fiaux et al. 2002) and even the formation of protein complexes in a cellular environment (Inomata et al. 2009). CSP also provides useful input in docking procedures (see below).

Nuclear Overhauser effect: A major goal in interaction studies is to define the structure of the complex formed. Experimental restraints can be provided by paramagnetic probes and residual dipolar coupling (see below) but observation of pairwise dipolar interactions between assigned nuclei via the nuclear Overhauser effect (NOE) remains the most potent and precise way to define the structure of a complex. An influential early example of the application of NOEs in structure determination was a study of a calmodulin peptide complex (Ikura et al. 1992). The detection of intermolecular NOEs is greatly facilitated by differential isotope labeling (Fig. 4a). Elegant NMR isotope filtering methods have been developed to differentiate the spectra of labeled and unlabeled protein (Cavanagh et al. 2007) and to detect intermolecular NOEs.

There are now hundreds of examples of structures of protein-protein complexes solved by solution state NMR. NMR has been particularly effective in studies of the disorder/order type of interaction shown in Fig. 1b. An example from our own work is the determination of the structure of a complex between a beta integrin tail and its intracellular talin ligand (Wegener et al. 2007). A notable example of the domain-domain class of interaction is the complex formed between the N-terminal domain of enzyme I and the histidine phosphocarrier protein (Clore and Iwahara 2009); this structure was defined using a wide range of NMR restraints including NOEs, residual dipolar coupling, and paramagnetic centers (see below).

Cross saturation: Another important mapping method is "cross saturation" where a relatively large unlabeled protein is selectively irradiated with radio frequency so that all the ^{1}H resonances are "saturated" (their net magnetization is reduced to zero). Association of an isotope-labeled smaller protein with the irradiated protein can be detected because ^{1}H groups in the interface region also become saturated (or "bleached") by dipolar interactions with the saturated resonances of the larger protein; this results in intensity reductions in specific resonances of the smaller protein (Shimada et al. 2009; Fig. 5). Variations of this approach have also been applied; for example by altering the isotopic labeling patterns in the two proteins (Ferrage et al. 2010).

Kinetics. NMR can also contribute significantly to an understanding of the kinetics of complex formation.



NMR Studies of Macromolecular Interactions – Introduction, Fig. 6 Illustration of residual dipolar coupling effects. Weak alignment of a protein is induced by the introduction of space restrictions within the sample; information about the size of the dipolar coupling, e.g., between ^{15}N and ^{1}H in an N–H group, can then be extracted. This information can be related to the orientation of the N–H bond vector with respect to the direction of the applied field, B_0

The lineshapes of peaks are sensitive to the exchange rates, especially k_{off}, and simulation programs are available to fit the observed shapes to kinetic parameters. An example is a study of the binding of different peptides to an SH2 domain in intermediate exchange (Ladbury et al. 1996). Transfer of magnetization between free and bound states can also be observed in the slow exchange regime using saturation transfer difference spectra and transferred NOEs (Pons et al. 2008). Relaxation dispersion measurements have also become popular (Sugase et al. 2007). In this experiment, T_2 relaxation is measured by a train of 180° pulses; measurement of T_2 as a function of pulse separation gives information about the ligand exchange rates on a timescale that complements other methods. The method can also be used to investigate 'invisible states' in exchanging systems (Hansen et al. 2008).

Mapping with amide-proton exchange. Protein interfaces can be mapped using the protection of amide-proton exchange that arises when protein-protein interactions shield the interface area from solvent (Zuiderweg 2002). However, amide-proton exchange rates are generally slower in a ligand complex than in the free protein and interpretation in terms of a precise interface has to be done with caution.

Residual dipolar coupling: Dipolar interactions between a pair of nuclei lead to relaxation and NOEs, but they also cause a splitting of NMR signals that depends on the internuclear separation and relative orientation of the pair in the applied magnetic field. In solution these dipolar splittings are normally averaged out by rapid tumbling but they can be partially reintroduced by inducing fractional alignment of the molecules in the NMR sample. Weak alignment can be introduced in various ways, including compressed gels, bicelles, and filamentous phage (see e.g., Bax 2003; Liu and Prestegard 2010). The small residual dipolar couplings (RDCs) seen using these weak alignment media (~10 Hz) give useful information about how some groups, e.g., NH, CH, are aligned relative to the external magnetic field. This is very useful in obtaining relative domain orientation using docking programs (see below).

Paramagnetic relaxation and encounter complexes. Paramagnetic probes are very useful tools for structural studies of macromolecules and their complexes (reviewed in Otting 2010). A free electron has a magnetic moment 658× larger than a ^{1}H, so large induced relaxation and chemical shifts can be observed for resonances of groups in the vicinity of a probe. Nitroxide spin labels can be used to explore exposed surfaces (Deschamps et al. 2005), or they can be covalently attached to a single, surface-exposed cysteine introduced into the protein. Paramagnetic metal ions are also very useful, especially with new tags designed to add specific binding sites for lanthanide ions or Mn^{2+}. A notable recent use of paramagnetic probes has been in the study of encounter complexes (Fig. 2c). The Clore group studied interactions between the N-terminal domain of enzyme I and HPr. This 40 kDa complex is involved in the transfer of phosphoryl groups between histidines. EDTA-based Mn^{2+} binding sites were introduced at three separate sites on HPr, and intra- and intermolecular relaxation effects measured (Clore and Iwahara 2009). The intramolecular relaxation data observed for HPr were consistent with the known structure but the intermolecular features indicated the formation of nonspecific encounter complexes. The Ubbink group observed large encounter areas in the electron transfer complex formed between yeast cytochrome *c* and cytochrome *c* peroxidase (Bashir et al. 2010). The significance of these observed encounter complexes is not yet entirely clear but they illustrate, along with the other factors shown in Fig. 2, how biological systems can fine-tune the kinetics and affinity of complex formation.

Docking. In many cases, NMR alone does not provide sufficient information to produce a well-defined

structure of a protein-protein complex; docking calculations can then be very powerful. The backbone of the protein structure is usually assumed to remain unchanged in the complex, but conformational flexibility of the protein side chains is allowed. Docking algorithms contain two main components: a search of all the possible geometries to create a large family of possible protein complex structures followed by an evaluation of these geometries based on energy terms. Several powerful docking programs are now available (e.g., HADDOCK; de Vries et al. 2010, Bonvin this volume); these can incorporate data from CSP, mutagenesis, and other known properties such as symmetry and the proximity of catalytic site residues. Docking protocols are also available that incorporate RDC data (Berlin et al. 2010).

Summary

NMR studies of interactions between proteins and various ligands give not only structural information about the complexes formed but also kinetic, dynamics, and thermodynamic information. A number of powerful NMR methods that can be applied to study complexes are briefly described.

Cross-References

- ▶ Cross Saturation/Transferred Cross Saturation (TCS)
- ▶ HADDOCK
- ▶ Protein Dynamics and Allostery
- ▶ SAR by NMR

References

Adam G, Delbrück M. Reduction of dimensionality in biological diffusion processes. In: Rich A, Davidson N, editors. Structural chemistry and molecular biology. San Francisco: Freeman; 1968. p. 198–215.

Bader S, Kühner S, Gavin AC. Interaction networks for systems biology. FEBS Lett. 2008;582:1220–4.

Bashir Q, Volkov AN, et al. Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome c and cytochrome c peroxidase. J Am Chem Soc. 2010;132:241–7.

Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci. 2003;12:1–16.

Berlin K, O'Leary DP, Fushman D. Structural assembly of molecular complexes based on residual dipolar couplings. J Am Chem Soc. 2010;132:8961–72.

Bonvin AM, Boelens R, Kaptein R. NMR analysis of protein interactions. Curr Opin Chem Biol. 2005;9:501–8.

Cavanagh J, Fairbrother W, et al. Protein NMR spectroscopy. 2nd ed. San Diego: Academic; 2007.

Clore GM, Iwahara J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev. 2009;109:4108–39.

de Vries SJ, van Dijk M, Bonvin AM. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc. 2010;5:883–97.

Deschamps ML, Pilka ES, et al. Probing protein-peptide binding surfaces using charged stable free radicals and transverse paramagnetic relaxation enhancement (PRE). J Biomol NMR. 2005;31:155–60.

Ferrage F, Dutta K, et al. Structural determination of biomolecular interfaces by nuclear magnetic resonance of proteins with reduced proton density. J Biomol NMR. 2010;47:41–54.

Fiaux J, Bertelsen EB, et al. NMR analysis of a 900 K GroEL GroES complex. Nature. 2002;418:207–11.

Hansen DF, Vallurupalli P, Kay LE. Using relaxation dispersion NMR Spectroscopy to determine structures of excited, invisible protein states. J Biomol NMR. 2008;41:113–20.

Ikura M, Clore GM, et al. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. Science. 1992;256:632–8.

Inomata K, Ohno A, et al. High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. Nature. 2009;458:106–9.

Kuriyan J, Eisenberg D. The origin of protein interactions and allostery in colocalization. Nature. 2007;450:983–90.

Ladbury JE, Hensmann M, et al. Alternative modes of tyrosyl phosphopeptide binding to a Src family SH2 domain: implications for regulation of tyrosine kinase activity. Biochemistry. 1996;35:11062–9.

Liu Y, Prestegard JH. A device for the measurement of residual chemical shift anisotropy and residual dipolar coupling in soluble and membrane-associated proteins. J Biomol NMR. 2010;47:249–58.

O'Connell MR, Gamsjaeger R, Mackay JP. The structural analysis of protein-protein interactions by NMR spectroscopy. Proteomics. 2009;9:5224–32.

Otting G. Protein NMR using paramagnetic ions. Annu Rev Biophys. 2010;39:387–405.

Pons J, Evrard-Todeschi N, et al. NMR and docking studies identify the binding of the peptide derived from activating transcription factor 4 to protein ubiquitin ligase beta-TrCP. Competition STD-NMR with beta-catenin. Biochemistry. 2008;47:14–29.

Reibarkh M, Malia TJ, et al. Identification of individual protein-ligand NOEs in the limit of intermediate exchange. J Biomol NMR. 2006;36:1–11.

Shimada I, Ueda T, et al. Cross-saturation and transferred cross-saturation experiments. Prog Nucl Magn Reson Spectrosc. 2009;54:123–40.

Sprangers R, Li X, et al. TROSY-based NMR evidence for a novel class of 20S proteasome inhibitors. Biochemistry. 2008;47:6727–34.

Sugase K, Lansing JC, et al. Tailoring relaxation dispersion experiments for fast-associating protein complexes. J Am Chem Soc. 2007;129:13406–7.
Tompa P, Fuxreiter M, et al. Close encounters of the third kind: disordered domains and the interactions of proteins. Bioessays. 2009;31:328–35.
Wegener KL, Partridge AW, et al. Structural basis of integrin activation by talin. Cell. 2007;128:171–82.
Zuiderweg ER. Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry. 2002;41:1–7.

Cross-References

▶ Macromolecular Crystallography: Overview
▶ NMR
▶ NMR-based Structural Proteomics
▶ Protein NMR – Introduction
▶ Protein Structure Comparison Methods
▶ Structural Genomics

NMR-based Structural Proteomics

Weontae Lee
Department of Biochemistry, Yonsei University, Seoul, South Korea

Definition

Unlike conventional proteomics, structural proteomics deduces biological function of unknown gene products from the three-dimensional structure of a protein by structural biology, bioinformatics, molecular biology, and biochemistry. The success of structural proteomics relies on that protein functions are related to their three-dimensional structures, providing crucial information about its biological functions. In this regard, structural proteomics approach is one of the powerful research areas in the post-genomic era, promising putative biochemical and cellular functions of a number of uncharacterized gene products. Recently, a number of pioneering groups have successfully predicted their biological functions based on three-dimensional structures of hypothetical proteins. The tools of structural proteomics are a number of structure determination techniques, including NMR spectroscopy. Especially, advances in hardware design, data acquisition methods, sample preparation, and automation of data analysis in NMR spectroscopy enable high-throughput structure determination of proteins in physiological condition. These efforts ensure that NMR spectroscopy will become an important methodology in structural genomics. Therefore, NMR-based structural proteomics will deposit a comprehensive structural database to lead the biological functions of hypothetical proteins identified by the human genome project.

Noncanonical Hydrogen Bonds

▶ XH/π Interactions

Nonlinear IR Spectroscopy of Proteins

▶ Ultrafast Two-Dimensional Infrared Spectroscopy of Proteins

Nonlinear Raman Spectroscopy: Coherent Anti-Stokes Raman Scattering (CARS)

Mikhail N. Slipchenko[1] and Ji-Xin Cheng[2]
[1]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA
[2]Department of Chemistry, Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

Synonyms

Raman spectroscopy

Definition

Nonlinear Raman spectroscopy is a spectroscopy based on nonlinear effects and involving Raman-active molecular transitions.

Coherent anti-Stokes Raman scattering is generated in samples due to third-order nonlinear susceptibility and requires three laser beams to produce a signal beam.

Introduction

Coherent anti-Stokes Raman scattering (CARS) provides analytical and spectroscopic information pertaining to Raman-active resonances in gases, liquids, and solids. Unlike second-order processes, CARS is a third-order nonlinear optical effect and, therefore, is applicable to isotropic, as well as anisotropic media. The signal due to CARS processes is considerably higher than spontaneous Raman scattering. Moreover, the CARS signal is spatially and temporally coherent and propagates like a laser beam. Collectively, these properties make CARS a superior tool for obtaining spectra of luminescent samples including fluorescent materials, impurities, combustion systems, and discharge systems.

In spite of many advantages, CARS has its drawbacks. One of the most serious disadvantages of CARS as an analytical tool is radiation due to nonresonant part of the third-order susceptibility which limits detection sensitivity of the technique. Strong perturbation of CARS spectrum due to interference between neighboring resonances, background and electronic transitions is another drawback which complicates spectral analysis.

Over the past years, the CARS technique has found two important applications, namely, combustion diagnostics and CARS microscopy.

The CARS Phenomenon and Theoretical Description of CARS Spectroscopy

The phenomenon of coherent Raman scattering was first reported by scientists at the Ford Motor Company in 1965 (Maker and Terhune 1965). Since its introduction in the early 1970s, CARS spectroscopy has been widely used as a spectroscopic tool (Levenson and Kano 1988). CARS is a four-wave mixing process in which a pump field $E(\omega_1)$, a Stokes field $E(\omega_2)$, and a probe field $E(\omega_3)$ interact with a sample and generate an anti-Stokes field $E(\omega_4)$ at the frequency of $\omega_4 = \omega_1 - \omega_2 + \omega_3$.

The energy diagrams of CARS are shown in Fig. 1. The CARS signal arises from the nonlinear polarization induced along a Cartesian coordinate i: (Boyd 1992)

$$P_i^{(3)}(\omega_4) = D\sum_{jkl}\chi_{ijkl}^{(3)}(-\omega_4;\omega_1;-\omega_2;\omega_3) E_j(\omega_1)E_k^*(\omega_2)E_l(\omega_3), \quad (1)$$

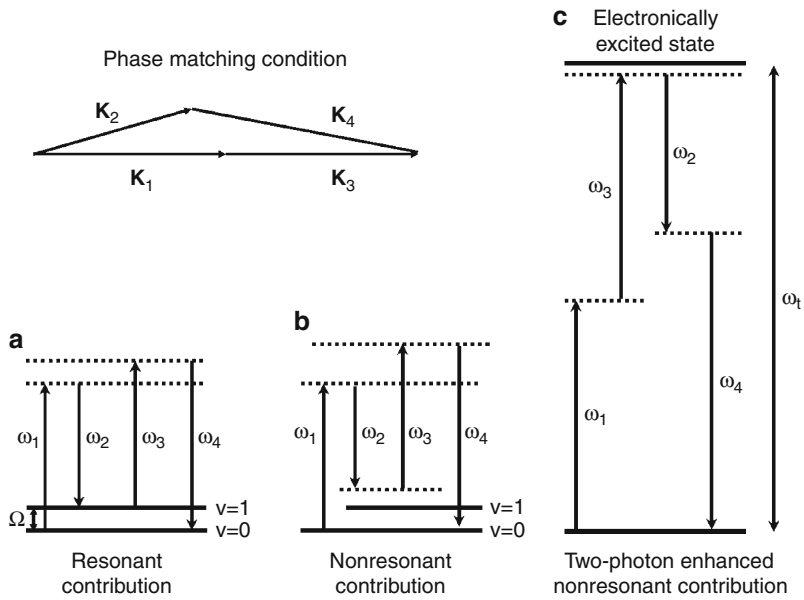
where $\chi_{ijkl}^{(3)}$ is the third-order susceptibility, which in an isotropic material can be expressed in the form:

$$\chi_{ijkl}^{(3)} = \chi_{1122}^{(3)}\delta_{ij}\delta_{kl} + \chi_{1212}^{(3)}\delta_{ik}\delta_{jl} + \chi_{1221}^{(3)}\delta_{il}\delta_{jk} \quad (2)$$

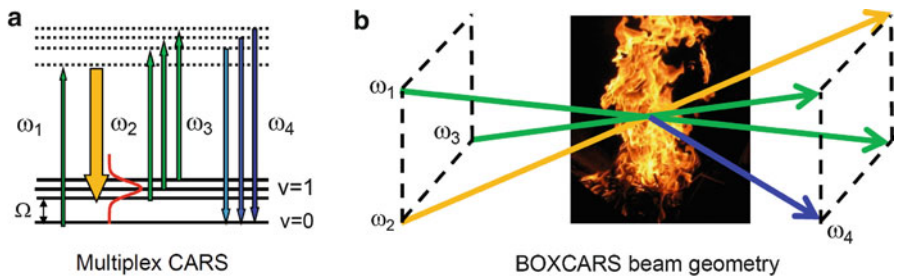
A general expression for the third-order susceptibility can be written as:

$$\chi^{(3)} = \frac{A_R}{\Omega - (\omega_p - \omega_s) - i\Gamma_R} + \chi_{nr}^{(3)} + \frac{A_t}{\omega_t - 2\omega_p - i\Gamma_t} \quad (3)$$

where Ω is the vibrational frequency; Γ_R and Γ_t are the half width at half maximum of the Raman line and that of the two-photon transition, respectively; A_R and A_t are constants representing the Raman scattering and the two-photon absorption cross sections. The first term in (3) is a vibrationally resonant contribution (Fig. 1a). The second term is a nonresonant contribution that is independent of Raman shift (Fig. 1b). The third term is an enhanced nonresonant contribution due to two-photon electronic resonance (Fig. 1c). In most experiments, the pump and probe fields are derived from the same laser beam, $E(\omega_1) \equiv E(\omega_3)$. Due to momentum conservation, CARS signal generation needs to fulfill the phase-matching condition, $|\mathbf{k}_4 - (2\mathbf{k}_1 - \mathbf{k}_2)|l << \pi$, where l is the interaction length and $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_4$ are wave vectors of the pump, Stokes, and CARS fields, respectively ($|\mathbf{k}| = \omega n/c$, where c is a speed of light and n is a refractive index) (Shen 1984; Levenson and Kano 1988; Mukamel 1995). Due to dispersion, especially in the condense phase samples, the collinear beam propagation results in small interaction length and the incident beams must be properly aligned to satisfy phase-matching conditions. Such spatial alignment also helps to spatially separate CARS signal beam from powerful input laser beams, which otherwise might leak into detector.



Nonlinear Raman Spectroscopy: Coherent Anti-Stokes Raman Scattering (CARS), Fig. 1 Energy diagram of CARS. (**a**) Resonant CARS characterized by $A_R/[\Omega - (\omega_1 - \omega_2) - i\Gamma_R]$. (**b**) Nonresonant CARS from an electronic contribution ($\chi^{(3)}_{nr}$), where the dotted lines indicate virtual states. (**c**) Electronic contribution ($A_t/[\omega_t - 2\omega_1 - i\Gamma_t]$) enhanced by a two-photon resonance of the pump beam associated with an excited electronic state



Nonlinear Raman Spectroscopy: Coherent Anti-Stokes Raman Scattering (CARS), Fig. 2 Multiplexed detection and BOXCARS beam geometry for combustion research

Application of CARS Spectroscopy to Combustion Studies

The application of CARS spectroscopy to study combustion was first demonstrated in 1973 by Taran and coworkers (Regnier and Taran 1973). In the following years, three major breakthroughs made the technique applicable to real systems: (1) development of multiplex CARS with broadband dye lasers which allowed single pulse measurements; (2) introduction of the BOXCARS beam geometry for better spatial resolution; and (3) adaptation of a technique for elimination of nonresonant background. The initial implementations of CARS were based on nanosecond (ns) laser systems and used a narrowband ($\sim$0.001 cm^{-1}) pump laser and a broadband ($\sim$150 cm^{-1}) Stokes laser to excite entire manifold of rovibrational transitions (see Fig. 2a).

For temperature measurements, nitrogen is a molecule of choice due to its abundance in air-fed reacting flows. To excite N_2, which has vibrational band origin at 2,330 cm^{-1}, the narrowband second harmonic from a Nd:Yag laser (at 532 nm), and the broadband emission from dye laser at around 607 nm can be used as a pump and Stokes beams. The pump beam is usually split in two in order to arrange beams in BOXCARS configuration (see Fig. 2b). Detailed review of the application of ns-CARS spectroscopy to combustion diagnostics is provided in Ref. Eckbreth (1996). The later development of ns-CARS has been focused on simultaneous detection of multiple species along with temperature and its applications to harsh environments. In order to excite rovibrational transitions in pairs of molecules such as [N_2, CO_2], [N_2, CH_4], [O_2, CO], [N_2, O_2], and [N_2, H_2], the dual-pump CARS

was developed. Later this technique was extended to triple-pump CARS, dual-pump dual-broadband CARS, and dual-broadband rotational CARS. For detailed review, see Ref. Roy et al. (2010).

With the introduction of commercially available compact picosecond (ps) and femtosecond (fs) lasers systems, ps- and fs-CARS measurements within time scales free from collisional effects became possible. The reduction of nonresonant background in ns-CARS is employed via various polarization schemes; however, these methods are ineffective in high pressure, high temperature environments of real combustors due to window-induced birefringence. Contrarily, ultrafast pulses allow to use time domain to eliminate nonresonant background due to different time dynamics of resonant and nonresonant signals. In addition, extremely high peak power of fs lasers leads to high signal levels that allow a single shot detection of CARS signal at kHz rates compared to ~10 Hz rates achievable with ns-CARS. Detailed review of application of fs-CARS spectroscopy to combustion can be found in Refs. Gord et al. (2008) and Roy et al. (2010).

CARS Microscopy

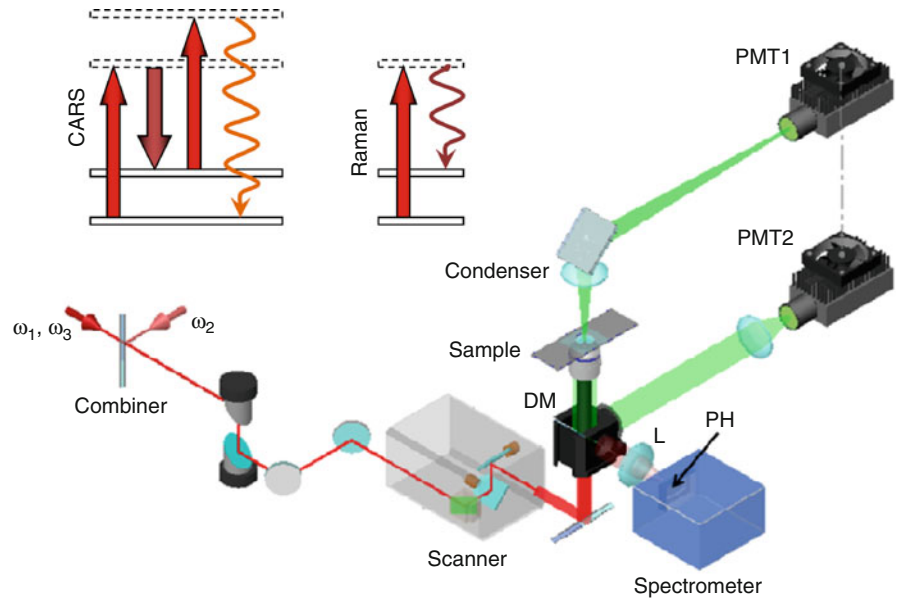
The application of the CARS process to microscopy opens up a new method for chemical imaging. In CARS microscopy, the temporally and spatially overlapped pump and Stokes laser pulses are tightly focused into a sample to generate a signal in a small excitation volume (<1 μm^3). A CARS image is acquired by raster scanning the sample or the laser beams. The vibrational contrast in CARS microscopy arises from the signal enhancement when $\omega_p - \omega_s$ is tuned to a Raman-active vibrational band. CARS microscopy offers the following advantages over traditional techniques: (1) *Chemical bond selectivity*. Because the contrast in CARS microscopy arises from intrinsic molecular vibrations, it is possible to produce images with chemically selective contrasts when the laser frequency difference ($\omega_p - \omega_s$) is tuned to different Raman bands. (2) *Large signal level*. As a coherent process, the CARS fields from different molecules have a well-defined phase relationship. The coherent addition results in a quadratic signal increase with respect to the density of molecular oscillators, versus the linear increase of spontaneous Raman signal. For a bulk liquid, the CARS signal can be larger than the spontaneous Raman signal by 9 orders of magnitude (Levenson and Kano 1988). (3) *High imaging speed*. The large signal level in CARS microscopy enables high-speed imaging which is important for live cell and tissue studies. In laser-scanning microscopy with high repetition rate (in MHz range) lasers, an image acquisition speed at one frame per second has been achieved (Cheng et al. 2002a) and has been recently increased to 20 frames per second (Evans et al. 2005). The high imaging speed not only avoids respiration-induced image distortion during in vivo imaging, but also allows real-time inspection of highly dynamic systems. (4) *Submicron 3D resolution.* The nonlinear optical excitation ensures that the signal is only generated at the center of the focus, offering CARS microscopy an inherent 3D spatial resolution. The lateral and axial resolutions with a 60X water immersion objective (numerical aperture, NA = 1.2) were measured to be 0.23 and 0.75 μm,(Cheng et al. 2002a) which allows detection of subcellular structures in a tissue environment. (5) *Multimodal imaging*. A unique advantage of CARS is that other NLO imaging modalities, including two-photon excited fluorescence (TPEF) and second harmonic generation (SHG), can be naturally incorporated into the same platform, enabling real-time visualization of various components in a tissue.

A modern high-speed, multifunctional CARS microscope (Fig. 3) for biological applications originated from multistep technical advances (Cheng and Xie 2004; Evans and Xie 2008) and theoretical study (Cheng et al. 2002b) of CARS microscopy.

Important features of the microscope are:

Collinear beam geometry. Noncollinear beam geometry is used in CARS spectroscopy to fulfill the phase-matching condition (Levenson and Kano 1988). When the interaction length becomes very small (~1 μm) under the tight focusing condition in microscopy, the phase-matching condition is fulfilled with a collinear beam geometry for forward-detected CARS (F-CARS) (Zumbusch et al. 1999). The collinear beam geometry greatly simplifies the optical alignment and is a key step in producing high-quality CARS images.

Laser-scanning on a confocal microscope platform. The most significant advantage of CARS over spontaneous Raman is large signal level of CARS allowing for high-speed vibrational imaging. This advantage has been realized by scanning pulsed lasers of high repetition rates (in MHz), which resulted in an image



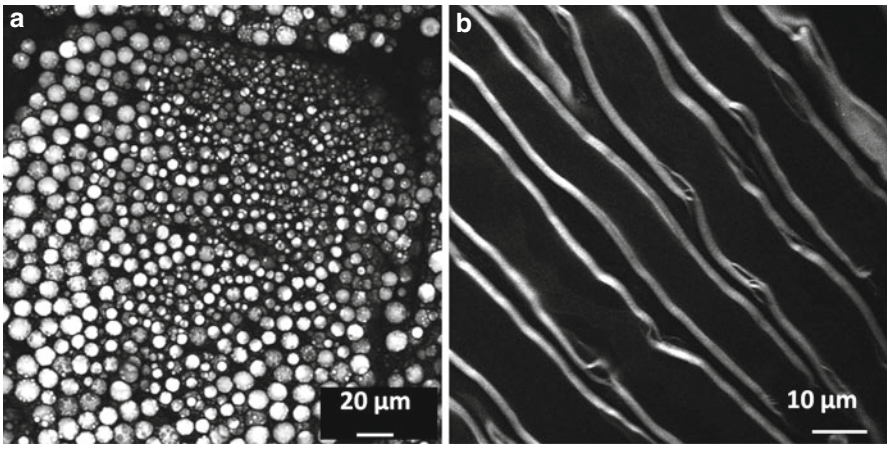
Nonlinear Raman Spectroscopy: Coherent Anti-Stokes Raman Scattering (CARS), Fig. 3 A high-speed CARS microscope. Pump and Stokes laser beams are collinearly combined, directed into a laser-scanning box, and focused into the samples. Emitted photons are collected by a condenser for forward-detected CARS (F-CARS) or by the focusing objective for epi-detected CARS (E-CARS). PMT1 and PMT2 are detectors for F-CARS and E-CARS, respectively. The pump laser is also used for generating spontaneous Raman signals from an object visualized by CARS. A spectrometer with a pinhole attached to the side port of the microscope is used for confocal recording of spontaneous Raman spectra. Although not shown, sum-frequency generation (SFG) and two-photon excited fluorescence (TPEF) imaging can be performed on the same platform. *PH* pinhole, *DM* dichroic mirror, *L* lense, *PMT* photomultiplier tube

acquisition speed of a frame per second (Cheng et al. 2002a). Video rate imaging can be realized by scanning the laser with a polygon mirror or a resonant scanner.

Near IR laser excitation. Using the near IR lasers allows to avoid two-photon electronic enhancement of the nonresonant background, reduces the photodamage induced by multiphoton absorption, and diminishes tissue scattering leading to increased optical penetration depth.

Forward-detection versus *epi-* (i.e., *backward-) detection*. For objects that are comparable or larger than the wavelengths of incident lasers, the CARS signal goes forward as a result of constructive addition of electric fields. Importantly, because the CARS signal is highly directional in the forward direction, an air condenser is sufficient for collection of the F-CARS signal. For example, F-CARS is suitable for imaging lipid bodies or myelin sheath in cultured cells or sliced tissues. For objects with an axial length much smaller than the excitation wavelengths, the phase-matching condition is fulfilled in both forward and backward directions. This backward propagating CARS signal from small objects constitutes the first mechanism for epi-detected CARS (E-CARS). E-CARS also arises from the discontinuity of $\chi^{(3)}$ at the interface, and backscattering of forward CARS photons. E-CARS is important for live animal imaging where the epi-detected signal arises from backscattering of the forward propagating CARS photons by the tissue.

Multimodal NLO imaging. An important advantage of CARS microscope is that other NLO modalities including two-photon excited fluorescence (TPEF) and sum-frequency generation (SFG) can be implemented on the same platform. Multimodality is important because different NLO imaging methods have their distinctive advantages: TPEF can be used to visualize proteins, ions with fluorescent labeling, or specific autofluorescent structures; SFG is selective to non-centrosymmetric molecular assemblies such as collagen fibrils or crystals; and CARS is naturally sensitive to lipid-enriched structures such as adipocytes. The pulsed laser beams for CARS could be simultaneously used for TPEF and electronic SFG imaging. As a two-beam modality,



Nonlinear Raman Spectroscopy: Coherent Anti-Stokes Raman Scattering (CARS), Fig. 4 CARS images of (**a**) cytoplasmic lipid droplets in the small intestine cells and (**b**) axonal myelin sheath in spinal white matter. The CARS signal arises from the symmetric CH_2 stretch vibration

CARS microscopy is mostly operated with picosecond (ps) pulses, either from two synchronized Ti:sapphire lasers or from a synchronously pumped OPO system. Picosecond pulse excitation not only provides sufficient spectral resolution, but also increases the ratio of resonant signal to nonresonant background. Although tunable ps laser systems operating in the NIR range are widely accepted for high-speed CARS imaging (for reviews see Cheng and Xie 2004; Cheng 2007; Evans and Xie 2008), the reduced efficiency of nonlinear optical (NLO) process caused by longer pulse duration hinders wide application of ps lasers to TPEF and SHG imaging. To overcome above limitation of ps-CARS, high-speed CARS imaging of lipids with fs laser pulses has been demonstrated. Commercial CARS microscopes with either ps or fs laser sources are now available through Leica and Olympus, respectively.

Coupling CARS imaging with spontaneous Raman spectroscopy. It has been realized that CARS and spontaneous Raman have their distinctive advantages: CARS permits high-speed imaging by focusing the energy on a single Raman band while spontaneous Raman permits complete fingerprint analysis at a pixel of interest. Based on this understanding, a compound Raman microscope in which the same ps laser source is used for high-speed coherent Raman imaging and confocal Raman spectral analysis has been developed and applied for analysis of lipid-rich structures such as sebaceous gland in mice (Slipchenko et al. 2009) and lipid species in *Caenorhabditis elegans* (Le et al. 2010a).

CARS microscopy has been demonstrated to be highly sensitive and selective to lipid-rich structures that are abundant of CH_2 groups. In particular, CARS microscopy permits chemically selective imaging of two important lipid-rich structures, cytoplasmic lipid droplets, and axonal myelin sheath (Fig. 4). With such capability, CARS microscopy has found broad applications in the field of lipid biology in recent years (Le and Cheng 2009; Wang et al. 2009; Le et al. 2010b).

Summary

CARS spectroscopy is a mature analytical tool. The applications of ns-, ps-, and fs-CARS spectroscopy to gas-phase reacting flows enabled noninvasive measurements of temperature and concentrations of major species present in the reacting flows with accuracy on the order of few percent. The CARS microscopy is based on chemical selectivity of CARS and offers a high-speed imaging and inherent 3D resolution as well as ability to be coupled with other imaging modalities on the same platform.

Cross-References

- ▶ Sum Frequency Generation Vibrational Spectroscopy
- ▶ Two-Photon Excitation Fluorescence STED

References

Boyd RW. Nonlinear optics. New York: Academic; 1992.

Cheng JX. Coherent anti-Stokes Raman scattering microscopy. Appl Spectrosc. 2007;61(9):197A–206.

Cheng JX, Xie XS. Coherent anti-Stokes Raman scattering microscopy: instrumentation, theory, and applications. J Phys Chem B. 2004;108:827–40.
Cheng JX, Jia YK, et al. Laser-scanning coherent anti-Stokes Raman scattering microscopy and applications to cell biology. Biophys J. 2002a;83(1):502–9.
Cheng JX, Volkmer A, et al. Theoretical and experimental characterization of coherent anti-Stokes Raman scattering microscopy. J Opt Soc Am B. 2002b;19:1363–75.
Eckbreth AC. Laser diagnostics for combustion temperature and species. Amsterdam: Gordon & Breach; 1996.
Evans CL, Xie XS. Coherent anti-Stokes Raman scattering microscopy: chemically selective imaging for biology and medicine. Annu Rev Anal Chem. 2008;1:883–909.
Evans CL, Potma EO, et al. Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy. Proc Natl Acad Sci USA. 2005;102(46):16807–12.
Gord JR, Meyer TR, et al. Applications of ultrafast lasers for optical measurements in combusting flows. Annu Rev Anal Chem. 2008;1(1):663–87.
Le TT, Cheng JX. Non-linear optical imaging of obesity-related health risks: review. J Innov Opt Health Sci. 2009;2:1–17.
Le TT, Duren HM, et al. Label-free quantitative analysis of lipid metabolism in living *Caenorhabditis elegans*. J Lipid Res. 2010a;51:672–7.
Le TT, Yue S, et al. Shedding new light on lipid biology by CARS microscopy. J Lipid Res. 2010b;51(11):3091–102.
Levenson MD, Kano SS. Introduction to nonlinear laser spectroscopy. San Diego: Academic; 1988.
Maker PD, Terhune RW. Study of optical effects due to an induced polarization third order in the electric field strength. Phys Rev. 1965;137:A801–18.
Mukamel S. Principles of nonlinear optical spectroscopy. New York: Oxford University Press; 1995.
Regnier PR, Taran JPE. Possibility of measuring gas concentrations by stimulated anti-Stokes scattering. Appl Phys Lett. 1973;23(5):240–2.
Roy S, Gord JR, et al. Recent advances in coherent anti-Stokes Raman scattering spectroscopy: fundamental developments and applications in reacting flows. Prog Energ Combust Sci. 2010;36(2):280–306.
Shen YR. The principles of nonlinear optics. New York: Wiley; 1984.
Slipchenko MN, Le TT, et al. High-speed vibrational imaging and spectral analysis of lipid bodies by compound Raman microscopy. J Phys Chem B. 2009;113:7681–6.
Wang HW, Fu Y, et al. Chasing lipids in health and diseases by coherent anti-Stokes Raman scattering microscopy. Vib Spectrosc. 2009;50:160–7.
Zumbusch A, Holtom GR, et al. Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering. Phys Rev Lett. 1999;82:4142–5.

Nonuniform Sampling

► Sparse Sampling in NMR

Nonvesicular Transport

► Lipid Trafficking in Cells

Non-voltage-Gated Cation Channels

► ENAC, Degenerins, ASICs, and Related Channels

Normal Modes of Molecules

► Molecular Vibrations and Their Interaction with Electromagnetic Radiation

NOS

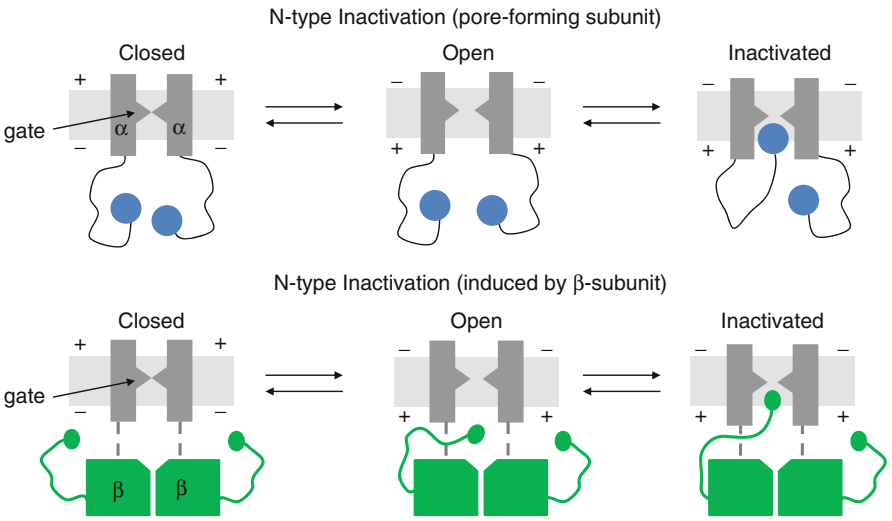
► Nitric Oxide Synthase - Computational Studies

N-Type Inactivation in Voltage-Gated Potassium Channels

KeWei Wang[1,2] and Jijie Chai[3]
[1]Department of Molecular & Cellular Pharmacology, Peking University School of Pharmaceutical Sciences, Beijing, China
[2]Department of Neurobiology, Peking University School of Medicine, Beijing, China
[3]National Institute of Biological Sciences, Beijing, China

Synonyms

Ball and chain inactivation; KChIP; Kv channel-interacting protein; Potassium channels, K channels; Tetramerization domain, T1 domain; Voltage-gated potassium channels, Kv channels



N-Type Inactivation in Voltage-Gated Potassium Channels, Fig. 1 N-type inactivation of voltage-gated K^+ channels conferred by either the N-terminus of the pore-forming α-subunit (*top panel*) or the N-terminus of β-subunits (*lower panel*). In each case the first 20 amino on the N-terminus form a ball-peptide that binds within the pore once the channel has opened – blocking conduction and inactivating the channel

Inactivation in Kv Potassium Channels

Voltage-gated K^+ channels (Kv) close rapidly by a dynamic process known as inactivation, which is a key component of channel function. Inactivating K^+ channels modulate membrane excitability. N-type inactivation involving the N (amino)-terminus of the channel (hence the name "N-type inactivation") was initially found in Shaker-related K^+ channels, the inactivation of which can be removed by intracellular enzymes or the N-terminal deletion of the α-subunit channel protein (Hoshi et al. 1990; Zagotta et al. 1990). It was later confirmed that rapid inactivation could also be conferred by a synthetic peptide having the same sequence as the proximal N-terminus of the α-subunit or the N-terminus of the β-subunit (Fig. 1). These experimental studies were unified with the development of the "ball-and-chain" model in which a positively charged inactivation particle (the ball) on a tether (the chain) binds and blocks the channel on the intracellular side of the membrane (Armstrong and Bezanilla 1977).

The inactivation ball, also known as the inactivation gate, responsible for closing of the channel, is formed by the first 20 amino acids on the cytoplasmic N-terminus of the channel and the contiguous 40 or more residues constitute the chain. The inactivation ball possesses two essential chemical characteristics. The first 10 residues are hydrophobic, and the remaining 10 are hydrophilic. The chain segment, which is rich in positive charge, pushes the ball toward its receptor site when the membrane is depolarized (Zhou et al. 2001). It has been shown that this N-type inactivation occurs through a sequential reaction of a dynamic nature in which the gate initially binds to the T1 domain surface by electrostatic interactions (Gulbis et al. 2000) and then enters through the lateral portals to the inner pore and occludes ion movement (Zhou et al. 2001). K^+ influx from extracellular elevation of extracellular potassium expels the inactivation gate from its binding site, which results in accelerated recovery form inactivation. The T1 domain which directly precedes the first transmembrane domain (S1) of Kv α-subunit is a highly conserved cytoplasmic portion consisting of ~130 amino acids responsible for driving and regulating the tetramerization of the pore-forming α-subunit. In Shaker-type K^+ channels, rapid N-type inactivation is followed by a slower C-type inactivation which involves a constriction of the extracellular mouth of the channel pore (Yellen et al. 1994, see also ▶ Potassium Channel Selectivity and Gating at the Selectivity Filter: Structural Basis).

Kv4 channels inactivate rapidly and recover quickly from inactivation. The T1 domain of the Kv4 N-terminus features intersubunit Zn^{2+} binding sites (Scannevin et al. 2004). The structure of the T1 domain includes four zinc ions bound to the T1 tetramer, and each Zn^{2+} located at the interface between two adjacent monomers is involved in intersubunit contacts. The Kv4 N-terminus is involved in a fast inactivation with features that manifest N-type inactivation. However, N-terminal deletions of Kv4 channels only partially eliminate or slow the fast phase of inactivation, rather than resulting in a complete loss of inactivation as observed in the Shaker channel, suggesting that Kv4 inactivation may differ from classical Shaker N-type inactivation.

N-Type Inactivation in Voltage-Gated Potassium Channels,

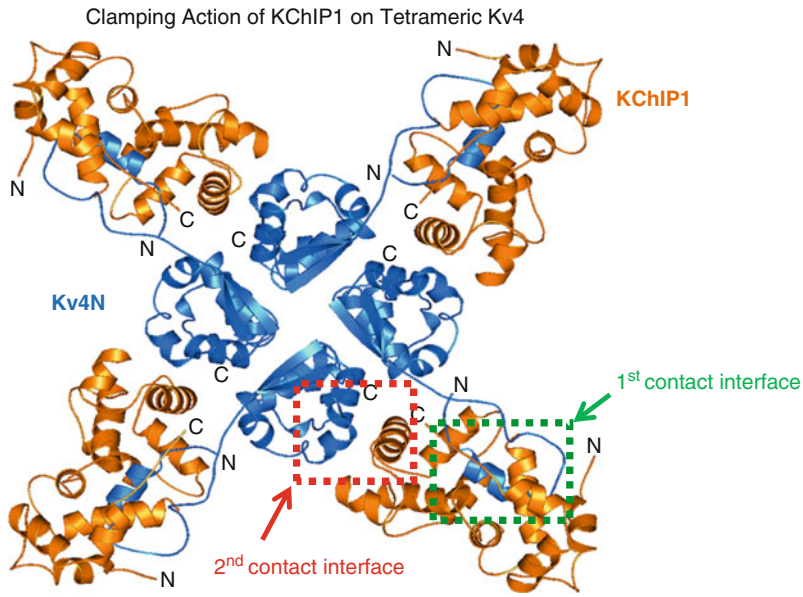


Fig. 2 The co-crystal structure of the Kv4.3 N-terminus/KChIP1complex, revealing the clamping action formed by the KChIPs and how the KChIPs interact with the N-terminus of Kv4.3 subunit to suppress channel inactivation

Cytosolic Kv channel-interacting proteins (KChIPs, 216 ~ 256 amino acids) that belong to the neuronal calcium sensor (NCS) family of calcium binding EF-hand proteins co-assemble with Kv4 (Shal) subunits to form a native complex that encodes somatodendritic A-type K^+ current, I_{SA}, in neurons and transient outward current, I_{TO}, in cardiac myocytes (An et al. 2000). The specific binding of KChIPs to the Kv4 N-terminus enhances surface expression, facilitates subunit assembly and regulates gating properties of Kv4 channels. The co-crystal structure of the Kv4.3 N-terminus/KChIP1complex reveals there is a clamping mode in which a single KChIP1 molecule acting as a monomer laterally clamps two neighboring Kv4.3 N-termini in a 4:4 manner, with two contact interfaces being involved in the interaction (Pioletti et al. 2006; Wang et al. 2007). Within the first interface, the proximal N-terminal hydrophobic peptides of Kv4.3 reach out into a deep elongated groove and are sequestered by binding to the deep hydrophobic pocket on the surface of KChIP1, resulting in physical suppression of Kv4 inactivation by KChIP1 (Fig. 2).

Cross-References

- ▶ Potassium Channel Selectivity and Gating at the Selectivity Filter: Structural Basis
- ▶ Potassium Channels: Their Physiological and Molecular Diversity

References

An WF, Bowlby MR, Betty M, Cao J, Ling HP, Mendoza G, Hinson JW, Mattsson KI, Strassle BW, Trimmer JS, et al. Modulation of A-type potassium channels by a family of calcium sensors. Nature. 2000;403:553–6.

Armstrong CM, Bezanilla F. Inactivation of the sodium channel. II. Gating current experiments. J Gen Physiol. 1977;70:567–90.

Gulbis JM, Zhou M, Mann S, Mackinnon R. Structure of the cytoplasmic beta subunit – T1 assembly of voltage-dependent K^+ channels. Science. 2000;289:123–7.

Hoshi T, Zagotta WN, Aldrich RW. Biophysical and molecular mechanisms of Shaker potassium channel inactivation. Science. 1990;250:533–8.

Pioletti M, Findeisen F, Hura GL, Minor Jr DL. Three-dimensional structure of the KChIP1-Kv4.3 T1 complex reveals a cross-shaped octamer. Nat Struct Mol Biol. 2006;13:987–95.

Scannevin RH, Wang K, Jow F, Megules J, Kopsco DC, Edris W, Carroll KC, Lu Q, Xu W, Xu Z, et al. Two N-terminal domains of Kv4 K(+) channels regulate binding to and modulation by KChIP1. Neuron. 2004;41:587–98.

Wang H, Yan Y, Liu Q, Huang Y, Shen Y, Chen L, Chen Y, Yang Q, Hao Q, Wang K, et al. Structural basis for modulation of Kv4 K^+ channels by auxiliary KChIP subunits. Nat Neurosci. 2007;10:32–9.

Yellen G, Sodickson D, Chen TY, Jurman ME. An engineered cysteine in the external mouth of a K^+ channel allows inactivation to be modulated by metal binding. Biophys J. 1994;66:1068–75.

Zagotta WN, Hoshi T, Aldrich RW. Restoration of inactivation in mutants of Shaker potassium channels by a peptide derived from ShB. Science. 1990;250:568–71.

Zhou M, Morais-Cabral JH, Mann S, MacKinnon R. Potassium channel receptor site for the inactivation gate and quaternary amine inhibitors. Nature. 2001;411:657–61.

Nuclear Magnetic Resonance Imaging

▶ Magnetic Resonance Imaging (MRI) Methodology

Nuclear Overhauser Effect

Patrik Lundström
Division of Molecular Biotechnology, Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden

Synonyms

NMR

Introduction

The nuclear Overhauser effect (NOE) is the transfer of nuclear spin polarization from one nuclear spin population to another via dipolar cross-relaxation. The phenomenon has been known since the early days of NMR and is a generalization of Albert Overhauser's proposal that nuclear spins could be polarized by irradiation of electron spins (Overhauser 1953), which was verified experimentally the same year (Carver and Slichter 1953). The theoretical framework for also describing transfer of polarization to one nuclear spin population from another such population was laid down two years later (Solomon 1955).

Theory

Spin-lattice relaxation for interacting spins can be treated theoretically by considering the rates of transitions between energy levels, W_i (Bloembergen et al. 1948). If the eigenstates for a two-spin system are labeled $|m_I m_S\rangle$, the energy level diagram of Fig. 1 can be drawn.

The different processes are denoted as flip transitions if $\Delta m_I = \pm 1$ or $\Delta m_S = \pm 1$, flip-flip transitions if $\Delta m_I + \Delta m_S = \pm 2$, and flip-flop transitions if $\Delta m_I + \Delta m_S = 0$, $\Delta m_I \neq 0$. The differential equation that governs the population of state $|\alpha\alpha\rangle$, $P_{\alpha\alpha}$, is

$$\frac{dP_{\alpha\alpha}}{dt} = -(W_S + W_I + W_2)P_{\alpha\alpha} + W_S P_{\alpha\beta} + W_I P_{\beta\alpha} + W_2 P_{\beta\beta} + K \quad (1)$$

where K is a constant that ensures that the population $P_{\alpha\alpha}$ returns to its equilibrium value. At equilibrium, all derivatives with respect to time are zero and therefore the value of K can be calculated as

$$K = (W_S + W_I + W_2)P^0_{\alpha\alpha} - W_S P^0_{\alpha\beta} - W_I P^0_{\beta\alpha} - W_2 P^0_{\beta\beta} \quad (2)$$

The differential equation can be modified to instead concern deviations from the equilibrium population by

$$\frac{d\Delta P_{\alpha\alpha}}{dt} = -(W_S + W_I + W_2)\Delta P_{\alpha\alpha} + W_S \Delta P_{\alpha\beta} + W_I P_{\beta\alpha} + W_2 \Delta P_{\beta\beta} \quad (3a)$$

For the other states one similarly gets

$$\frac{d\Delta P_{\alpha\beta}}{dt} = -(W_S + W_I + W_0)\Delta P_{\alpha\beta} + W_S \Delta P_{\alpha\alpha} + W_I P_{\beta\beta} + W_0 \Delta P_{\beta\alpha} \quad (3b)$$

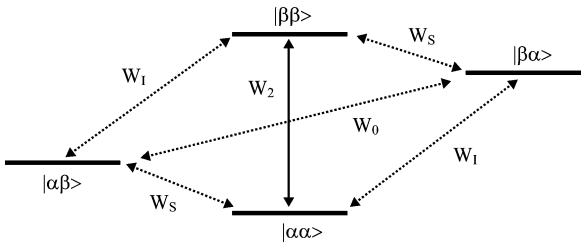
$$\frac{d\Delta P_{\beta\alpha}}{dt} = -(W_S + W_I + W_0)\Delta P_{\beta\alpha} + W_S \Delta P_{\beta\beta} + W_I P_{\alpha\alpha} + W_0 \Delta P_{\alpha\beta} \quad (3c)$$

$$\frac{d\Delta P_{\beta\beta}}{dt} = -(W_S + W_I + W_2)\Delta P_{\beta\beta} + W_S \Delta P_{\beta\alpha} + W_I P_{\alpha\beta} + W_2 \Delta P_{\alpha\alpha} \quad (3d)$$

Since the expectation values for the operators corresponding to longitudinal magnetization, $I_z(t)$ and $S_z(t)$, are equal to $I_z(t) = P_{\alpha\alpha} + P_{\alpha\beta} - P_{\beta\alpha} - P_{\beta\beta}$ and $S_z(t) = P_{\alpha\alpha} - P_{\alpha\beta} + P_{\beta\alpha} - P_{\beta\beta}$, the equations can be reformulated as

$$\frac{d\Delta I_z(t)}{dt} = -(W_0 + 2W_I + W_2)\Delta I_z(z) - (W_2 - W_0)\Delta S_z(t) \quad (4a)$$

$$\frac{d\Delta S_z(t)}{dt} = -(W_0 + 2W_S + W_2)\Delta S_z(z) - (W_2 - W_0)\Delta I_z(t) \quad (4b)$$



Nuclear Overhauser Effect, Fig. 1 Energy level diagram for the two-spin system *IS* with all possible transition rates indicated

Usually the rate constants are written as

$$\rho_I = W_0 + 2W_I + W_2 \tag{5a}$$

$$\rho_S = W_0 + 2W_S + W_2 \tag{5b}$$

$$\sigma_{IS} = W_2 - W_0 \tag{5c}$$

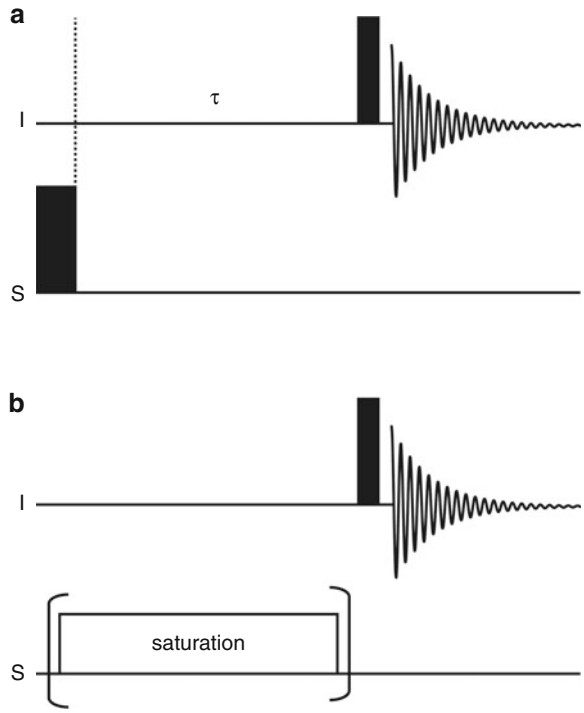
where ρ_I and ρ_S are the longitudinal auto-relaxation rate constants and σ_{IS} is referred to as the longitudinal cross-relaxation rate constant. The auto-relaxation rates, ρ_I and ρ_S, may or may not be equal and the cross-relaxation rate constant σ_{IS} is positive if $W_2 > W_0$, negative if $W_2 > W_0$, and vanishes if they are equal (cf.). With these definitions, the expressions governing the time evolution of ΔI_z and ΔS_z become

$$\frac{d\Delta I_z(t)}{dt} = -\rho_I \Delta I_z(t) - \sigma_{IS}\Delta S_z(t) \tag{6a}$$

$$\frac{d\Delta S_z(t)}{dt} = -\rho_S \Delta S_z(t) - \sigma_{IS}\Delta I_z(t) \tag{6b}$$

These equations are called the *Solomon equations* (Solomon 1955) and can be used to describe longitudinal relaxation of dipole-dipole coupled spins.

In the presence of cross-relaxation, the return of I_z to thermal equilibrium after a perturbation is typically not monoexponential with rate constant ρ_I but will depend on the spin state of S as exemplified with the following experiments (Cavanagh et al. 2007). For the *selective inversion recovery* experiment, where a 180° pulse is applied to I but not S, followed by a relaxation delay, t, and a read pulse one gets



Nuclear Overhauser Effect, Fig. 2 Pulse sequences for (**a**) the transient NOE experiment and (**b**) the steady state NOE difference experiment. *Narrow (wide) filled rectangles* denote 90° (180°) pulses while the *open rectangle* in (**b**) represents saturation by closely spaced pulses or continuous wave irradiation. The experiment is repeated without this element

$$\frac{I_z(t)}{I_z^0} = 1 - \exp[-(\rho_I - \sigma_{IS})t] \times \{1 + \exp(-2\sigma_{IS}t)\} \tag{7}$$

whereas for the *non-selective inversion recovery* experiment, where 180° pulses are applied to I as well as to S one gets

$$\frac{I_z(t)}{I_z^0} = 1 - 2\exp[-(\rho_I + \sigma_{IS})t] \tag{8}$$

Although relaxation is monoexponential in this case, this experiment cannot be used by itself to determine ρ_I. In the *transient NOE* experiment (Fig. 2a) S is selectively inverted by a 180° pulse followed by delay and a read pulse. I_z as a function of this delay is then governed by

$$\frac{I_z(t)}{I_z^0} = 1 - \exp[-(\rho_I - \sigma_{IS})t] \times \{1 - \exp(-2\sigma_{IS}t)\} \quad (9)$$

This experiment can be used to measure the cross-relaxation rate since the initial rate of change is equal to $2\sigma_{IS}$. The steady state NOE enhancement can be measured in the *steady state NOE difference* experiment (Fig. 2b). In this experiment, the S spin is selectively irradiated for a period that should be long compared with ρ_I and σ_{IS} so that $S_z(t) = 0$, after which a read pulse is applied to I. The signal intensity will then be proportional to the steady state magnetization, I_z^{ss}. The experiment is then repeated without saturation of the S spin so that the signal intensity is proportional to equilibrium magnetization, I_z^0. The steady state NOE enhancement, η_{IS}, is then calculated as

$$\eta_{IS} = \frac{I_z^{ss} - I_z^0}{I_z^0} \quad (10)$$

From semiclassical relaxation theory (Redfield 1957; Wangsness and Bloch 1953) it is possible to calculate relaxation rate constants as functions of the dipole-dipole coupling strength and spectral density functions sampled at appropriate frequencies as

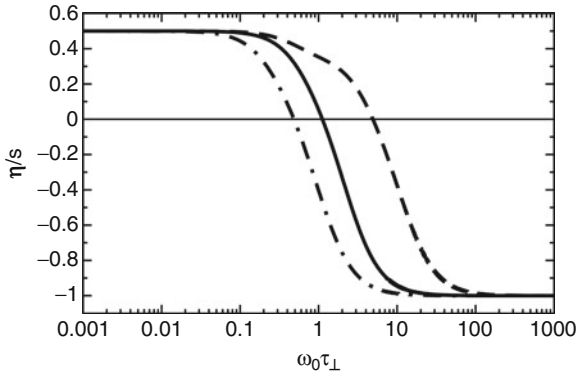
$$\rho_I = \frac{d^2}{4}(3J(\omega_I) + 6J(\omega_I + \omega_S) + J(\omega_I - \omega_S)) \quad (11a)$$

$$\rho_S = \frac{d^2}{4}(3J(\omega_S) + 6J(\omega_I + \omega_S) + J(\omega_I - \omega_S)) \quad (11b)$$

$$\sigma_{IS} = \frac{d^2}{4}(6J(\omega_I + \omega_S) - J(\omega_I - \omega_S)) \quad (11c)$$

$$\eta_{IS} = \frac{\gamma_S}{\gamma_I}\frac{\sigma_{IS}}{\rho_I} \quad (11d)$$

where $d = \mu_0 \hbar \gamma_I \gamma_S < r_{IS}^{-3} > /4\pi$, μ_0 is the permeability of free space, $\hbar$ is the Planck constant divided by 2π, γ_I and γ_S are the magnetogyric ratios for spins I and S, r_{IS} is the internuclear distance, and $J(\omega)$ is the spectral density sampled at frequency ω. For a rigid spherical top the spectral density function is modeled as



Nuclear Overhauser Effect, Fig. 3 Homonuclear NOE enhancement for a symmetric top as a function of $\omega_0\tau_\perp$, where $\tau_\perp$ is defined as the correlation time for rotation about an axis orthogonal to the symmetry axis and $\tau_\parallel$ is the correlation time for rotation about the symmetry axis. For a spherical top $\tau_\perp = \tau_\parallel = \tau_c$. The calculations have been performed for $\frac{\tau_\parallel}{\tau_\perp} = 0.1$ (*dashed line*), $\frac{\tau_\parallel}{\tau_\perp} = 1$ (*solid line*) and $\frac{\tau_\parallel}{\tau_\perp} = 10$ (*dashed-dotted line*) at a static magnetic field of 11.7 T and an angle between the internuclear vector and the symmetry axis of 90°

$$J(\omega) = \frac{2}{5}\left(\frac{\tau_c}{1 + \omega^2\tau_c^2}\right) \quad (12)$$

where τ_c is the correlation time for molecular tumbling. For the homonuclear case the equations can be further simplified to

$$\rho_I = \rho_S = \frac{d^2}{4}(J(0) + 3J(\omega_0) + 6J(2\omega_0)) \quad (13a)$$

$$\sigma_{IS} = \frac{d^2}{4}(6J(2\omega_0) - J(0)) \quad (13b)$$

where $\omega_0 = \gamma_I B_0$ and B_0 is the static magnetic field. It can be shown that if the spins I and S are of the same species, enhancement is equal to 0.5 in the *extreme narrowing* limit, $|\omega_0\tau_c| \ll 1$, whereas in the *spin diffusion* limit, $|\omega_0\tau_c| \gg 1$, enhancement is equal to −1. For a spherical top, enhancement disappears when $\omega_0\tau_c = 1.12$ while the zero-crossing occurs at different values if molecular tumbling is anisotropic as shown in Fig. 3. Similar effects are found for the NOE buildup rate (Neuhaus and Williamson 2000).

A noteworthy feature of the above expressions is that although σ_{IS} is proportional to the inverse sixth power of the distance between spins I and S, (Eq. 11d) predicts that NOE enhancement is independent of

internuclear distance. This is however only true in the case of an isolated two-spin system where the sole relaxation mechanism is the dipole-dipole interaction between the two spins. In practice, the expression for η_{IS} must be modified to

$$\eta_{IS} = \frac{\gamma_S}{\gamma_I} \frac{\sigma_{IS}}{\rho_I + \rho_I^*} \tag{14}$$

where ρ_I^* is longitudinal relaxation of I due to other mechanisms than the dipolar coupling to S. These may be contributions from chemical shift anisotropy, dipole-dipole interactions with external nuclear spins and with unpaired electrons from dissolved oxygen or paramagnetic ions. Since ρ_I^* is independent of the internuclear distance between I and S, it follows that the enhancement decreases as the internuclear distance increases. There is thus no straightforward answer to which is the maximum allowed distance between I and S for the observation of the NOE. However, the concept is seldom meaningful for internuclear distances of more than 5 Å.

The Heteronuclear NOE

The above equations hold equally well if the nuclei I and S are of different species but the fact that the magnetogyric ratios γ_I and γ_S are different must be taken into account. Specifically, this means that higher enhancement will be realized for the nucleus with the lower magnetogyric ratio. The NOE can thus be used to increase the sensitivity in experiments probing low magnetogyric ratio nuclei although coherent polarization transfer methods such as INEPT usually are more efficient (▶ INEPT). It is also noteworthy that if the relative signs of γ_I and γ_S differ, enhancement will be negative in the extreme narrowing limit. A subtlety in this case is that the precession frequencies also are of opposite sign, implying that $|\omega_I \pm \omega_S| \neq |\omega_I| \pm |\omega_S|$. This must be taken into account when evaluating the spectral density functions and indicates that NOE enhancement as a function of $\omega_0\tau_c$ for $\frac{\gamma_i}{\gamma_j} = -K$ is not simply the mirror image of the case when $\frac{\gamma_i}{\gamma_j} = -K$ (Neuhaus and Williamson 2000).

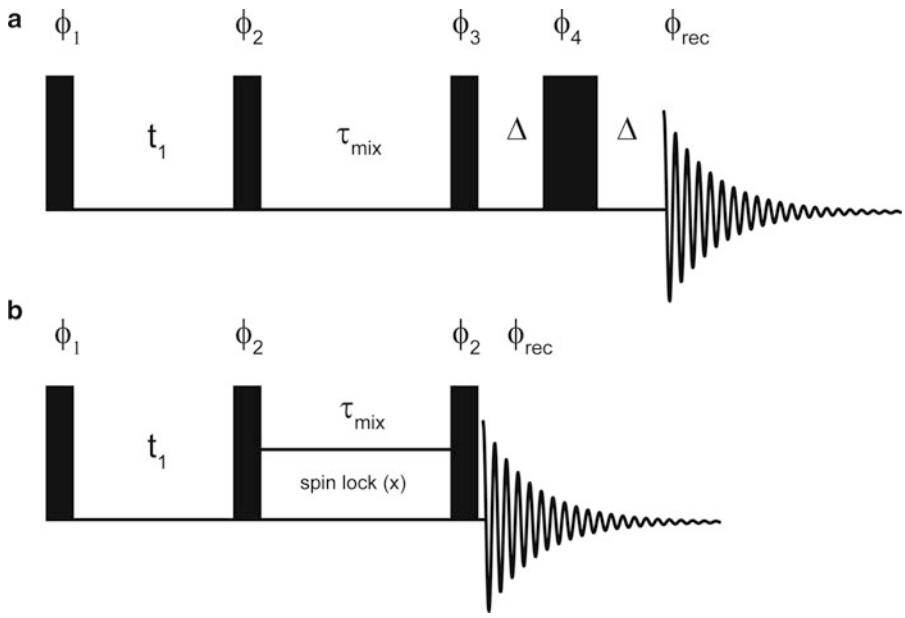
For a heteronuclear spin-pair in a macromolecule the frequencies $\omega_I \pm \omega_S$ are typically significantly larger than the inverse of the correlation time for molecular tumbling. For example, for ^{1}H and ^{15}N at a field of 11.7 T, $|\omega_H + \omega_N|$ and $|\omega_H - \omega_N|$ correspond to $2.8 \cdot 10^9$ and $3.4 \cdot 10^9$ s^{-1}, respectively. The steady state heteronuclear NOE enhancement is thus sensitive to motions that are faster than τ_c and is commonly used to probe such motions. Typically, it is measured by recording a two-dimensional version of the experiment in Fig. 2b with saturation of the S spin performed by application of closely spaced 120° pulses for a period of several seconds. In order to measure the steady state NOE accurately, it is imperative that relaxation to thermal equilibrium and steady state, respectively, is complete. Recycle delays in excess of 10 s which should be compared to one or a few seconds for most protein NMR experiments should thus be employed. This experiment is combined with measurements of the longitudinal and transverse auto-relaxation rates and the data is usually interpreted according to the model-free formalism, in which the spectral density function is modeled with parameters describing molecular tumbling as well as correlation times and order parameters of internal motions on time scales that are faster than molecular tumbling (Clore et al. 1990; Lipari and Szabo 1982).

The Solomon Equations for a Multi-spin System

It is straightforward to generalize the energy level diagram of Fig. 1 and the Solomon equations to more than two spins. It is noteworthy that the dipole-dipole interaction between two spins can never induce transitions for a third spin, that is, only pairwise W_0, W_1, and W_2 processes have to be considered. For a multi-spin system, the equation describing ΔI_z is modified to

$$\frac{d\Delta I_z(t)}{dt} = -\rho_I \Delta I_z(t) - \sigma_{IS}\Delta S_z(t) - \sum_X \sigma_{IX}\Delta X_z(t) \tag{15}$$

where the sum runs over all external spins, X and the definition of ρ_I is modified to also include couplings to these spins. In the initial time regime, NOE buildup is no different than for the two-spin case. However, at later time-points indirect mechanisms must be considered (Cavanagh et al. 2007). For instance the mechanism dubbed *spin diffusion* (▶ Spin Diffusion) is the result of cross-relaxation being relayed by a third spin, that is, $S \rightarrow X \rightarrow I$ and may erroneously be interpreted



Nuclear Overhauser Effect, Fig. 4 Basic pulse sequences for NOESY and ROESY experiments. *Narrow (wide) rectangles* correspond to pulses with flip angles of 90° (180°). (**a**) In this NOESY sequence, which includes a Hahn echo, the basic eight steps of the phase cycle is ϕ_1 = x,−x,−x,x, ϕ_2 = x,x,−x−x, ϕ_3 = x, ϕ_4 = y,y,y,y,−y,−y,−y,−y, ϕ_{rec} = x,−x. (**b**) For the ROESY experiment the basic phase cycle is ϕ_1 = x,−x, ϕ_2 = x,x, ϕ_{rec} = x,−x. For both experiments the phase cycle is extended by CYCLOPS on all pulses and the receiver. Frequency discrimination in F1 can be achieved by TPPI or States-TPPI. Bipolar gradients can be employed during t_1 to prevent radiation damping and for the NOESY experiment an additional gradient may be applied at the beginning of τ_{mix}

as an NOE between *I* and *S*. The effect is however useful for rapidly transferring saturation throughout a molecule and to a ligand as a way of detecting transient interactions in cross saturation experiments (Takahashi et al. 2000). Another indirect effect is called *back transfer* and is the result of magnetization being relayed back to the original spin, that is, $I \rightarrow S \rightarrow I$. This process makes the NOE weaker than for an isolated *IS* two-spin system.

The NOESY Experiment

The nuclear Overhauser effect spectroscopy (NOESY) experiment is a two-dimensional experiment used to measure cross-relaxation rates (Macura et al. 1981). The basic pulse sequence is shown in Fig. 4a. After an excitation pulse, the spins are frequency labeled during t_1 and subsequently returned to the z-axis. Magnetization transfer, $S \rightarrow I$, due to cross-relaxation occurs for period τ_{mix} after which a read pulse followed by an optional Hahn echo is applied.

Since buildup of cross-peak intensities is linear with proportionality constant $2\sigma_{IS} \propto r_{IS}^{-6}$ in the initial time regime, cross-peak intensities are often used as a measure of internuclear distance (▶ Interspin Distance Determination by EPR). This must however be done with caution. For instance, cross-peaks that only appear at long mixing times may be due to spin diffusion and should not be considered. Furthermore, the identical dependence on r_{ij}^{-6} for all pairs of spins only holds under the assumption that the internal motions of all spins are equal. Usually, cross-peak intensities are converted into distances conservatively by dividing them into three classes corresponding to strong, medium and weak based on their intensities and assigning them distance ranges of 1.8–2.7, 1.8–3.3 Å, and 1.8–5.0 Å, respectively. By always using the shortest possible distance as the lower bound, internal motions that reduce cross-peak intensities are taken into account. The r_{IS}^{-6} dependence on cross-peak intensity must also be considered when calculating distances of single averaged signals. Averaging is in fact slightly involved and a detailed analysis shows that conformational changes

on a time scales slower than molecular tumbling, such as flips of aromatic rings, should be averaged as $\langle r_{IS}^{-6}\rangle^{-\frac{1}{6}}$ whereas motions occurring on faster time scales, such as rotations of methyl groups, should be averaged as $\langle r_{IS}^{-3}\rangle^{-\frac{1}{3}}$ (Neuhaus and Williamson 2000).

Since the homonuclear NOE vanishes for $\omega_0\tau_c = 1.12$ the NOESY experiment can in practice be realized for small or large molecules. Especially in the spin diffusion limit, NOESY is the method of choice for measurement of cross-relaxation rates since (1) two or more dimensions are necessary to avoid signal overlap, (2) it is not feasible to perform selective inversion of all nuclei for large molecules, (3) the time needed to complete the phase cycle is not a limiting factor because of low sensitivity, and (4) buildup of NOE enhancements is rapid in the spin diffusion limit. Whether the one-dimensional transient NOE experiment or NOESY is best suited for small molecules is often not so obvious and depends on how many resonances must be selectively inverted among other things (Neuhaus and Williamson 2000).

The ROESY Experiment

In the rotating frame Overhauser effect spectroscopy (ROESY) experiment (Bothner-By et al. 1984) (Fig. 4b) the spins are subject to continuous wave irradiation during τ_{mix}. The amplitude of this so-called spin-lock field, ω_1, is usually on the order of 10 kHz. The spins will then experience an effective field $\omega_{eff} = (\omega_1^2 + \Omega^2)^{0.5}$ that makes the angle $\theta = \arctan\left(\frac{\omega}{\Omega}\right)$ with the static magnetic field. If the spin-lock field is applied on resonance, $\Omega = 0$, the dipolar contributions to auto- and cross-relaxation for a homonuclear IS two-spin system are given by (Neuhaus and Williamson 2000)

$$\begin{aligned}\rho_I^{ROE} &= \frac{d^2}{8}\left\{\frac{9}{2}J(2\omega_1)+\frac{1}{2}J(0)+9J(\omega_0)+6J(2\omega_0)\right\}\\ &\approx \frac{d^2}{8}\{5J(0)+9J(\omega_0)+6J(2\omega_0)\}\end{aligned} \tag{16a}$$

$$\begin{aligned}\sigma_{IS}^{ROE} &= \frac{d^2}{8}\left\{\frac{9}{2}J(2\omega_1)-\frac{1}{2}J(0)+6J(\omega_0)\right\}\\ &\approx \frac{d^2}{8}\{4J(0)+6J(\omega_0)\}\end{aligned} \tag{16b}$$

where the superscript ROE is used to distinguish the rotating frame rate constants from their laboratory frame counterparts. The equations make it clear that ROESY experiment is an excellent choice for molecules of intermediate size such that $\omega_0\tau_c = 1.12$ since the cross-relaxation constant does not vanish. Because of fast auto-relaxation, it is less well suited for very large molecules. Other complications are due to side effects of spin-locking. For instance, if it is not applied on resonance, the magnetization will be locked in a tilted frame and the auto- and cross-relaxation rates will be averages of ρ_I^{ROE} and ρ_I and σ_{IS}^{ROE} and σ_{IS} weighted by the tilt angle the effective field makes with the static field. Another complication arises since the strong J coupling (▶ J Coupling) is active in the presence of a spin-lock field and may lead to magnetization transfer not due to cross-relaxation. However, continuous wave irradiation is not particularly effective for promoting transfer of magnetization by this mechanism except for spins resonating with very similar frequencies making this problem less important.

Summary

The nuclear Overhauser effect is one of the most useful features of NMR spectroscopy and is used to transfer nuclear spin polarization from one population to another. The NOE is sensitive to molecular motions and is often used to characterize these. Since it is a through-space rather than a through-bond effect, it reports on proximity of nuclei and not on how many covalent bonds that separate them. One of the most important applications is therefore structure calculations of proteins and other molecules (▶ NMR-based Structural Proteomics).

Cross-References

- ▶ INEPT
- ▶ J Coupling
- ▶ NMR in Drug Discovery – Introduction
- ▶ NMR Studies of Macromolecular Interactions – Introduction
- ▶ NMR-based Structural Proteomics
- ▶ Nucleic Acid NMR – Introduction
- ▶ Protein NMR – Introduction
- ▶ Spin Diffusion

References

Bloembergen N, Purcell EM, Pound RV. Relaxation effects in nuclear magnetic resonance absorption. Phys Rev. 1948;73:679–715.

Bothner-By AA, Stephens RL, Lee JM, Warren CD, Jeanloz RW. Structure determination of a tetrasaccharide – transient nuclear overhauser effects in the rotating frame. J Am Chem Soc. 1984;106:811–3.

Carver TR, Slichter CP. Polarization of nuclear spins in metals. Phys Rev. 1953;92:212–3.

Cavanagh J, Fairbrother WJ, Palmer III AG, Rance M, Skelton NJ. Protein NMR spectroscopy. Principles and practice. Burlington: Elsevier Academic; 2007.

Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM. Deviations from the simple two-parameter model-free approach to the interpretation of ^{15}N nuclear magnetic relaxation of proteins. J Am Chem Soc. 1990;112:4989–91.

Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules.1. Theory and range of validity. J Am Chem Soc. 1982;104:4546–59.

Macura S, Huang Y, Suter D, Ernst RR. Two-dimensional chemical-exchange and cross-relaxation spectroscopy of coupled nuclear spins. J Magn Reson. 1981;43:259–81.

Neuhaus D, Williamson MP. The nuclear overhauser effect in structural and conformational analysis. New York: Wiley; 2000.

Overhauser AW. Polarization of nuclei in metals. Phys Rev. 1953;92:411–5.

Redfield AG. On the theory of relaxation processes. IBM J Res Dev. 1957;1:19–31.

Solomon I. Relaxation processes in a system of two spins. Phys Rev. 1955;99:559–66.

Takahashi H, Nakanishi T, Kami K, Arata Y, Shimada I. A novel NMR method for determining the interfaces of large protein-protein complexes. Nat Struct Biol. 2000;7:220–3.

Wangsness RK, Bloch F. The dynamical theory of nuclear induction. Phys Rev. 1953;89:728–39.

Nucleic Acid Circular Dichroism

Alison Rodger
Department of Chemistry, University of Warwick, Coventry, UK

Synonyms

DNA

Definition

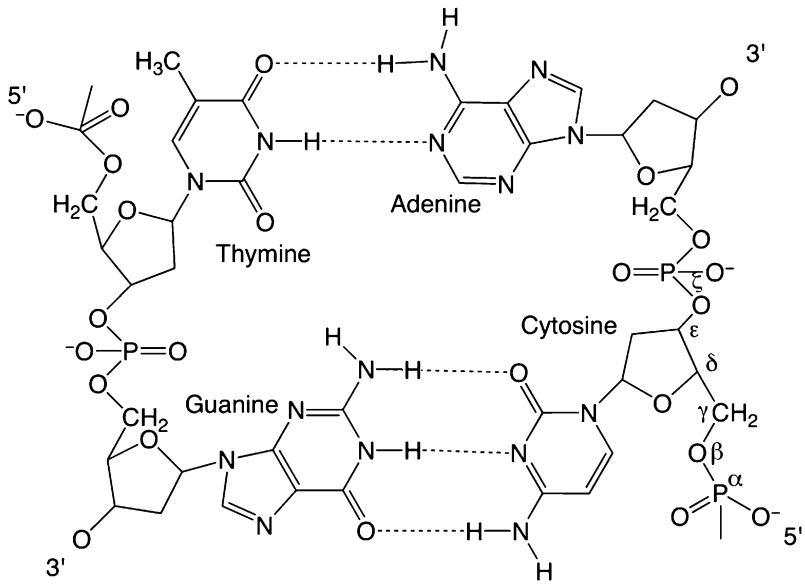
Deoxyribonucleic acid, DNA, is composed of planar aromatic nucleic acid bases linked to a sugar and a phosphate as illustrated in Fig. 1. Ribonucleic acid, RNA, differs from DNA by an extra oxygen on the sugar and also thymine is replaced by uracil which does not have the methyl substituent. Most of the DNA and RNA transitions we measure with UV spectroscopy is due to the π–π^* transitions of the bases. These transitions have no intrinsic CD signal since the bases are planar and hence achiral. However, the isolated bases acquire asymmetry in their electronic transitions by coupling with the chiral ribose sugar units of the backbone. The magnitude of $\Delta\varepsilon_{max}$ (the molar difference in absorbance of left and right circularly polarized light) for each base is of the order of 2 mol^{-1} dm^3 cm^{-1} at 270 nm; the purine bases have a negative signal whereas the pyrimidine ones have a positive CD in that region. When the bases are linked by phosphodiester bonds to form DNA or RNA, an additional (and usually larger) source of chirality results from the coupling between bases that are stacked in a helix (Berova et al in press; Nordén et al 2010).

Basic Characteristics

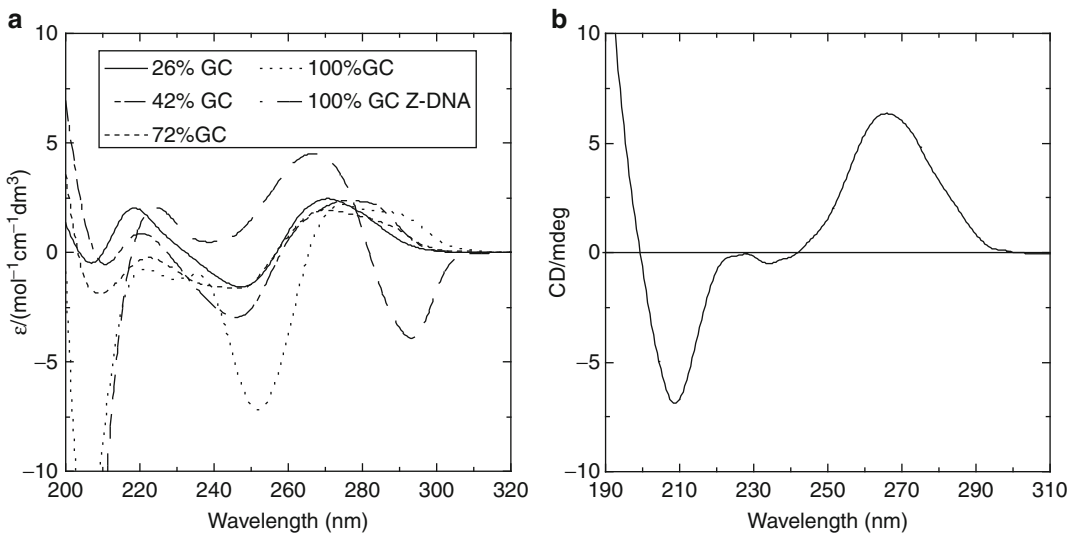
The shape of the CD spectrum measured for a DNA or RNA sample contains information not only about its base pair content but also about the geometry of the structure. In fact, the orientation of the bases relative to one another (i.e., the DNA polymorph) is the most significant factor determining the spectral shape. Since CD spectra are easy to measure, CD is often the simplest technique to use to probe DNA conformational changes as a function of ionic strength, solvent, ligand concentration, etc. in solution.

The simplest application of CD to DNA (or RNA) structure determination is to identify which polymorph is present in a sample. Since the DNA CD from 190 to 300 nm is due to the skewed orientation of the bases, if the DNA is untwisted or the bases are tilted, a change would be expected in the CD spectrum. Some examples are given in Fig. 2. Z-DNA is negative to a first approximation where the CD of B-DNA is positive. The CD of a typical RNA or A-form DNA has a sharp negative band at 210 nm.

A CD spectrometer is an ideal tool for probing condensed DNA structures, although what is measured then is not strictly CD, but rather a form of light scattering. The results of intramolecular condensation



Nucleic Acid Circular Dichroism, Fig. 1 DNA nucleotide bases indicating backbone parameter labels used to describe the orientation of DNA bases with respect to one another

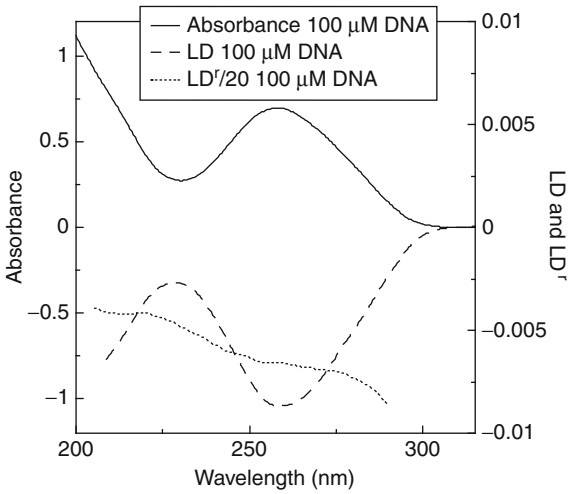


Nucleic Acid Circular Dichroism, Fig. 2 (**a**) CD spectrum of DNAs as a function of GC content (pH = 6.8, 5 mM NaCl). *Clostridium perfringens* (27% GC), calf thymus DNA (42% GC), *Micrococcus luteus* (72% GC), poly[d(G–C)]$_2$ (100% GC). Also shown is Z-form poly[d(G–C)]$_2$ induced by 50 μM $[Co(NH_3)_6]^{3+}$. (**b**) *CD* of a 372 nucleotide mRNA. $A_{260\ nm} = 1$ at 25°C for this sample

are particles of size comparable with the wavelength of the light. The particles interact differently with left and right circularly polarized light and, at the wavelengths where the chromophores absorb light, give a large "CD" signal. The sign of the CD indicates the helical handedness of the condensed DNA particle, with a negative signal above 250 nm corresponding to a left-handed helix (Arscott et al 1995). CD is also useful for probing structural changes such as those from quadruplex to non-quadruplex structures. However, it should be noted that the accepted spectral signatures for such transitions need to be validated by considering, for example, NMR data from samples at the same concentrations and conditions as the structures exist as equilibria between different arrangements of the bases.

Cross-References

- Circular Dichroism Spectroscopy of Biomacromolecules
- Circular Dichroism Spectroscopy: Units
- DNA-Ligand Circular Dichroism
- DNA-Ligand Flow Linear Dichroism
- Far UV Protein Circular Dichroism
- Linear Dichroism
- Linear Dichroism Spectra - Measurement
- Linear Dichroism Spectroscopy: Theory
- Near UV Protein CD
- Nucleic Acid Linear Dichroism
- Oriented Circular Dichroism Spectroscopy
- Polarized Light, Linear Dichroism, and Circular Dichroism
- Protein Circular Dichroism Analysis



Nucleic Acid Linear Dichroism, Fig. 1 Absorbance, LD and reduced LD, $LD^r = LD/A$ spectra of calf thymus DNA (100 μM) in a 1 cm path length cuvette (absorbance) and a 1 mm path length Couette flow cell (LD)

References

Arscott PG, Ma C, Wenner JR, Bloomfield VA. DNA condensation by cobalt hexaammine(III) in alcohol-water mixtures: dielectric constant and other solvent effects. Biopolymers. 1995;36:345–65.

Berova N, Woody RW, Polavarapu P, Nakanishi K. Comprehensive chiroptical spectroscopy (in press).

Nordén B, Rodger A, Dafforn TR. Linear dichroism and circular dichroism: a textbook on polarized spectroscopy. Cambridge: Royal Society of Chemistry; 2010.

Nucleic Acid Linear Dichroism

Alison Rodger
Department of Chemistry, University of Warwick, Coventry, UK

Synonyms

Anisotropic spectroscopy; DNA

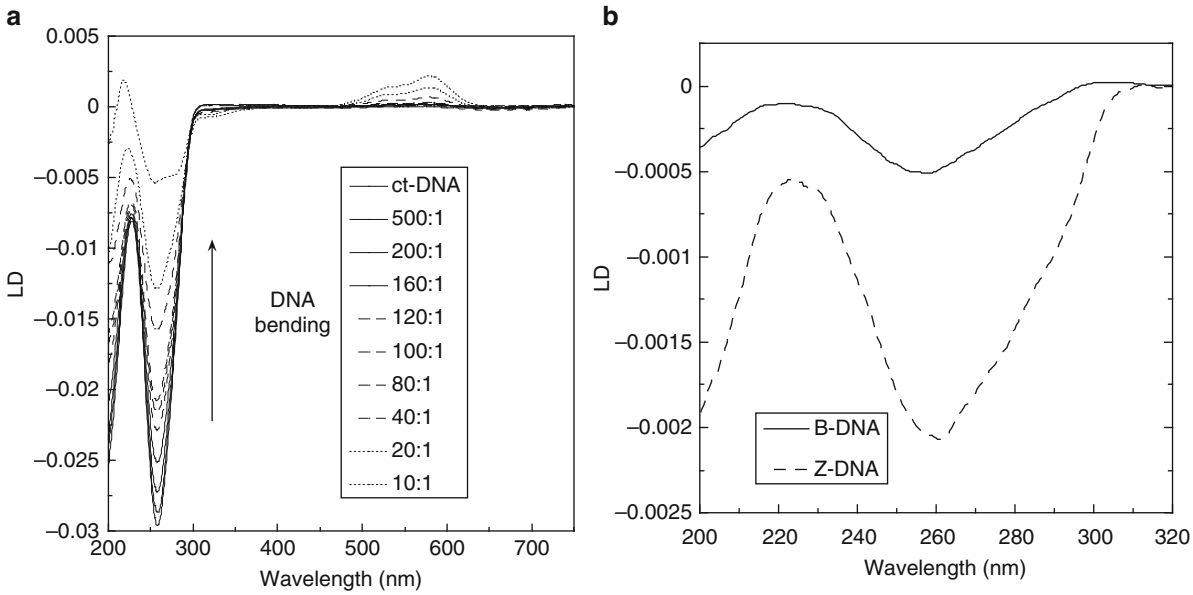
Definition

A linear dichroism (LD) spectrum shows the difference in absorbance of light polarized parallel and perpendicular to an orientation direction (Rodger and Nordén 1997; Nordén et al 2010). The idealized structure of B-DNA has the base pairs oriented perpendicular to the DNA helix axis to a first approximation. This means the π–π* transitions take place in the plane perpendicular to the helix axis, and so these are expected to give a negative LD signal as illustrated in Fig. 1. LD can be used to probe changes in orientation of the DNA bases.

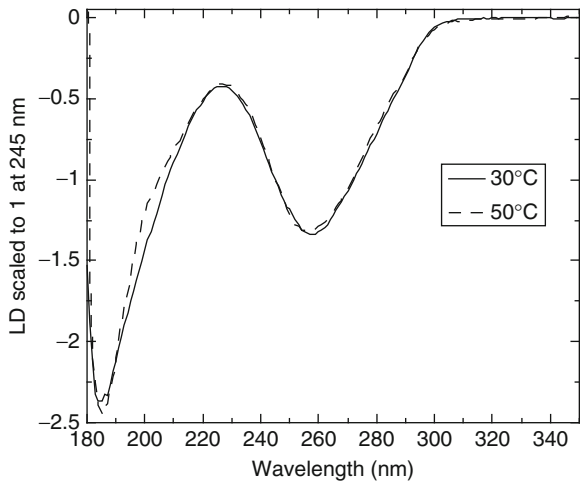
The LD spectrum of B-DNA shown in Fig. 1 is similar to but not exactly the same shape as the negative of its absorbance spectrum, $-A$. The same information is summarized in the reduced LD spectrum which is fairly constant across the 260 nm band, but not completely so. This tells us that, in solution, the bases are not rigidly perpendicular to the helix axis as implied by the idealized B-DNA structure. Careful analysis of solution LD studies has led to the conclusion that the bases of B-DNA in solution lie at an average angle of ~80° (or even less) from the helix axis (Chou and Johnson 1993) (though we typically assume 86° in a calculation (Matsuoka and Nordén 1982)).

The magnitude of the LD signal of any sample is also affected by the degree of orientation as summarized by the orientation parameter, S.

$$LD^r = \frac{LD}{A_{iso}} = \frac{3}{2}S\left(3\cos^2\alpha - 1\right) \quad (1)$$



Nucleic Acid Linear Dichroism, Fig. 2 (**a**) The LD of calf thymus B-DNA (500 μM, 20 mM NaCl, 1 mM sodium cacodylate buffer pH = 7) upon addition of a tetracationic di-iron triple helicate $[Fe(LL)_3]^{4+}$ (LL = $C_{25}H_{20}N_4$) at DNA:ligand ratios shown on the figure (Hannon et al. 2001) (**b**) LD spectra of poly[d(G–C)]$_2$ (50 μM, 5 mM NaCl, 1 mM cacodylate) as B-DNA and Z-DNA after the addition $[Co(NH_3)]^{3+}$ (50 μM)



Nucleic Acid Linear Dichroism, Fig. 3 LD spectra of *Micrococcus luteus* DNA (200 μM in water) as a function of temperature normalized to 1 at 245 nm. Data were collected in a 0.5-mm path length Couette flow cell spinning at ∼3,000 rpm at the Astrid synchrotron. Quartz cutoff for sample capillary is at 182 nm in these spectra (Rittman et al. 2012)

The magnitude of the orientation parameter, *S*, can be used to provide information about the length, flexibility, or bendability of DNA or about the introduction of kinks into DNA. If *S* increases as the result of some change imposed on the system (such as the addition of a DNA-binding ligand), then the LD increases because the DNA has become more oriented. This generally means it has lengthened, or stiffened (or both), or some bend or kink has been removed. Similarly, molecules that bind to DNA and bend it reduce *S*. Figure 2 illustrates both a bending effect and a stiffening effect (in this case converting to B-DNA to Z-DNA).

Subtle changes in DNA LD spectral shapes can be used to probe changes in structure such as that induced by increases or decreases in single stranded regions as illustrated by the salt and temperature dependence of the DNA spectra of Fig. 3 (Rittman et al. 2012).

Cross-References

- ▶ DNA-Ligand Circular Dichroism
- ▶ DNA-Ligand Flow Linear Dichroism
- ▶ Linear Dichroism
- ▶ Linear Dichroism Spectra - Measurement
- ▶ Linear Dichroism Spectroscopy: Theory
- ▶ Polarized Light, Linear Dichroism, and Circular Dichroism

References

Chou PJ, Johnson WC. Base inclinations in natural and synthetic DNAs. J Am Chem Soc. 1993;115:1205–14.
Hannon MJ, Moreno V, Prieto MJ, Molderheim E, Sletten E, Meistermann I, Isaac CJ, Sanders KJ, Rodger A. Intramolecular DNA coiling mediated by a metallo supramolecular cylinder. Angew Chem. 2001;40:879–84.
Matsuoka Y, Nordén B. Linear dichroism studies of nucleic acids. II. Calculation of reduced dichroism curves of A and B form DNA. Biopolymers 1982;21:2433–52.
Nordén B, Rodger A, Dafforn TR. Linear dichroism and circular dichroism: a textbook on polarized spectroscopy. Cambridge: Royal Society of Chemistry; 2010.
Rittman M, Hoffmann SV, Gilroy EL, Hicks MR, Finkenstadt B, Rodger A. Probing the structure of long DNA molecules in solution using synchrotron radiation linear dichroism. PCCP. 2012;14:353–66.
Rodger A, Nordén B. Circular dichroism and linear dichroism. Oxford: Oxford University Press; 1997.

Nucleic Acid NMR – Introduction

Masato Katahira and Tsukasa Mashima
Kyoto University, Institute of Advanced Energy, Kyoto, Japan

Synonyms

DNA

Definition

Nucleic acid NMR is the study of the DNA/RNA structure in solution with the NMR method.

Introduction

► NMR is a powerful tool to study the ► DNA/► RNA structure in solution. It should be reminded that there are many examples in which the structure of a certain DNA/RNA in solution is quite different from that in crystal. This entry describes how the DNA/RNA structure can be determined by the NMR method. In addition to the established ways of the analysis, recent progresses and tips are also documented.

Firstly, three different ways of the DNA/RNA preparation are described. Secondly, the method of the resonance assignment, that is the crucial step for the NMR analysis, is explained. Thirdly, the way to identify the mode of the base pair, that is the key information to construct the structure, is described. Fourthly, the methods to derive various structural restraints are explained. Fifthly, the way to determine the structure by calculation by means of structural restraints is described. Finally, the recent examples of the DNA/RNA structures determined by NMR are shown. The example of real-time monitoring of the cytosine to uracil conversion of DNA by a human anti-HIV enzyme is also shown.

Sample Preparation

DNA and RNA can be made with a DNA/RNA synthesizer. By using the ^{13}C, ^{15}N-labeled phosphoramidite, either uniform or residue-specific ^{13}C, ^{15}N-labeling can be achieved. The yield of the overall synthesis becomes very low, when the length of DNA/RNA is long. Thus, this method is suitable for making relatively short DNA/RNA.

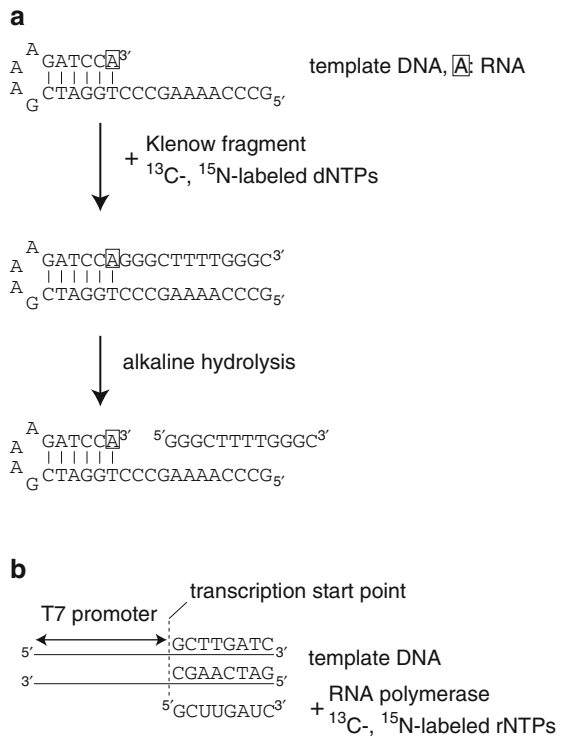
DNA can also be made with enzymatic synthesis with a DNA polymerase, a Klenow fragment, using a proper template (Fig. 1a) (Zimmer and Crothers 1995). By locating a ribonucleotide at the end of the template, the synthesized target DNA can be separated from the template by alkaline hydrolysis. The PCR method can also be applied for the synthesis (Werner et al. 2001). By using ^{13}C, ^{15}N-labeled dNTPs as starting materials, either uniform or nucleotide-type-specific ^{13}C, ^{15}N -labeling can be achieved. DNA is purified with denatured polyacrylamide gel electrophoresis.

RNA can also be made with enzymatic synthesis with an RNA polymerase of T7 phage using a proper template (Fig. 1b) (Varani et al. 1996; Furtig et al. 2003). Efficiency of the synthesis of short RNA is usually very low; thus this method is suitable for making relatively long RNA.

It is recommended to put an NMR tube in a 15% H_2O_2 solution overnight to inactivate a trace of RNase before use and to extensively wash the tube with water afterward.

Resonance Assignments

A repeating unit of DNA/RNA, a nucleotide, and five bases are shown in Fig. 2a and b, respectively.



Nucleic Acid NMR – Introduction, Fig. 1 *Sample preparation*. (**a**) Enzymatic synthesis of DNA with a DNA polymerase, a Klenow fragment. The 3′-end residue of the template, boxed, is a ribonucleotide. (**b**) Enzymatic synthesis of RNA with an RNA polymerase

The most useful tool for the assignments of DNA/RNA is the sequential base proton (either H6 of pyrimidine or H8 of purine)-H1′ connectivities comprising of intraresidue H6(*i*)/H8(*i*)-H1′(*i*) and interresidue H6(*i*)/H8(*i*)-H1′(*i* − 1) cross peaks in NOESY (Varani et al. 1996; Cromsigt et al. 2001; Furtig et al. 2003). An example of the trace of the connectivities is shown in Fig. 3a (Mashima et al. 2009). On the basis of the H1′ assignment, the other sugar protons, H2′, H2″, H3′, H4′, H5′, and H5″, can be assigned using DQF-COSY and TOCSY. The obtained assignments can be confirmed by the sequential base proton-H2′/H2″ and base proton-H3′ connectivities.

When a sugar takes on C3′-*endo* puckering as often found in RNA, $^{3}J_{H1'H2'}$ is almost 0 Hz, thus the H1′–H2′ correlation is hard to observe in either DQF-COSY or ▶ TOCSY. Even in this case, HCCH-COSY and HCCH-TOCSY can give the H1′–H2′ correlation and the H2′ assignment as a result. These spectra are also powerful for the assignments of the other sugar protons.

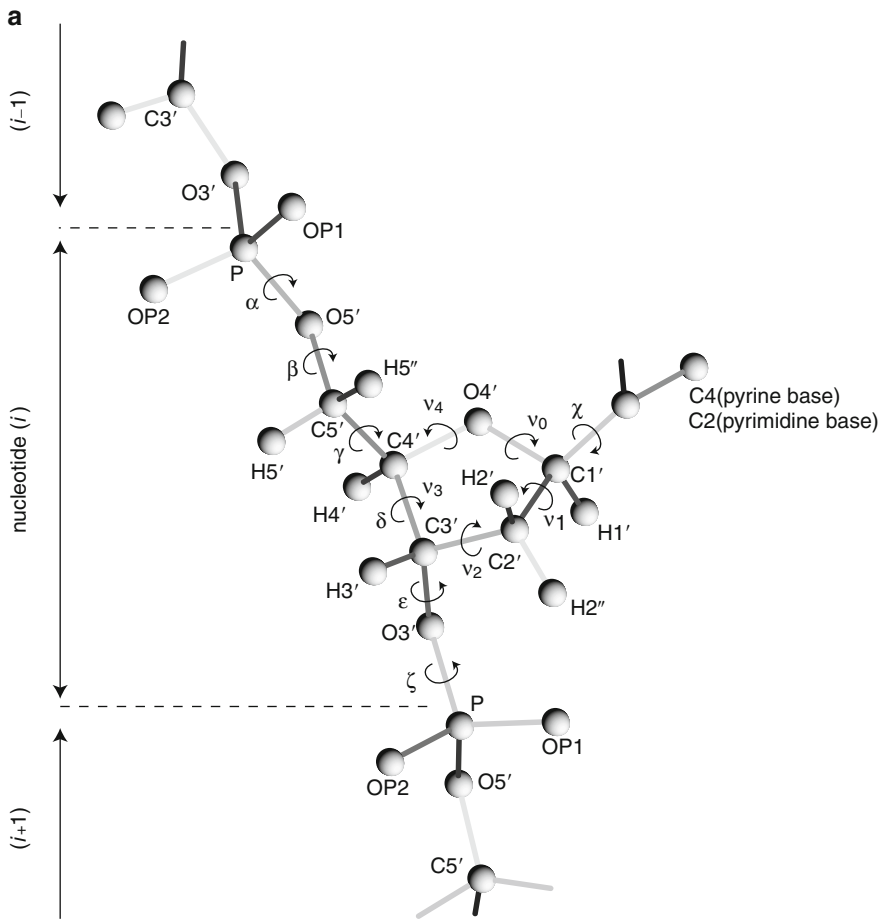
The sequential phosphorus H3′/H4′/H5′/H5″ connectivities comprising of intraresidue P(*i*)-H5′(*i*)/H5″(*i*)/H4′(*i*) and interresidue P(*i*)-H3′(*i* − 1) cross peaks in Het-COR can further confirm the sugar proton assignments, with the assignments of phosphorus being obtained at the same time. The sequential phosphorus C3′/C4′/C5′–H3′/H4′/H5′/H5″ connectivities observed in HCP can also confirm the assignments. For relatively short DNA/RNA, these two methods using the phosphorus can give the assignments on the basis of totally through-bond connectivities without relying on through-space connectivities obtained from NOESY. For longer DNA/RNA, the methods using the phosphorus become less sensitive because they rely on relatively small 2-, 3-, or 4-bond coupling constants. Intraresidue correlation between H1′ and H6/H8 can also be obtained purely relying on through-bond connectivities for relatively small DNA/RNA.

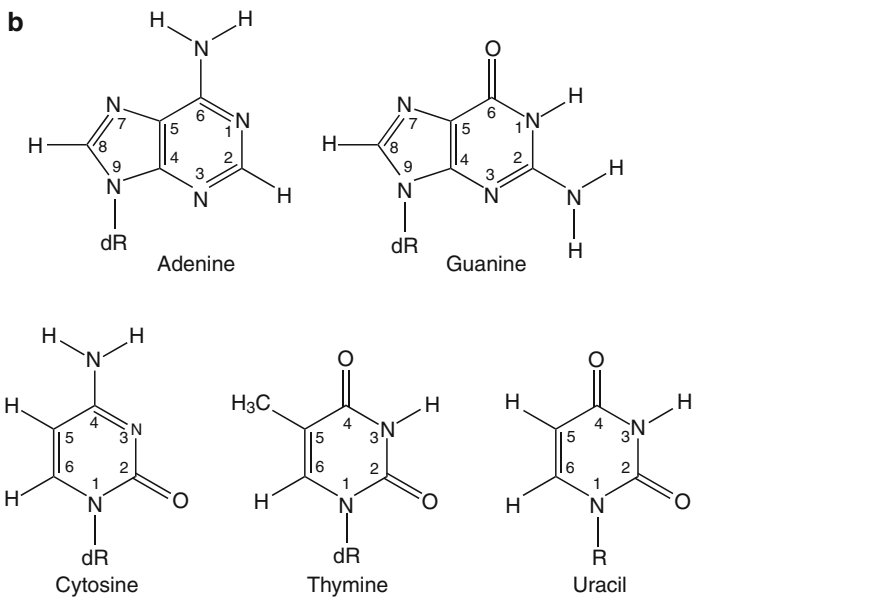
H5 of pyrimidine is assigned from an H5–H6 correlation peak in TOCSY, and H2 of adenine is from an H2–H8 correlation peak in HCCH-TOCSY. HMBC is also useful to assign H2 for non-labeled DNA/RNA through H8–C4 and H2–C4 correlation peaks.

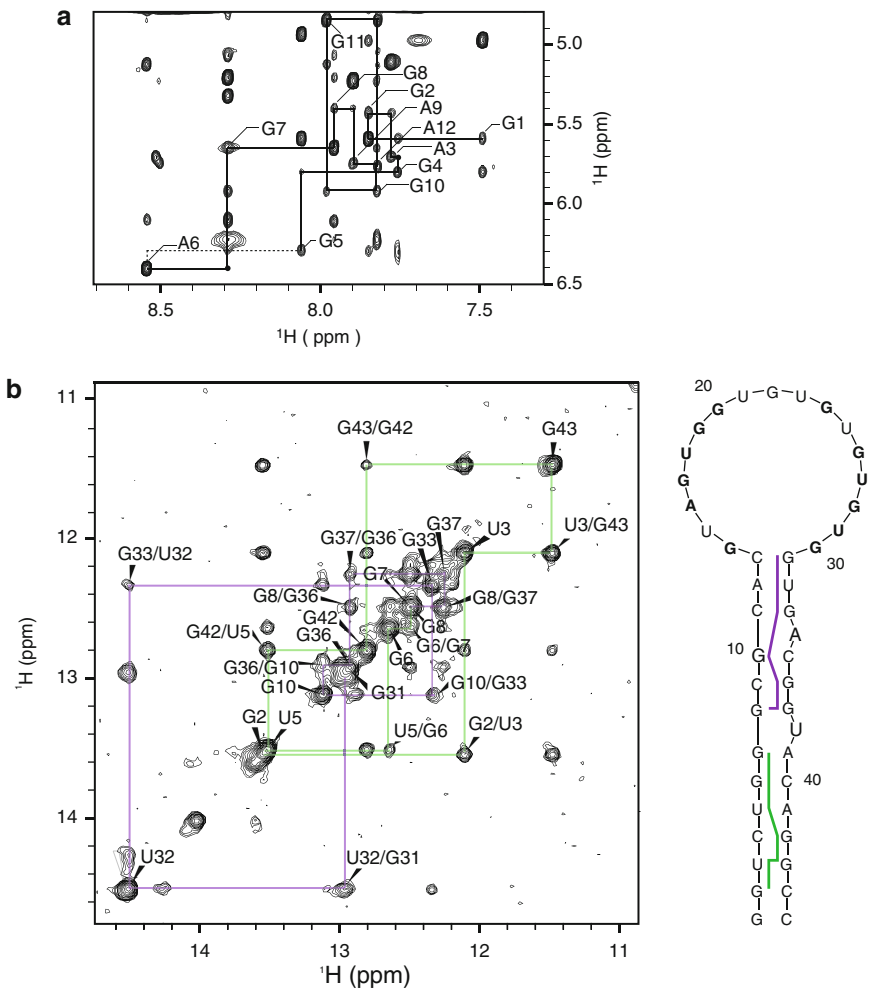
^{15}N ▶ HSQC is useful to discriminate exchangeable imino protons originating from guanine and thymine (uracil) bases. The imino nitrogen resonance of a guanine appears at 146–149 ppm, while that of a thymine (uracil) does at 156–162 ppm. The residue-specific assignment of the imino proton can be obtained from HCCNH-TOCSY on the basis of already assigned H6/H8 through the magnetization transfer of either H8 → C8 → C4 → C6 → N1 → H1 (imino proton) for a guanine or H6 → C6 → N1 → C2 → N3 → H3 (imino proton) for a thymine (uracil). The imino proton of a guanine can also be assigned with HNN-COSY in which both H8 and H1 (imino) give the correlation peak to the common N3. The imino proton of a thymine (uracil) can also be assigned with ^{15}N HSQC in which both H5 and H3 (imino) give the correlation peak to the common N3. The imino proton of a guanine of non-labeled DNA/RNA can be assigned with JR-HMBC in which both H8 and H1 give the correlation peak to the common C5. When a certain secondary structure can be assumed, imino protons can be assigned on the basis of imino-imino connectivities in NOESY between neighboring base pairs (Fig. 3b) (Nakano et al. 2011).

Discrimination of exchangeable amino protons originating from cytosine, adenine, and guanine bases

Nucleic Acid NMR – Introduction, Fig. 2 *A nucleotide and bases*. (**a**) Structure of a nucleotide, a repeating unit of DNA/RNA. (**b**) Structures of five bases







Nucleic Acid NMR – Introduction, Fig. 3 *Assignments.* (**a**) Sequential base proton (either H6 of pyrimidine or H8 of purine)-H1′ connectivities in NOESY for r(GGAGGAGGAGGA), the intraresidue base-H1′ cross peak being labeled with the residue number (Mashima et al. 2009). (**b**) Imino-imino connectivities between neighboring base pairs for the stem region of an RNP aptamer against ATP (Nakano et al. 2011)

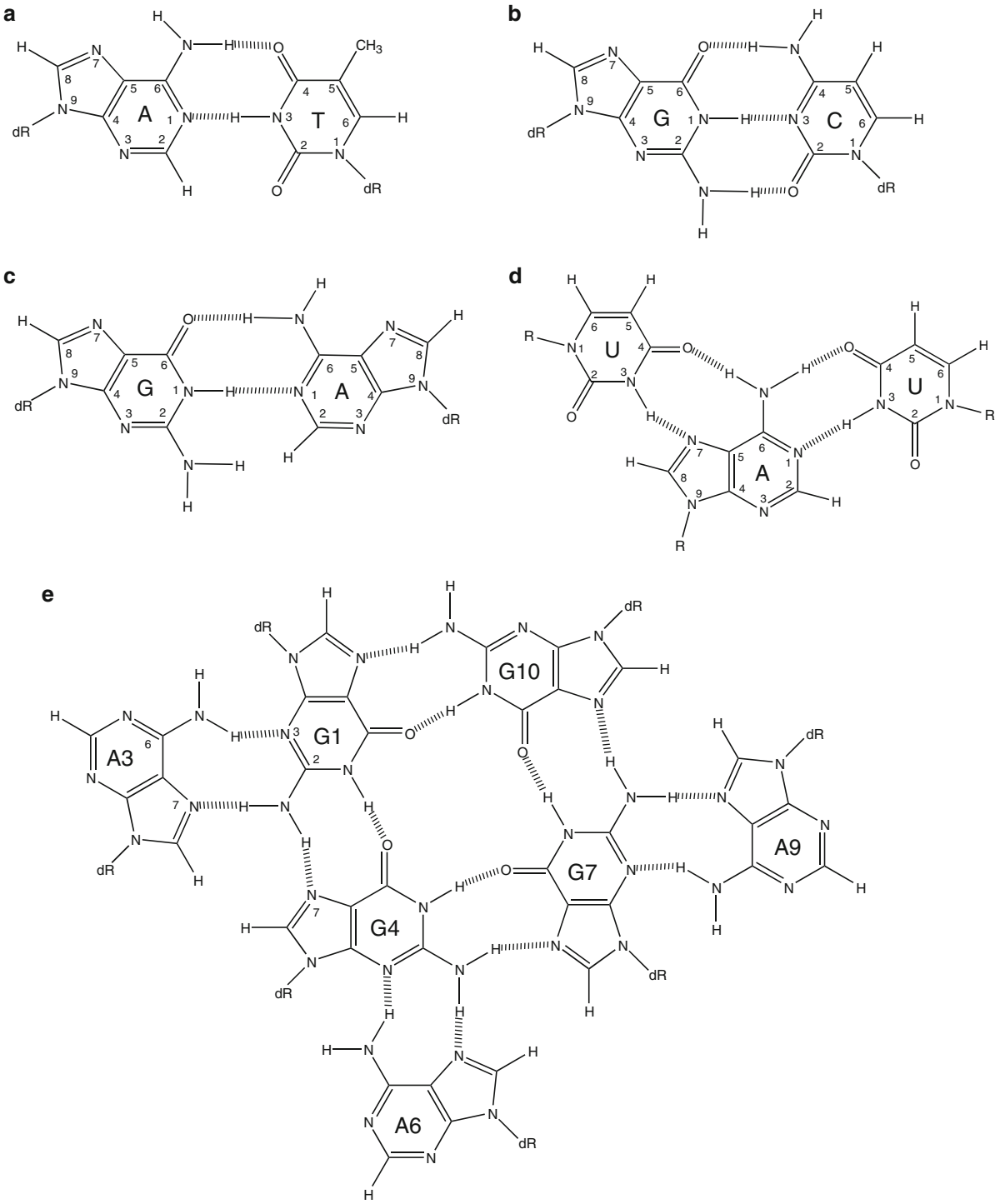
can also be obtained from ^{15}N HSQC. The amino nitrogen resonance of cytosine, adenine, and guanine appears at 94–98, 82–84, and 72–76 ppm, respectively. The residue-specific assignment of the amino protons of a cytosine is obtained from the strong intrabase NOESY cross peak with already assigned H5. Adenine amino protons are assigned from HCCNH-TOCSY on the basis of already assigned H2 and H8. Guanine amino protons are assigned from the strong intrabase NOESY cross peak with already assigned H1 (imino).

C2, C5, C6, and C8 are assigned from ^{13}C HSQC. C2 appears in a low-field region (152–156 ppm) compared to C6 and C8 (134–144 ppm), which is utilized to discriminate H2 from H6/H8 in an early stage of the assignment. C1′, C2′, C3′, C4′, and C5′ are assigned from ^{13}C HSQC, HCCH-COSY, and HCCH-TOCSY. ^{15}N HSQC gives H6–N1 (pyrimidine), H8-N7/N9 (purine), and H2–N1/N3 (adenine) correlations, which is utilized to discriminate H2, H6, and H8 in an early stage of the assignments.

Finally, site-specific incorporation of 1–2% ^{13}C- or ^{15}N-labeled nucleotide at each position of DNA/RNA with a synthesizer is also a powerful and cost-effective tool for the assignment (Phan and Patel 2002).

Identification of Base Pairs

For an A:T base pair (Fig. 4a), a strong AH2-T(U)H3 (imino) NOESY cross peak is observed. For a G:C base pair (Fig. 4b), strong CH41/CH42 (amino)-GH1 (imino) NOESY cross peaks are observed. These cross peaks indicate the formation of the either A:T or G:C base pair.



Nucleic Acid NMR – Introduction, Fig. 4 *Base pairs*. (**a**) A:T base pair, (**b**) G:C base pair, (**c**) G:A base pair, (**d**) U:A:U base triple, and (**e**) G(:A):G(:A):G(:A):G heptad

A base pair can be identified more directly and solidly with ▶ HNN-COSY by observing a spin-spin coupling across a ▶ hydrogen bond involved in the formation of the base pair (Grzesiek et al. 2001). In the case of the A:T base pair, the T(U)H3 (imino)-AN1 correlation peak appears due to the existence of the ▶ spin-spin coupling, $^{2h}J_{NN}$, between T(U)N3 and AN1. In the case of the G:C base pair, the GH1 (imino)-CN3 correlation peak appears due to the existence of the $^{2h}J_{NN}$ coupling between GN1 and CN3. Even if exchangeable T(U)H3 (imino) is not observed for an unstable A:T(U) base pair, the base pair formation can still be deduced from the observation of the AH2 (non-exchangeable)-T(U)N3 correlation peak originating from the existence of the $^{2h}J_{NN}$ coupling between AN1 and T(U)N3.

Nonstandard base pairs can also be identified with HNN-COSY. For a G:A base pair (Fig. 4c), the GH1 (imino)-AN1 correlation peak appears due to the $^{2h}J_{NN}$ coupling between GN1 and AN1. In the case of a U:A:U base triple (Fig. 4d), the UH3 (imino)-AN1 correlation peak appears due to the UN3 (imino)-AN1 $^{2h}J_{NN}$ coupling for a Watson-Crick base pair, while the UH3 (imino)-AN7 correlation peak appears due to the UN3 (imino)-AN7 $^{2h}J_{NN}$ coupling for a Hoogsteen base pair. More complex architecture such as a heptad (Fig. 4e) (Matsugami et al. 2001) can also be identified with HNN-COSY. Observation of the G1H21/H22 (amino)-A3N7 correlation peaks originating from the G1N2-A3N7 $^{2h}J_{NN}$ coupling and the A3H61/H62 (amino)-G1N3 correlation peaks originating from the A3N6-G1N3 $^{2h}J_{NN}$ coupling indicates the formation of the G1:A3 base pair. Presence of the G1H21/H22 (amino)-G4N7 correlation peaks originating from the G1N2-G4N7 $^{2h}J_{NN}$ coupling indicates the formation of the G1:G4 base pair. In this way, the formation of the heptad was solidly proved (Sotoya et al. 2004).

As an application of HNN-COSY, a method for the direct unambiguous discrimination between intra- and intermolecular hydrogen bonds in symmetric multimers has been developed (Sotoya et al. 2004). This method relies on the extent of the reduction of the HNN-COSY cross peak intensity between fully labeled, and an equimolar mixture of labeled and non-labeled samples. Irrespective of the level of multimerization, this method can clearly discriminate the intra- and intermolecular hydrogen bonds.

Structural Restraints

Distance restraints are the most powerful source for the structure determination of DNA/RNA (Varani et al. 1996; Cromsigt et al. 2001; Furtig et al. 2003). Under the two-spin approximation, distance restraints are calculated from

$$r_{ij} = r_{ref}\left(NOE_{ref}/NOE_{ij}\right)^{1/6}$$

where r_{ij} is the distance between proton i and j, r_{ref} is a reference distance, NOE_{ref} is the intensity of a cross peak of NOESY recorded with a short mixing time (~50 ms) between a reference proton pair, and NOE_{ij} is that between protons i and j. As the reference proton pair, H5 and H6 of a cytosine (reference distance of 2.47 Å) or H1′–H2″ (2.20 Å) is used. The calculated distance contains errors due to spectral noise, spin diffusion, and local dynamics. Therefore, a certain distance is added to the calculated one to set an upper limit distance. For example, 0.8 Å is added for the calculated distance of 0–2.7 Å, 1.5 Å is added for the 2.7–3.2 Å distance, 2.0 Å is added for the 3.2–4.0 Å distance, and 2.5 Å for 4.0–6.0 Å. For a cross peak that can be observed only in NOESY with a longer mixing time (~200 ms), 0.5 Å is further added to the upper limit distance. The lower limit distance is set to 1.8 Å (Matsugami et al. 2001). For exchangeable protons, WATERGATE NOESY or JRNOESY is used to obtain the distance (Varani et al. 1996).

The structure of each nucleotide unit can be described with the seven dihedral angles, α-ζ for a backbone and χ around the glycoside bond (Fig. 2a). δ is related to sugar puckering. Because of the experimental difficulty to obtain reliable dihedral angle restraints, inclusion of these restraints into the structure determination is limited to relatively small DNA/RNA. Qualitative analysis of $^{3}J_{H5'P}$ and $^{3}J_{H5''P}$ based on Het-COR gives restraints for β for a certain case (Varani et al. 1996). Qualitative analysis of $^{3}J_{H4'H5'}$ and $^{3}J_{H4'H5''}$ based on DQF-COSY gives restraints for γ (Varani et al. 1996). The sugar pucker can be determined from the analysis of $^{3}J_{H1'H2'}$, $^{3}J_{H2''H3'}$, and $^{3}J_{H3'H4'}$ derived from DQF-COSY and E. COSY (Varani et al. 1996). It is also useful to know that C3′ and C4′ resonances appear in the low-field region for C2′-*endo* in comparison with the case of C3′-*endo*. Thus, restrains for δ related to the

sugar puckering are obtained. Qualitative analysis of $^{3}J_{H3'(i)P(i+1)}$ based on Het-COR gives restraints for ε (Varani et al. 1996; Cromsigt et al. 2001; Furtig et al. 2003; da Silva 2007).

When the mode of a base pair is identified as described above, distance restraints between donor and acceptor atoms of the hydrogen bonds are included.

Planarity restraint that forces all atoms involved in a base pair to locate onto one plane could also be included.

When a molecule is weakly aligned in solution, the dipolar-dipolar interaction does not average to zero anymore, but the ▶ residual dipolar coupling (RDC) can be detected. The RDC between i and j spins, D_{ij}, is given by

$$D_{ij}(\theta,\psi) = D_a\{(3\cos^2\theta - 1) + 3/2\ R\sin^2\theta\cos 2\psi\} \quad (1)$$

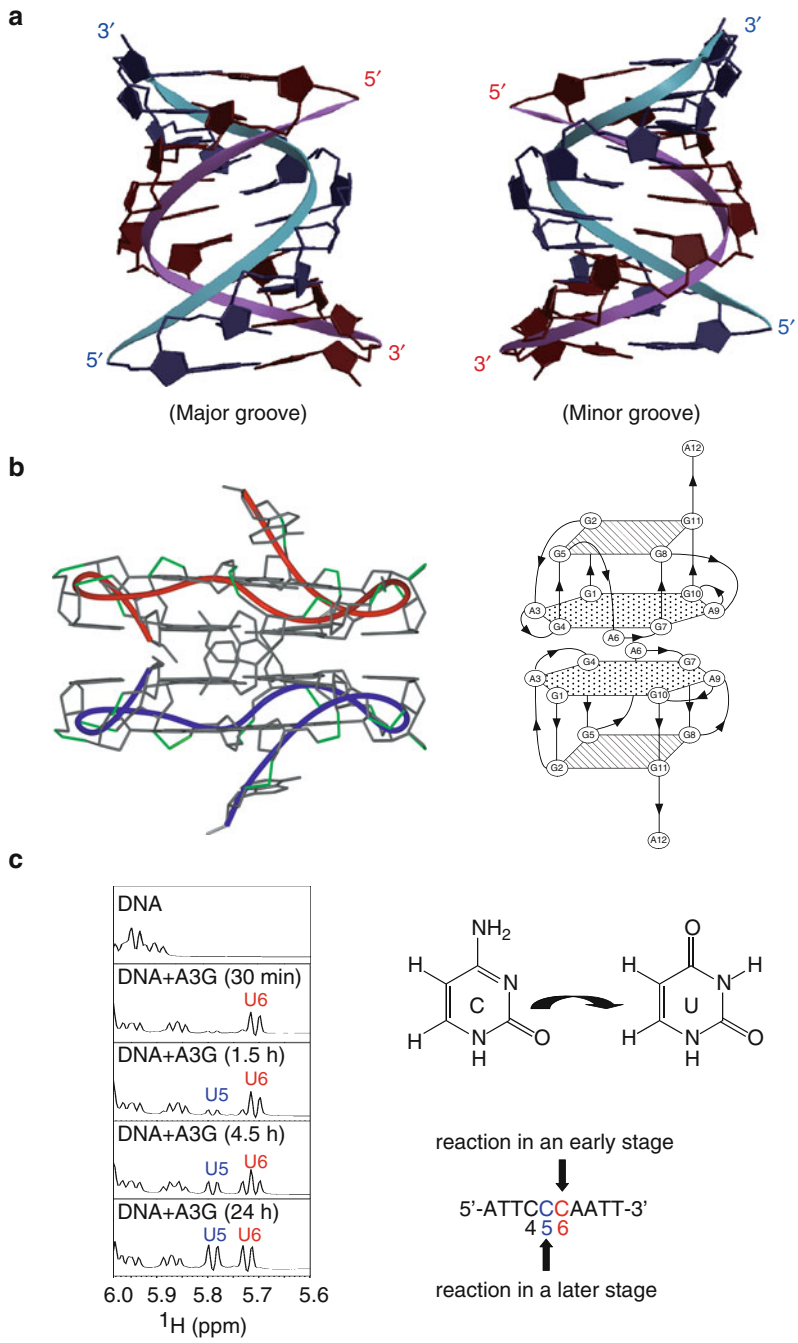
where D_a is the magnitude of the dipolar coupling tensor, R is the rhombicity, and θ and ψ are polar coordinates that describe the orientation of the ij vector in the principal axis system of the molecular alignment tensor (Bax et al. 2001). When five parameters, D_a, R, and three Euler angles that define the orientations of principal axes of the molecular alignment tensor are determined, the experimentally obtained RDC value for the ij vector such as either a directly bonded N-H vector or a C-H vector restrains the orientation of the corresponding vector according to the Eq. 1. The restraint for the orientation of each vector is imposed against the common frame (the molecular alignment tensor); therefore these restraints can be regarded as "global restraints" in contrast to the distance restraints derived from NOE which are "local restraints." The filamentous bacteriophage Pf1 aligns in solution due to its magnetic anisotropy. Then, when the Pf1 phage (10–20 mg/mL) is added to DNA/RNA solution, the weak alignment of DNA/RNA is achieved due to the steric interaction of DNA/RNA with the aligned Pf1 phage. The negatively charged Pf1 is useful to align DNA/RNA that is also negatively charged, because the unfavorable attachment of DNA/RNA to the alignment medium can be avoided by electrostatic repulsion. The RDC values for one-bond N–H and C–H can be derived as the difference in the coupling (scalar coupling plus RDC) between under the isotropic and aligned conditions. 1D spectrum, HSQC without decoupling along the indirect dimension, IPAP (inphase/antiphase) HSQC and ▶ TROSY are used to obtain the RDC values (Bax et al. 2001; Lukavsky and Puglisi 2005).

Cations play a critical role for the formation of certain DNA/RNA structures. For example, K^+ plays a critical role for the formation of the quadruplex. NH_4^+ can be used to probe monovalent ion binding to DNA/RNA. Observation of intermolecular NOE between the ammonium ion protons and those of DNA/RNA gives the information on how cations bind to and stabilize the certain DNA/RNA structure (Feigon et al. 2001).

Structure Determination

Experimentally obtained structural restraints are transformed into pseudo-energy terms, which are combined with empirical energy terms to give the total energy. With this total energy, the restrained molecular dynamics is run. The structure determination is performed with simulated annealing protocol in which the system is heated to high temperature (2,000–3,000 K) and gradually cooled down to low temperature. Through this process, the molecule overcomes the energy barrier to escape from the local minimum and reaches the global minimum. The programs such as XPLOR-NIH, ▶ CYANA, ▶ CNS, and AMBER are used for this calculation. Ten to twenty final structures are selected from 100 to 200 calculations on the basis of the criterion of the smallest residual total energy. The quality of the obtained structures is evaluated on the basis the root mean square deviation (r.m.s.d.) and violations of restraints. Reasonably good structures give the r.m.s.d. less than 0.5 Å for all heavy atoms, no distance restraints violation greater than 0.5 Å, and no dihedral angle restraints violation greater than 5° (Varani et al. 1996; Cromsigt et al. 2001; Furtig et al. 2003).

Because the proton density of DNA/RNA is lower than that of protein, there is less chance to obtain the distance restraints derived from NOE for DNA/RNA than for protein. Additionally, the shape of DNA/RNA tends to be rod-like rather than globular, and then the distance restraints are likely to be "local" rather than "global." Therefore, the inclusion of the structural restraints derived from RDC could be more critical



Nucleic Acid NMR – Introduction, Fig. 5 *Examples of nucleic acid structures and monitoring of base conversion*. (**a**) Structure of double-stranded 4′-thioDNA with the sequence of d (CGCGAATTCGCG) (Matsugami et al. 2008). (**b**) Dimeric structure of RNA aptamer, R12, against a prion protein with the sequence of r (GGAGGAGGAGGA) (Mashima et al. 2009). (**c**) Real-time monitoring of cytosine to uracil conversion in single-stranded DNA with the sequence of d (ATTCCCAATT) through deamination by anti-HIV enzyme, APOBEC3G (Furukawa et al. 2009)

for the precise structure determination of DNA/RNA than for that of protein. In the case of a globular protein, it can be assumed that the directions of NH and CH vectors are uniformly distributed to all orientations. Under this assumption, D_a and R can be determined on the basis of the histogram of the observed RDC for many NH and CH pairs. Then just three Euler angles, that define the orientations of principal axes of the molecular alignment tensor, should be included into the structural calculation as newly added variables in order to incorporate structural restraints derived from RDC (Bax et al. 2001). In the case of DNA/RNA that tends to exhibit a rod-like shape, however, the uniform orientation of NH and CH vectors cannot

be assumed and thus D_a and R cannot be determined in the same way. Alternatively, the structure calculation is carried out without the RDC restraints in an early phase to obtain a tentative structure. Then, the optimum combination of D_a and R that gives the best agreement between the experimentally obtained RDCs and the RDCs predicted for the tentative structure is determined with a grid search. In a later phase, the RDC restraints are included into structure calculation with the determined D_a and R (Zweckstetter and Bax 2000; Bax et al. 2001; McCallum and Pardi 2003).

Examples of DNA/RNA Structures and Real-Time Monitoring of Base Conversion by Anti-HIV Enzyme

Figure 5a shows the structure of double-stranded 4′-thioDNA comprising 2′-deoxy-4′-thionucleotides with the sequence of d(CGCGAATTCGCG) (Matsugami et al. 2008). Most sugars take on the C3′-*endo* conformation. The major groove is narrow and deep, while the minor groove is wide and shallow. Thus, 4′-thioDNA takes on A-form characteristic of RNA, both locally and globally, rather than conventional B-form. The unique biochemical properties of 4′-thioDNA, including nuclease resistance that is a promising character as a functional oligonucleotide, are rationalized in the light of the elucidated structure.

Figure 5b shows the dimeric structure of 12-mer RNA, r(GGAGGAGGAGGA) (R12), which functions as an aptamer against a prion protein (Mashima et al. 2009). R12 forms an intramolecular parallel quadruplex under physiological ionic conditions. The quadruplex contains G:G:G:G tetrad and G(:A):G:G(:A):G hexad planes. Two quadruplexes form a dimer through intermolecular hexad-hexad stacking. The structure has suggested how the RNA aptamer, R12, traps the prion protein with high affinity. The atomic coordinates of R12 would be useful for the development of R12 as a therapeutic agent against prion diseases and Alzheimer's disease.

Figure 5c shows time course of ^{1}H NMR spectra indicating deamination of two cytidine residues of single-stranded DNA, d(ATTCCCAATT), in a strict 3′→5′ order by human ▶ anti-HIV enzyme, APOBEC3G, in an NMR tube (Furukawa et al. 2009). The enzymatic reaction is started by the addition of APOBEC3G and monitored in real time with NMR signals. This result indicates how APOBEC3G converts the cytidine base of the minus strand DNA of HIV-1 and destroys the genetic information in order to exert the anti-HIV activity.

Summary

The DNA/RNA structure in crystal does not always reflect that in solution. NMR is the only way to provide the DNA/RNA structure in solution at an atomic resolution. Therefore, this methodology is crucial to understand various biological events that need the knowledge on the DNA/RNA structure. The mode of a base pair, that is key information to construct the structure, can be solidly determined on the basis of the observation of a spin-spin coupling across a hydrogen bond with HNN-COSY. RDCs obtained for weakly aligned DNA/RNA in solution can provide valuable "global" structural information. In combination with these structural restraints, distance and dihedral angle restraints can provide a precise DNA/RNA structure in solution.

Cross-References

- ▶ Cross-Correlated Relaxation
- ▶ Cryogenic NMR Probes
- ▶ CYANA
- ▶ INEPT
- ▶ Infrared Spectroscopy of Protein Dynamics: Ultrafast Kinetics
- ▶ J Coupling
- ▶ Modeling of RNA Molecules
- ▶ Molecular Dynamics Simulations of Lipids
- ▶ Multidimensional NMR Spectroscopy
- ▶ NMR
- ▶ NMR in Drug Discovery – Introduction
- ▶ NMR Methods for Kinetic Analysis
- ▶ NMR Studies of Macromolecular Interactions – Introduction
- ▶ NMR-based Structural Proteomics
- ▶ Nuclear Overhauser Effect
- ▶ Protein NMR – Introduction
- ▶ Total Correlation Spectroscopy (TOCSY) in NMR Protein Structure Determination
- ▶ TROSY

References

Bax A, Kontaxis G, et al. Dipolar couplings in macromolecular structure determination. Methods Enzymol. 2001;339:127–74.

Cromsigt J, van Buuren B, et al. Resonance assignment and structure determination for RNA. Methods Enzymol. 2001;338:371–99.

da Silva MW. NMR methods for studying quadruplex nucleic acids. Methods. 2007;43:264–77.

Feigon J, Butcher SE, et al. Solution nuclear magnetic resonance probing of cation binding sites on nucleic acids. Methods Enzymol. 2001;338:400–43.

Furtig B, Richter C, et al. NMR spectroscopy of RNA. Chembiochem. 2003;4:936–62.

Furukawa A, Nagata T, et al. Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. EMBO J. 2009;28:440–51.

Grzesiek S, Cordier F, et al. Scalar couplings across hydrogen bonds. Methods Enzymol. 2001;338:111–33.

Lukavsky PJ, Puglisi JD. Structure determination of large biological RNAs. Methods Enzymol. 2005;394:399–416.

Mashima T, Matsugami A, et al. Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein. Nucleic Acids Res. 2009;37:6249–58.

Matsugami A, Ouhashi K, et al. An intramolecular quadruplex of (GGA)4 triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. J Mol Biol. 2001;313:255–69.

Matsugami A, Ohyama T, et al. Unexpected a-form formation of 4′-thioDNA in solution, revealed by NMR, and the implications as to the mechanism of nuclease resistance. Nucleic Acids Res. 2008;36:1805–12.

McCallum SA, Pardi A. Refined solution structure of the iron-responsive element RNA using residual dipolar couplings. J Mol Biol. 2003;326:1037–50.

Nakano S, Mashima T, et al. Structural aspects for the recognition of ATP by ribonucleotide receptors. J Am Chem Soc. 2011;133:4567–79.

Phan AT, Patel DJ. A site-specific low-enrichment ^{15}N, ^{13}C isotope-labelling approach to unambiguous NMR spectral assignments in nucleic acids. J Am Chem Soc. 2002;124:1160–1.

Sotoya H, Matsugami A, et al. Method for direct discrimination of intra- and intermolecular hydrogen bonds, and characterization of the G(:A):G(:A):G(:A):G heptad, with scalar couplings across hydrogen bonds. Nucleic Acids Res. 2004;32:51113–18.

Varani G, Aboul-ela F, et al. NMR investigation of RNA structure. Prog Nucl Magn Reson Spectrosc. 1996;29:51–127.

Werner MH, Gupta V, et al. Uniform $^{13}C/^{15}N$-labeling of DNA by tandem repeat amplification. Methods Enzymol. 2001;338:283–304.

Zimmer DP, Crothers DM. NMR of enzymatically synthesized uniformly ^{13}C, ^{15}N-labeled DNA oligonucleotides. Proc Natl Acad Sci USA. 1995;92:3091–5.

Zweckstetter M, Bax A. Prediction of sterically induced alignment in a dilute crystalline phase: aid to protein structure determination by NMR. J Am Chem Soc. 2000;122:3791–2.

Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA

Nieng Yan
School of Medicine, Tsinghua University, Beijing, China

Synonyms

NCS2 (nucleobase-cation-symporter-2) or NAT (nucleobases-ascorbate-transporter); UraA, uracil permease in *Escherichia coli*

Definition: Based on Primary Sequence Similarity and Biological Function

NCS2 or **NAT** family (TC # 2.A.40) (http://www.tcdb.org/search/result.php?tc=2.A.40) contains over 100 known members across all kingdoms of life (Saier 2000). In bacteria and fungi, NCS2 proteins mainly transport uracil, xanthine, and other derivatives of purine or pyrimidine, whereas in mammals, the NCS2 members SVCT1 and SVCT2 (sodium-dependent vitamin C transporter) transport vitamin C (L-ascorbate).

Representative members of NCS2 family are (http://www.tcdb.org/search/result.php?tc=2.A.40.6):

2.A.40.1.1: UraA – Uracil permease in *Escherichia coli*

2.A.40.1.2: PyrP – High affinity uracil permease in *Lactococcus lactis*

2.A.40.1.3: RutG– Putative pyrimidine permease in *Escherichia coli*

2.A.40.2.1: YcpX– Purine permease in *Clostridium perfringens*

2.A.40.3.1: PubX (XanP) – Xanthine permease in *Bacillus subtilis*

2.A.40.3.2: PucJ – Uric acid permease in *Bacillus subtilis*

2.A.40.4.1: UapA – High affinity uric acid-xanthine permease in *Emericella nidulans*

2.A.40.4.2: YicE – The putative xanthine permease in *Escherichia coli*

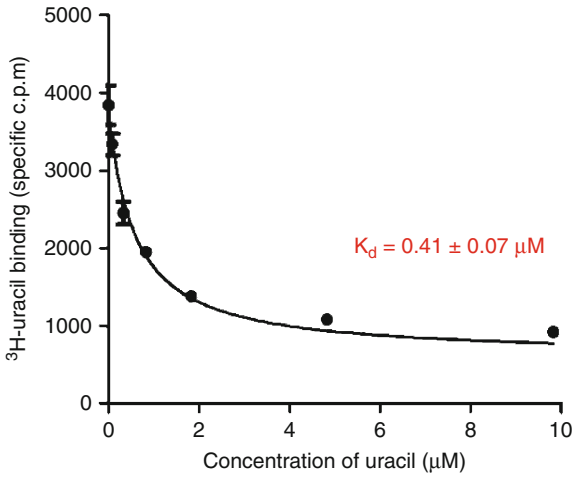
2.A.40.4.3: YgfO (XanQ) – Purine (xanthine) transporter in *Escherichia coli*

2.A.40.5.1: UapC – General purine permease in *Emericella nidulans*

2.A.40.6.1: SVCT1 – L-ascorbate: Na^+ symporter in *Rattus norvegicus*
2.A.40.6.2: SVCT2 – L-ascorbate: Na^+ symporter in *Homo sapiens*
2.A.40.6.3: LPE1 – Leaf permease protein 1, high affinity uric acid-xanthine transporter in *Zea mays*

UraA is the uracil transport protein in *E. coli*. Imported uracil is metabolized via the pyrimidine salvage pathways (Andersen et al. 1995). 5-Fluorouracil, an important anticancer drug, is also a substrate of UraA. UraA shares a sequence identity of 23.1%, 20.5%, and 21.9%, with human SVCT1, SVCT2, and rat SNBT1, respectively. The transport reaction is:

Uracil (out) + H^+ (out) → Uracil (in) + H^+ (in)



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 1 The binding affinity between uracil and WT UraA was measured by scintillation proximity assay (SPA). A competition experiment was performed in which the final concentration of ^{3}H-uracil was kept constant at 0.17 µM (40 Ci mmol^{-1}), and the concentration of non-labeled uracil was increased from 0 to 10 µM. Fitting of the data gave rise to a dissociation constant of approximately 0.41 ± 0.07 µM

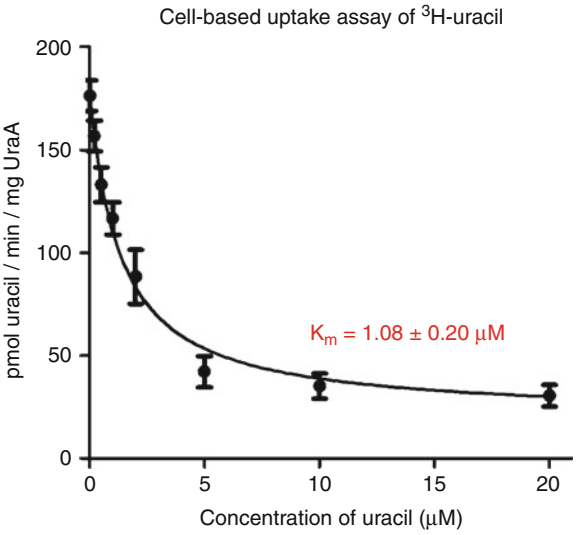
Introduction

The NAT Signature Motif

Genetic, biochemical, and in silico studies of UapA and YgfO identified a NAT signature motif comprising 11 residues, [A/G/S][Q/E/P]-N-X-G-X-X-X-X-T- [R/K/G] (Gournas et al. 2008).

Transport and Ligand Binding Assays

Uracil binds to the recombinant UraA protein with a dissociation constant (K_d) of approximately 0.41 ± 0.07 µM as measured by the scintillation proximity assay (SPA) (Fig. 1). In a cell-based uptake assay, UraA can transport ^{3}H-uracil with a K_m of 1.08 ± 0.2 µM (Fig. 2).



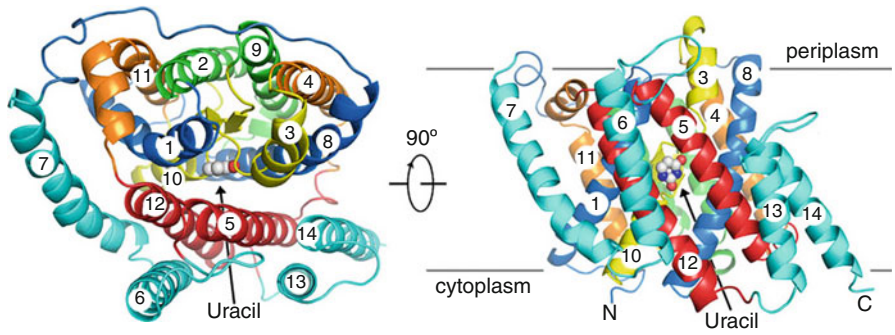
Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 2 *Uptake of ^{3}H-uracil by cells. E. coli ΔuraA* was transformed with WT-UraA in the pQLINK vector

X-ray Crystal Structure of UraA

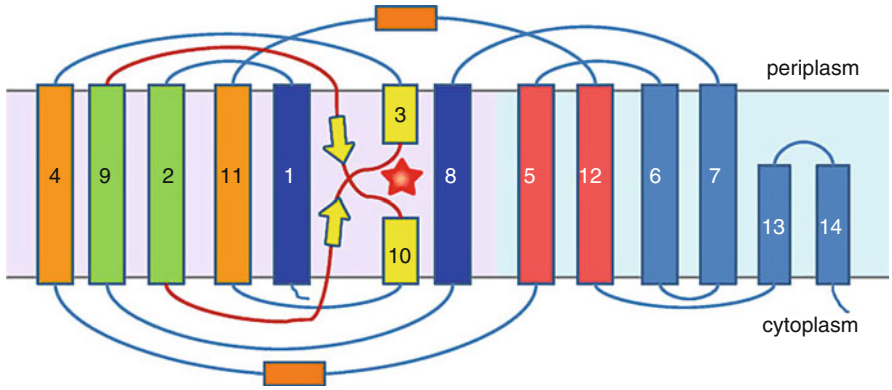
Crystals of recombinantly expressed and purified UraA were obtained in space group $P6_422$ in the presence of 0.4% β-NG (n-nonyl-β-D-glucopyranoside) detergent and diffracted X-rays beyond 2.9 Å. The structure of uracil-bound UraA was refined to 2.9 Å resolution (Fig. 3).

A Novel Transporter Fold

UraA consists of 14 membrane helices, with both N- and C-termini located on the cytoplasmic side. The two C-terminal helices α13 and α14 penetrate only halfway into the membrane and form a hairpin (Fig. 3). There is little apparent structural similarity between UraA and Mhp1, a representative member of the Nucleobase-Cation-Symport (NCS1) family (Weyand et al. 2008), despite the apparent similarity of substrates for NCS2 and NCS1 family members. An exhaustive search of the PDB using DALI suggested that UraA has a novel fold compared to all known structures of secondary active transporters (Fig. 4).



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 3 The crystal structure of UraA was determined at 2.9 Å. Two perpendicular views, one from periplasm and one from the lipid membrane, are shown. Corresponding TMs within the two inverted repeats (e.g., TM1 and TM8) are colored identically. The bound uracil is shown



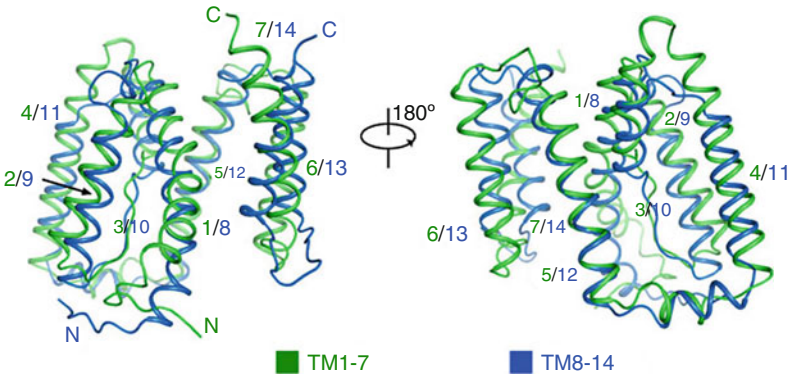
Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 4 Topology diagram of UraA. The TMs are arranged into the core domain (in *light purple background*)and the gate domain (in *light cyan background*)

The most prominent feature is placement of two short antiparallel β-strands, on TM3 and TM10 respectively, at the center of the structure (Fig. 4). Inner membrane proteins are mostly α-helical. UraA is likely the only known transporter that contains β-strands in the transmembrane region. Each β-strand is preceded by an extended loop and followed by a short α-helix, α3 and α10, respectively (Fig. 4). The 11-residue NAT motif, which was predicted to be a loop in the cytoplasm, constitutes α10 in the structure of UraA. Helices α3 and α10 are aligned as if they were a single discontinuous transmembrane helix, with the substrate bound at the breakage point (Fig. 4).

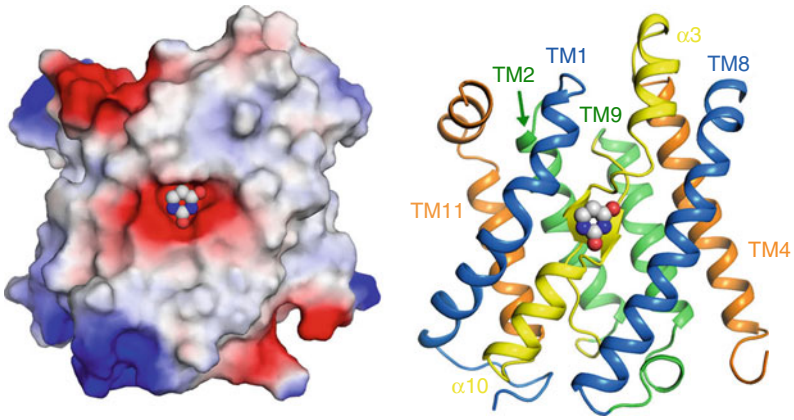
The 14 TMHs of UraA are spatially organized into two halves, which are named the core domain and the gate domain. The core domain comprises TMHs 1-4 and 8-11, whereas the gate domain includes the other 6 TMHs. Intriguingly, UraA also contains two internal, structural repeats: TMs 1-7 and TMs 8-14, which are related to each other by an approximately 180° rotation around an axis parallel to the membrane bilayer (Fig. 5). These two repeats can be superimposed with a root-mean-squared deviation (rmsd) of 2.9 Å over 129 Cα atoms.

Substrate Recognition

Uracil is located in a concave pocket formed by the antiparallel β-strands and their connecting loops in TMH3 and TMH10. The pyrimidine ring of uracil is approximately in parallel with the β-strands and surrounded by negative electrostatic potentials (Fig. 6). Recognition of the substrate is almost exclusively mediated by residues from the core domain, involving both polar and hydrophobic contacts from TM1, TM3, TM8, TM10, and TM12 (Fig. 7). Two negatively charged residues, Glu241 and Glu290, appear to anchor uracil by each making two hydrogen bonds to uracil. His245 is also located in the vicinity of uracil, likely contributing to its coordination through water-mediated hydrogen bonds or through interactions with Glu241. Furthermore, the two oxygen atoms of uracil are further hydrogen-bonded to the amide nitrogen atoms of Phe73 and Gly289.



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 5 Superimposition of the two inverted repeats within UraA



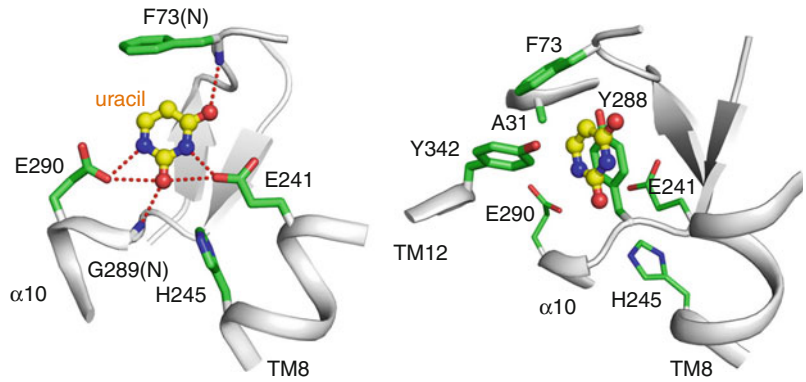
Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 6 Uracil is located in a concave within the core domain and surrounded by negative electrostatic potential. The core domain of UraA is shown in both surface electrostatic potential (*left*) and rainbow cartoon (*right*)

In addition to the polar contacts, uracil is surrounded by hydrophobic residues including Ala31 of TMH1, Phe73 of TMH3, Tyr288 of TMH10, and Tyr342 of TMH12. Phe73 insulates the substrate from the periplasm, whereas the phenyl ring of Tyr288 is roughly in parallel with the pyrimidine ring of uracil. Tyr342 is the only residue from the gate domain that contributes to uracil coordination (Fig. 7).

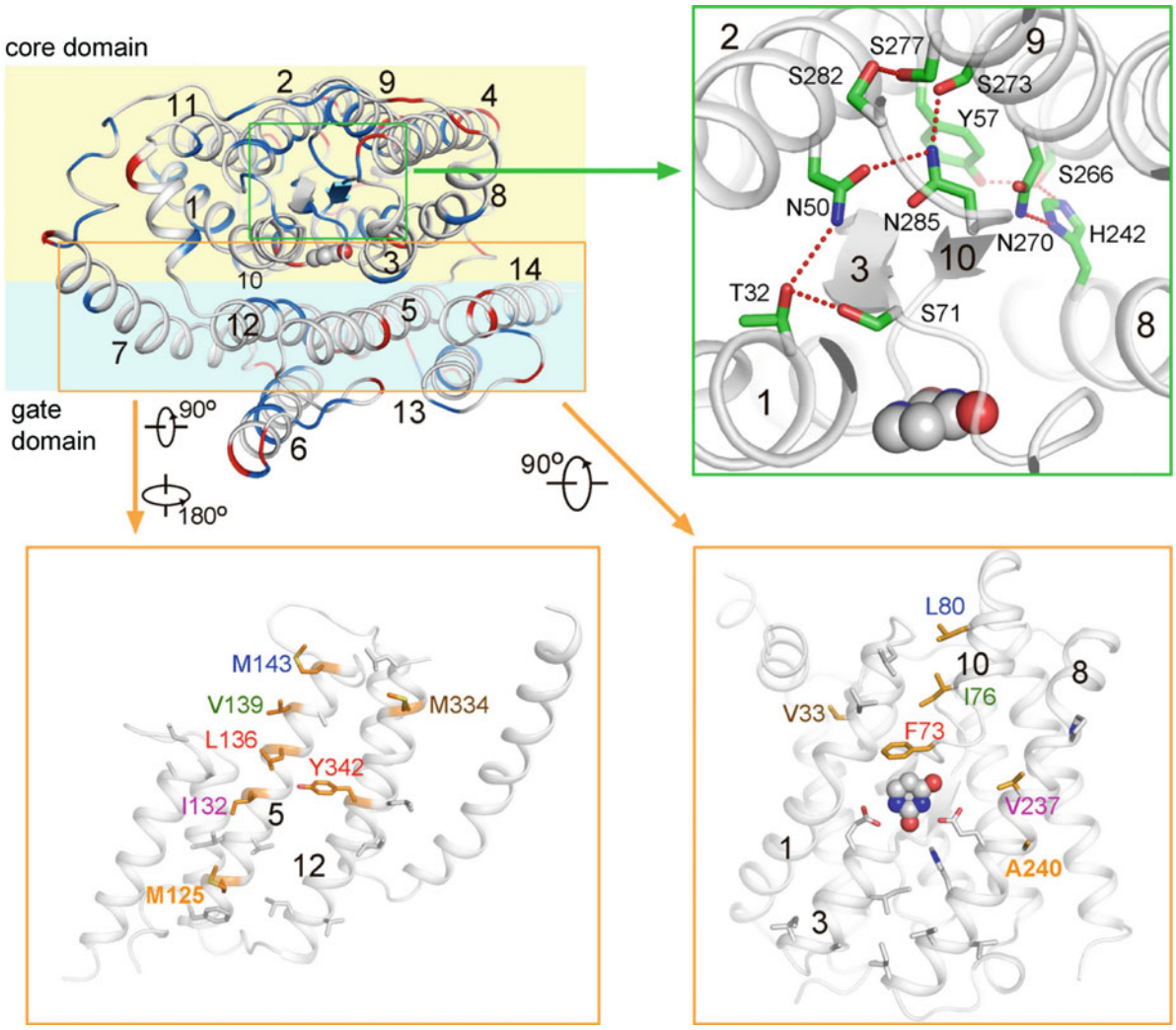
Domain Organization

The core and gate domains are approximately in parallel to each other with a planar interface (Fig. 8). Surprisingly, there are a large number of buried hydrogen bonds within the core domain, with the β-strands of TMH3 and TMH10 serving as the organizing center for these interactions. The side chain of Asn285 in TMH10 forms two hydrogen bonds with Asn50 in TMH2 and Ser273 in TMH9. The hydroxyl groups of Ser277 and Ser282 form an extra hydrogen bond between TMH9 and TMH10. The hydroxyl group of Thr32 in TMH1 makes two hydrogen bonds with the side chains of Asp50 in TMH2 and Ser71 in TMH3. In addition, the imidazole nitrogens of His242 in TMH8 are hydrogen-bonded to Ser266 and Asn270 in TMH9. This network of interactions may help maintain the conformation of the extended loops in TMH3 and TMH10 and restrain the core domain TMHs from intra-domain movement.

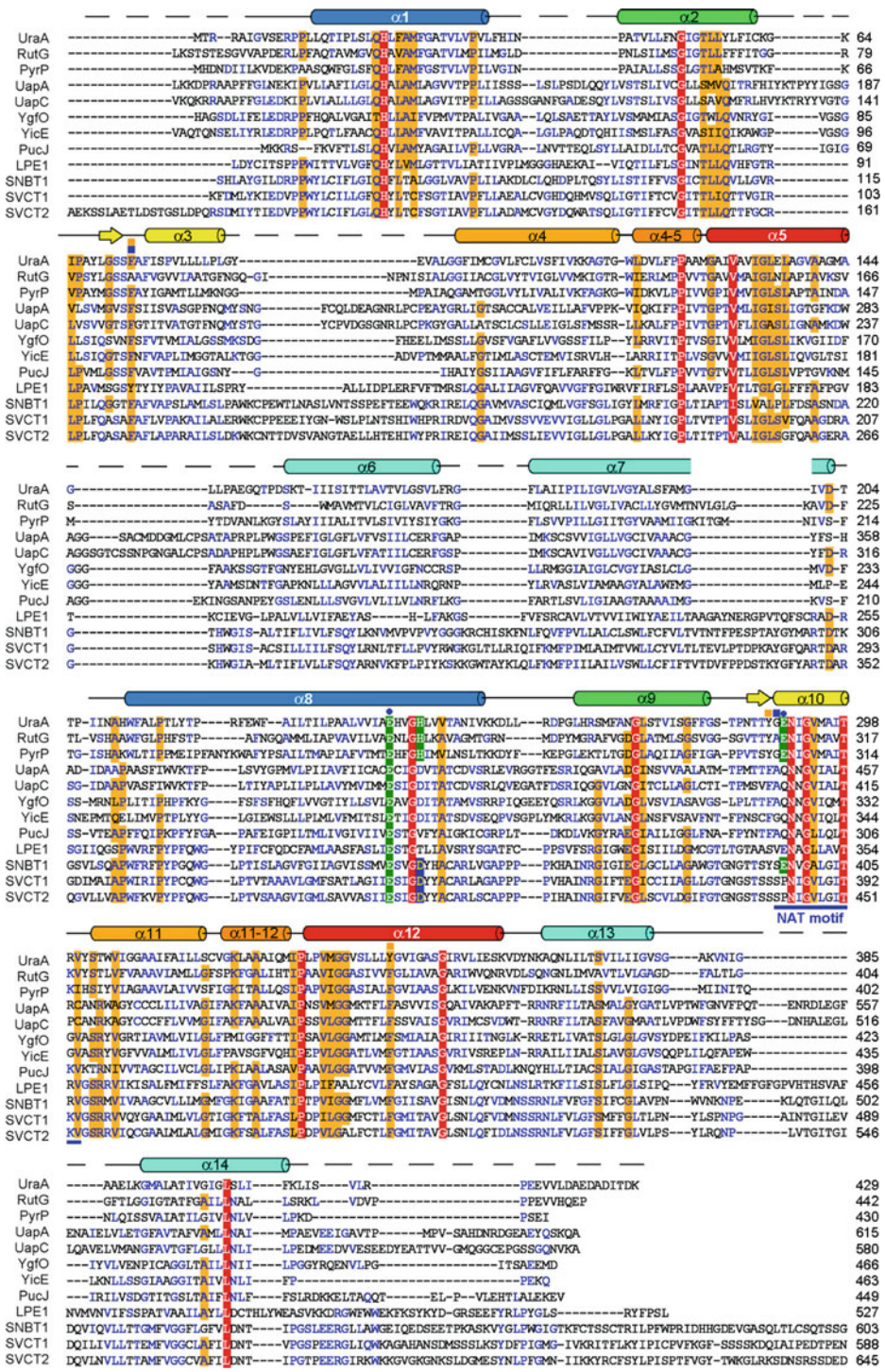
In contrast to the extensive polar interactions within the core domain, the inter-domain interface is populated mainly with hydrophobic residues from TMHs 1, 3, 8, 10 in the core domain and TMHs 5, 12 in the gate domain (Fig. 8). The contacts are mostly located on the periplasmic side of uracil, suggesting an inward-facing conformation of the structure. Except the two uracil-binding Glu residues and His245, the substrate-facing surfaces of the core and gate domains are predominantly composed of hydrophobic residues, a feature that might be important for the transport mechanism of UraA.



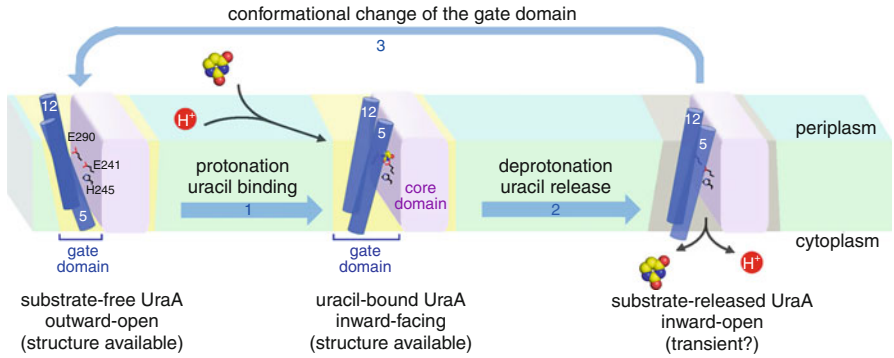
Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 7 Uracil is coordinated by both polar (*left*) and van der Waals (*right*) contacts. Hydrogen bonds are shown in *red dashes*. Uracil is shown in *yellow ball-and-sticks*



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 8 Domain organization of UraA. *Upper left panel*: A periplasmic view is shown. The charged residues (Asp, Glu, Lys, and Arg) are shown in *red*, and polar residues (Ser, Thr, Asn, Gln, His, and Tyr) are shown in blue. Uracil is shown in *white* spheres. *Upper right panel*: The network of hydrogen bonds within the core domain. *Lower panels*: The substrate-facing surfaces of the core and the gate domains are shown. The corresponding residues that contact each other on the interface are labeled with the same colors



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 9 Sequence alignment of the *E. coli* UraA with NAT homologs from other species and organisms. Secondary structural elements of UraA are indicated above the sequence alignment. Invariant and highly conserved amino acids are shaded in red and orange, respectively. The residues that are hydrogen-bonded with uracil with their side chains and main chains are indicated by blue spheres and blue squares, respectively. The residues that bind to uracil with van der Waals contacts are indicated by *orange squares*



Nucleobase-Cation-Symport-2 Family and the Uracil: Proton Symporter, UraA, Fig. 10 A working model to illustrate the putative transport mechanism of UraA. Glu241, His245, and Glu245 are shown to emphasize their essential role in substrate recognition and proton translocation. For simplicity, only TM5 and TM12 are shown to represent the gate domain

Putative Transport Mechanism

UraA is a proton-coupled symporter, in which the translocation of proton depends on the residues that are capable of protonation and deprotonation. A remarkable revelation of the structure is the clustering of Glu241, His245, and Glu290 (Fig. 7), all of which are able to undergo cycles of protonation and deprotonation. In particular, Glu241 and Glu290 are both directly committed to uracil recognition, whereas His245 may be involved in uracil binding through water-mediated hydrogen bond. This arrangement suggests that these residues may play a key role in both substrate recognition and proton translocation. Only Glu241 is invariant among all NAT members. His245 is replaced by Asp, Thr, or Val in other NAT proteins, whereas Glu290 is substituted by Gln or Pro (Fig. 9). A closer examination, however, revealed function-based conservation. Glu290 is conserved in the three known pyrimidine permeases RutG, PyrP, and SNBT1, suggesting a conserved mode of recognition for pyrimidine. In addition, His245 is conserved in RutG and PyrP, both of which are pyrimidine:proton symporters. By contrast, in the sodium symporters SNBT1, SCVT1, and SCVT2, His245 is replaced by a conserved Asp residue (Fig. 9). This analysis strongly suggests conserved mechanisms of transport, with His245 likely playing a key role in proton symport and Asp playing a similar role in sodium symport.

The structure of uracil-bound UraA represents an inward-open conformation. In order to take up uracil from the environment, UraA has to adopt at least one more conformation, that is, the one open to the periplasm. The distinct domain organization of UraA suggests a simple operation involving a rigid body rotation of the gate domain relative to the core domain centering the bound uracil in order to achieve the outward-open conformation (Fig. 10), although subtle conformational changes of certain residues such as Phe73 may be inevitable.

It is noteworthy that the two Glu residues and His245 are surrounded by hydrophobic residues (Fig. 9). Therefore, by analogy to LacY (Abramson et al. 2003), we speculate that these residues are likely to be neutral in the outward-open conformation of UraA, with Glu protonated. Based on this assumption, we propose a working model to explain the proton-coupling and symport mechanism of uracil by UraA (Fig. 10). The default conformation of the substrate-free UraA may be outward-open and protonated. The absence of substrate binding makes closure of the gate domain onto the core domains energetically unfavorable, due to the two Glu residues. Upon binding of substrate, the gate domain undergoes a conformational change, achieving an inward-open conformation as seen in the uracil-bound structure. A proton translocation is likely to occur among the Glu and His residues. Deprotonation may cause local conformational changes around Glu241, His245, and Glu290, and lead to the dissociation of uracil. An inward-open, deprotonated conformation is probably unstable, and thus only exists transiently. It quickly resumes the outward-open and protonated state. In the sodium symporters of the NAT family, the sodium ion, instead of proton, is required to neutralize the conserved Glu and Asp residues and for the binding of the substrate.

This model predicts that the core domain provides the molecular basis for substrate selectivity and proton or sodium translocation, whereas conformational changes of the gate domain allow transport of the substrate molecule. We recognize the speculative nature of this model, as many important questions remain unanswered. For example, what is the molar ratio between proton/sodium ion and the substrate molecule during each transport cycle? How do the conserved Glu and His residues trigger conformational changes during cycles of protonation and deprotonation? Many more biochemical experiments, as well as a high-resolution structure of substrate-free UraA, are required to address these questions. Nonetheless, the structural and biochemical characterizations of UraA reported here provide an important framework for the mechanistic understanding of the NAT family transporters.

References

Abramson J, Smirnova I, et al. Structure and mechanism of the lactose permease of *Escherichia coli*. Science. 2003;301:610–5.

Andersen PS, Frees D, et al. Uracil uptake in *Escherichia coli* K-12: isolation of uraA mutants and cloning of the gene. J Bacteriol. 1995;177:2008–13.

Gournas C, Papageorgiou I, et al. The nucleobase-ascorbate transporter (NAT) family: genomics, evolution, structure-function relationships and physiological role. Mol Biosyst. 2008;4:404–16.

Saier Jr MH. Families of transmembrane sugar transport proteins. Mol Microbiol. 2000;35:699–710.

Weyand S, Shimamura T, et al. Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. Science. 2008;322:709–13.

Nutlins

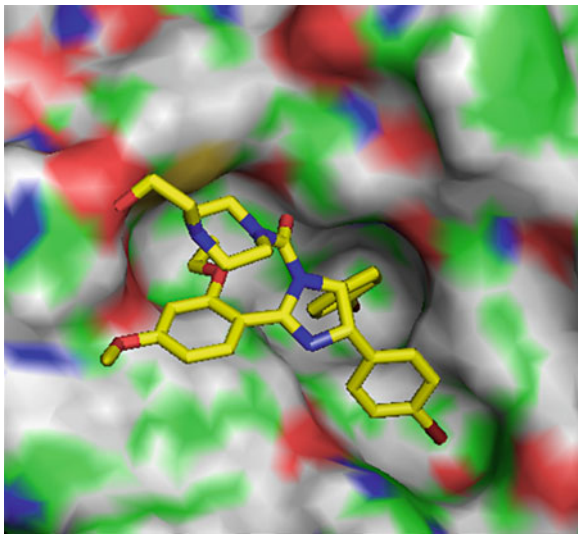
Lyubomir T. Vassilev
Roche Research Center, Hoffmann-La Roche Inc, Nutley, NJ, USA

Synonyms

Cis-imidazoline MDM2 antagonists

Definition

The tumor suppressor p53 plays a pivotal role in protecting us from cancer development (Vogelstein et al. 2000). It is a potent transcription factor which is activated following stress and regulates multiple genes implicated in cell cycle control, DNA repair, apoptosis, and senescence. Owing to its central role as a cellular gatekeeper, the p53 pathway is the most frequent target of genetic alterations in cancer. Approximately half of all human tumors express a dysfunctional p53 protein as a result of a mutation or deletion in its gene. In non-stressed cells, p53 levels are tightly regulated by another cellular protein, MDM2 (murine double minute 2). MDM2 controls p53 stability and activity through a feedback mechanism by which both proteins mutually regulate their cellular levels. p53 activates the transcription of *mdm2* gene while the MDM2 protein binds to p53 and physically blocks its N-terminal transactivation domain. MDM2 is also a p53-specific E3 ligase which targets the tumor suppressor for ubiquitin-dependent degradation. This feedback loop is frequently dysregulated in cancer. Many tumors overproduce MDM2 protein and effectively impair p53 function. Therefore, restoration of p53 activity by inhibiting the p53–MDM2 interaction represents an attractive new approach to cancer therapy. X-ray crystallography studies have shown that the transactivation domain of p53 forms an amphiphilic α-helix that projects Phe^{19}, Trp^{23}, and Leu^{26} residues into a relatively deep hydrophobic binding pocket on the MDM2 surface (Kussie et al. 1996). Nutlins are synthetic *cis*-imidazoline derivatives designed to bind tightly into the p53 pocket of MDM2 (Vassilev et al. 2004). They inhibit the formation of p53–MDM2 complexes in vitro with IC_{50} in the 100–300 nM range. The crystal structure of MDM2-nutlin-2 complexes has revealed that nutlins project functional groups into the binding pocket that effectively mimic the interaction of the three p53 amino acids critical for the p53–MDM2 binding (Fig. 1). Only one of the two nutlin enantiomers can bind to MDM2 with high affinity (e.g., nutlin-3a) while the other enantiomer (nutlin-3b) has 150–200-fold lower affinity and is thus practically inactive. Nutlins inhibit p53–MDM2 binding in cancer cells, leading to stabilization of p53 and activation of the p53 pathway, cell cycle arrest, and apoptosis. They have shown a remarkable selectivity for the p53–MDM2 interaction. Their effects have been observed only in cells expressing wild-type p53 but not



Nutlins, Fig. 1 X-ray structure of nutlin-2 bound to the p53 pocket of MDM2 protein

in p53 mutant or null cells. Treatment of established human cancer xenografts with nontoxic doses of nutlin-3 have led to effective tumor growth inhibition and regression. (Vassilev et al. 2004).

Cross-References

- ▶ Ubiquitinylation
- ▶ X-Ray Diffraction and Crystallography of Oligosaccharides and Polysaccharides

References

Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science. 1996;274:948–53.

Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science. 2004;303:844–8.

Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature. 2000;408:307–10.