Chapter 12
Future Prospects

Norma M. Allewell, Igor A. Kaltashov, Linda O. Narhi, and Ivan Rayment

Abstract This chapter illustrates the dynamic, evolving nature of molecular biophysics by providing perspectives on future prospects in three major areas: X-ray and neutron scattering, mass spectrometry, and therapeutic drug development. In all three areas, major advances in the biological sciences, development of powerful new experimental and computational tools, and urgent real-world challenges are driving rapid progress. These developments have enabled and encouraged biophysicists to focus increasingly on studying systems of various sizes and the interactions between their components, rather than simply on their isolated constituents. As the examples demonstrate, these interactions are often transient, and may occur in massive macromolecular complexes, between macromolecules, or between macromolecules and ligands. A diverse set of emerging and advancing technologies are likely to spur future developments. These include advances in methods that enable individual molecules to be studied at atomic resolution; high throughput methods, increasing automation, development of massive databases that allow comparison and analysis of data of many types gathered worldwide; and increasingly powerful computational methods that enable ever larger systems to be modeled at high

N.M. Allewell (✉)
Department of Cell Biology and Molecular Genetics, University of Maryland,
College Park, MD 20852, USA
e-mail: allewell@umd.edu

I.A. Kaltashov
Department of Chemistry, University of Massachusetts-Amherst,
710 North Pleasant Street, Amherst, MA 01003, USA

L.O. Narhi
Research and Development, Amgen, Inc., One Amgen Center Drive,
Thousand Oaks, CA 91320, USA
e-mail: lnarhi@amgen.com

I. Rayment
Department of Biochemistry, University of Wisconsin-Madison,
433 Babcock Drive, Madison, WI 53706-1544, USA

N.M. Allewell et al. (eds.), *Molecular Biophysics for the Life Sciences*,
Biophysics for the Life Sciences 6, DOI 10.1007/978-1-4614-8548-3_12,

resolution. Finally, the emerging field of synthetic biology will create exciting opportunities to create, explore, and manipulate the biophysics of novel systems.

Keywords Advances in computation • Database development • High throughput automation • Macromolecular interactions • Mass spectrometry • Membrane proteins • Single molecule methods • Structural biology • Therapeutic drug development • X-ray and neutron scattering

Molecular biophysics is a dynamic, evolving area of science that continues to undergo rapid change in terms of the kinds of questions that can be asked, and the experimental and computational tools that are available to address them. This chapter presents perspectives on current challenges and future prospects in three major areas that, in combination, provide a snapshot of where the field is now and where it is moving. The first section, on X-ray and neutron scattering, emphasizes the importance of biological questions in driving advances in these technologies. Many of these biological questions focus on interactions, often transient, in massive macromolecular complexes, between macromolecules, or between macromolecules and their ligands. These themes are amplified in the second section, which describes the explosive development of mass spectrometry as a powerful tool for characterizing the conformation and dynamics of membrane proteins, large macromolecular assemblies, highly heterogeneous proteins, and molecular interactions *in vivo*. Although the biological themes in the two sections are similar, their juxtaposition reveals the complementarity of these two major experimental approaches and the insights that they provide. The last section, on the use of biophysical methods in therapeutic protein development, illustrates another important trend, the rapidly increasing importance of molecular biophysics in solving real-world problems. These applications have also driven development of the technology, in this case towards small volume, high throughput methods. In all of these examples, the questions being asked have resulted in advances in the technology, which have led to increased understanding, and consequently the ability to address even more complicated questions. In this way, molecular biophysics and the fields to which it is applied constantly interact to advance together.

12.1 X-Ray and Neutron Scattering

X-ray diffraction and neutron scattering are relatively mature methods and thus their future prospects will be driven primarily by the biological questions that must be envisaged or answered. Scattering methods are also well advanced; however, considerable technical development can be anticipated in the foreseeable future, particularly in improved methods and facilities for data collection and advances in sample preparation. Together, it can be anticipated that problems that appear almost insurmountable at present will become routine. The most compelling change will be increasing

use of scattering methods by newcomers who have not previously used these methods, as a result of more widespread understanding of the fundamentals and consequent development of automated structural determination. These prospects are outlined for crystallography, fiber diffraction, and small angle scattering, with challenges that lie at the forefront of scattering and diffraction methods described first.

12.1.1 Challenges at the Frontier of Structural Biology

For the most part, these challenges are conceptually well established, many are under active investigation and great progress is expected in the near future. Areas where there are major opportunities to enhance understanding of cellular function include the architecture of microtubule organizing centers, kinetochores, nuclear pore complexes, multiprotein membrane complexes found at the interfaces between cells, and spliceosomes. Inherent in all of these research areas are interactions between macromolecules and ligands. Interactions between the components of a cell will remain the focus of structural biology for many years to come and represent a real change in what is expected from a structural investigation. In the early days of X-ray crystallography, it was sufficient to determine the structures of the components. Initially every structure of a protein or nucleic acid was considered a major advance with little regard to the ligands or macromolecular interactions involved. However, every protein, nucleic acid, oligosaccharide, and small molecule ligand interacts with something else in the cellular system. Consequently, structure determination today goes hand in hand with biochemical and cellular studies that examine the hypotheses that arise from the structures themselves. This is because the focus has moved away from methodological development back towards understanding biological phenomena. This progress has been accompanied by an increase in the size and complexity of the biological systems that can be investigated.

The traditional approach in macromolecular structure has been to divide the problem into the smallest pieces that yield useful information and are amenable to study and then to construct a conceptual model of the original larger entity from the pieces. As techniques for determining the structures of large complexes have improved the size of the structures that can be studied has steadily increased so that less division is required. This trend is likely to continue. The challenge with complexes such as the nuclear pore complex [1] or kinetochores [2] is to isolate stable subassemblies that will crystallize in a form that yields useful structural information. Great progress has been made, but larger complexes that will ultimately lead to a complete model are still required. The next frontier in many areas will be to define the transitory interactions between molecules.

Most of the structures of complexes that have been determined thus far represent stable complexes (dissociation constants in the low micromolar or nanomolar range), but many interactions in biology are much weaker, transitory, or modulated by posttranslational modification or small molecule ligands. This is

particularly true when interactions involve an ensemble of weaker interactions such as those associated with the cytoskeleton. These studies will necessarily interface with results from electron microscopy that can provide a big picture of a macromolecular assembly. Enhanced use of molecular modeling will eventually become a vital tool in these studies, because, even with large complexes, interactions between a comparatively small number of side chains or functional groups can profoundly influence the behavior of a biological system. Most interactions in biology are dominated by hydrogen bonds and hydrophobic interactions. Hence high resolution structures of components will continue to be essential, but these will have to be incorporated into a larger model.

12.1.2 Macromolecular Crystallography: X-Rays

Conventional structural determination will almost certainly become increasingly routine. The major developments in this area will be dominated by robotic protein preparation, crystal growth and handling, automated data collection, and structural determination. This approach has been pioneered by the efforts in structural genomics, but is rapidly becoming the standard mode of operation for data recorded with synchrotron radiation. These techniques allow non-expert users to incorporate X-ray structural studies in their research protocols.

The most challenging technical problems in X-ray crystallography lie with massive macromolecular complexes, transitory interactions between molecules, and problems that yield vanishingly small crystals. At the frontiers of difficult structures, considerable advances are expected, driven by current developments in detector technology (pixel array detectors) coupled with the ability to record high quality data from exceedingly small crystals (1–5 μm). As a consequence, considerably less material is needed for a complex structural study than was once deemed necessary (micrograms to milligrams). Increasing emphasis will be placed on determining the structures of large macromolecular complexes, recognizing that protein:protein and protein:nucleic acid interactions dominate much of cell biology. The crystals of most of these complexes will not diffract to high resolution and will thus require new approaches to determining low resolution structures. Development of suitable metrics for assessing the reliable information content of these structures will be critical for the outside reader or user of this structural data.

Another area that will see rapid growth is the crystallographic study of integral membrane proteins. These have lagged behind soluble cytosolic and extracellular proteins because they are difficult to prepare and crystallize. Even when they do crystallize, most crystals of membrane proteins do not diffract well. The improvements in detector and crystallization technology are expected to have a profound impact in this area. Structural studies of membrane proteins will also be strongly influenced by developments in low resolution structural determination.

Synchrotron radiation has revolutionized X-ray structural determination, but even though high resolution data can now be recorded in a few minutes with pixel

array detectors, radiation damage is still a major problem. Recent developments in free-electron lasers that deliver ultrashort flashes in the femtosecond range of radiation have the potential to overcome this challenge [3]. With the use of ultrashort pulses of X-rays, the data can be recorded before the crystal has a chance to disintegrate or suffer radiation damage. This technology also creates an opportunity to examine even smaller crystals (less than 1 μm) and will facilitate study of macromolecules that are difficult to crystallize. This is a highly challenging approach since it requires combining scattering data from millions of diffraction experiments; however, improvements in automated data collection and sample handling are expected to simplify this approach for important structural problems.

12.1.3 Small Angle Scattering: X-Rays

As is the case for macromolecular crystallography, the results from small angle scattering will increasingly be utilized by investigators who are not experts in the field and thus will require improvements in automation and validation to ensure high quality routine data collection and appropriate interpretation of the results [4]. A large part of the effort to increase the use of small angle scattering will be associated with continued development of algorithms needed for *ab initio* modeling of the scattering data and interfacing the results with those derived from other biophysical techniques such as NMR and crystallography. In-house facilities have shown dramatic improvements in recent years and have significantly increased the number of users, however, synchrotron radiation will continue to play a major role because of the improved signal to noise and speed of data collection. A standard set of validation tools and protocols for depositing and reporting the results from small angle scattering studies will be needed to optimize the investment in this technique.

12.1.4 Neutron Scattering Methods

Neutrons provide an enormously powerful alternative to X-rays because of the greater scattering power of hydrogen and deuterium relative to other elements in biological molecules. The high scattering power of these elements makes possible contrast variation in scattering studies and the localization of hydrogen atoms in X-ray structures. In contrast, hydrogen atoms are not observed in X-ray studies except at ultra-high resolution. The only restriction on the routine usage of neutron scattering is limited access to neutron sources and the length of time required for adequate data collection. Nuclear reactors have been the mainstay of neutron sources throughout the world, but more recently spallation sources have been coming online. These accelerator-driven sources provide beams of pulsed neutrons that are considerably more intense than those available from other sources. It can be expected that these sources will encourage greater use of neutrons in the biophysical studies of macromolecules.

12.2 Mass Spectrometry

Although one of the "youngest" analytical techniques in the experimental arsenal of biophysics, mass spectrometry (MS) has already established itself as an indispensable tool, providing answers to challenging problems that cannot be addressed using other approaches. The list of targets suitable for MS analysis continues to expand, with many applications that seemed ground-breaking only a few years ago now becoming routine and commonplace. As the entire field of biophysics continues to advance, MS is expanding its scope of inquiry to include such challenging targets as membrane proteins, large macromolecular assemblies, and many others. For MS, as for other areas of biophysics, the greatest challenge is to break away from the reductionist paradigm and embrace the complexity of living systems.

12.2.1 Characterization of Conformation and Dynamics of Membrane Proteins by MS

Although membrane proteins constitute about one-third of the entire proteome, the three-dimensional structures of only 357 unique membrane proteins were available as of September, 2012. The architecture and dynamics of membrane proteins are defined by a wide range of intermolecular forces, including interactions with the hydrophobic interior of the membrane, its polar solvent interface region, as well as internal and external water molecules. As a result, membrane proteins generally have very poor solubility characteristics, making any experimental study of the architecture, dynamics, and interactions of these macromolecules extremely difficult. Traditionally, solubilization and isolation of membrane proteins relied on detergents, but many earlier attempts to characterize detergent-solubilized membrane proteins by MS had very little success because of the suppressive effect of detergents [5]. While various techniques that remove detergents prior to MS analysis remain the most popular strategy for dealing with this problem, such a dramatic change in the environment of the protein inevitably leads to the loss of higher order structure. Fortunately, small amounts of detergents can be tolerated by MS at least in some cases, allowing direct ESI MS analyses of non-covalently bound membrane protein assemblies to be carried out after reconstituting them in a minimum amount of detergent [6]. A similar approach was used recently to study very large non-covalent assemblies of transmembrane proteins [7].

Despite initial successes in using detergents for direct MS characterization of membrane proteins, one must be aware of some potential pitfalls, the most serious of which is the denaturing action of many (if not all) detergents. An ideal membrane mimetic would not only form a bilayer-based structure, but also reflect the physical properties of the specific biological membrane. Several MS-based experimental approaches are currently under investigation as potential probes of the structure and behavior of membrane proteins with bilayer-based membrane mimics. These

include limited proteolysis to identify membrane-bound protein segments, chemical probes to obtain topological information on various protein segments, and hydrogen–deuterium exchange to provide information on interfacial positioning and stability of transmembrane polypeptides in lipid bilayers. Another recently introduced bilayer-based membrane-mimicking system is a nanodisc where the bilayer structures are maintained by membrane scaffold proteins modeled after apolipoprotein A1. However, the best environment to study the behavior of membrane proteins is indisputably the specific biological membrane itself. Although characterization of various properties of membrane-bound proteins within the context of their native environment using MS was a technical impossibility until very recently, several examples of such studies have been published in the past few years [8, 9].

12.2.2 Mass Spectrometry Above 1 MDa

12.2.2.1 Characterization of Large Macromolecular Assemblies

Large protein assemblies play crucial roles in a variety of cellular functions. For example, each cellular protein emerges from a large assembly upon its birth (ribosome), enters another large assembly at the end of its life (proteasome), and interacts with a number of other macromolecular assemblies throughout its lifetime. While the ability of electrospray ionization mass spectrometry (ESI MS) to detect and characterize relatively modest non-covalent assemblies of proteins and other biopolymers (e.g., protein/DNA complexes) was recognized over 20 years ago and has been used actively since, large macromolecular assemblies representing complete self-contained units of biological machinery (such as ribosomes and proteasomes) remained out of reach of MS analysis for a much longer period of time.

The situation began to change in the past decade as a result of pioneering work of Robinson [10] and Heck [11], who demonstrated that careful control of ionization conditions and use of mass analyzers with extended m/z range may allow very large non-covalent complexes to be preserved in the gas phase, and meaningful structural information to be extracted for protein assemblies whose masses exceed several MDa. Although still far from being a routine method of analysis of large macromolecular complexes, the so-called native mass spectrometry is now capable of dealing with complex objects ranging from proteasomes and ribosomes to intact viral capsids.

12.2.2.2 MS of Highly Heterogeneous Proteins

Despite the dramatic expansion of the mass limit of macromolecules for which meaningful information can be provided by MS, the bar remained disappointingly low until recently for MS analysis of several classes of proteins. These include extensively glycosylated proteins and protein–polymer conjugates, which

frequently exhibit remarkable degrees of structural heterogeneity. Heterogeneity poses a formidable challenge to MS-based studies of higher order structure, dynamics, and interactions of such proteins, frequently making the mundane task of mass measurement an extremely challenging undertaking. Among several recent developments in this field, a particularly promising approach combines reduction of complexity of the protein ion ensemble (by mass selecting a narrow fraction of the entire ionic population) and gas phase chemistry (charge reduction via electron capture or electron transfer) [12].

Another MS technique that holds great promise vis-à-vis dealing with macromolecular complexity is ion mobility (IM) MS [13]. While the majority of current applications of this technique exploit its ability to provide information on the physical size of macromolecular ions in the gas phase, the potential utility of IM MS to provide an additional separation stage prior to MS detection, thereby reducing complexity of heterogeneous systems, is frequently overlooked. Nevertheless, the ability of IM MS to separate various isoforms of biopolymers has been acknowledged and has already been used to facilitate MS characterization of covalent structure of large glycoproteins [14] and protein–polymer conjugates [15].

12.2.2.3 Mass Spectrometry *In Vivo*

A very important aspect of macromolecular interactions *in vivo* is their extreme complexity due to the large number of participating players. While most biophysical studies have traditionally used the so-called reductionist approach by focusing attention only on the minimal number of players deemed absolutely essential for a particular process or interaction, the limitations of this approach are now becoming commonly acknowledged. Emergence of the new paradigm that embraces, rather than downplays, the complexity of biological processes has been catalyzed by the completion of genome sequencing for several organisms, which highlighted the enormous repertoire of biomolecules making up living cells.

One approach to dealing with the complexity of real-living systems that enjoyed great popularity in the past decade, is functional proteomics [16–22]. Above and beyond proteomic approaches that provide a global picture of biomolecular interactions in living systems, a number of groups are beginning to invest significant effort in expanding the existing experimental strategies to study biomolecules in their native environment. These include the possibilities for investigation of protein structure and interactions in living cells provided by chemical cross-linking with MS detection [23], or chemical labeling and footprinting methods [24]. Efficient delivery of cross-linking and/or labeling reagents to the cell without disrupting its normal functioning or indeed killing it remains a formidable challenge. This obstacle places a significant limitation on the number of reagents that can be used in such *in vivo* measurements. One particularly attractive approach to overcoming this problem would tap into the arsenal of the emerging field of synthetic biology by reprogramming the genetic code of the cell, forcing it to produce and incorporate into proteins amino acids with reactive side chains that can be used as in situ chemical probes [25].

The past two decades witnessed many triumphs of MS in various subfields of biophysics and structural biology, and it is certain that this technique will remain a valuable contributor in these fields, catalyzing their progress in the years to come and bringing about new exciting discoveries. Despite reaching a respectable age, biological MS remains very dynamic and constantly adapts to the ever changing landscape in the life sciences, always remaining at the forefront and ready to deal with the most challenging problems.

12.3 The Future of Biophysical Analysis in Therapeutic Protein Development

The development of therapeutic proteins is an endeavor that includes extensive use of biophysical techniques, as has been described in numerous publications (see, for example [26]). As is the case for all proteins, proteins being developed and used for therapeutics are complex macromolecules that require appropriate primary, secondary, and tertiary structure to maintain their function and stability, as discussed in Chap. 2. The ultimate goal of therapeutic development is the creation of a molecule that is safe and efficacious, and that will maintain its structural integrity during manufacturing, storage (usually two years, often in solution, and under variable conditions), and administration. Different biophysical tools are employed during the different stages of development of this important class of drugs, depending on the amount of material and time available, and the goal of the analysis.

The protein therapeutic development lifecycle includes several steps, beginning with the identification of a biological target. After the target has been chosen, the molecule with the greatest chance of succeeding as a drug and with the desired biological activity must be selected from multiple candidates with different primary sequences. Following the choice of candidate, process and formulation development, and characterization are the next steps, with selection of delivery device and route of administration coming next. During all of these steps, the integrity of the protein, in terms of its secondary, tertiary, and quaternary structure needs to be maintained. The last step in this process is clinical trials and then, if successful, commercialization. During these later stages of development, the focus is on product consistency and lot release assays, exploration of different delivery devices and therapeutic indications, comparability assessments, and support for product and process failure investigations.

Biophysical tools are used at all of these stages in the therapeutic protein lifecycle. Currently characterization is done by removing an aliquot of the sample and analyzing specific properties with different techniques, and then using heuristics to combine the results. This can be time-consuming and involves multiple aliquots and a fair amount of material. The desired future state for biophysical assessment of protein therapeutics would include the ability to do multiple analyses on the same sample at all stages of development. This would increase the reliability of the results because different attributes could be directly compared, and also enable the testing

Table 12.1 The goals and challenges of the "biophysics of the future of therapeutic protein development"

Goal/challenge	Current	Desirable
Analysis at high/low concentration	Many biophysical techniques require dilution to 0.5–1 mg/mL range, others require concentration to >10 mg/mL	At actual formulation concentration
High throughput	Many techniques are labor-intensive and low throughput	Automation, high throughput data collection and analysis
Noninvasive	Requires removing sample from device (vial or syringe)	In situ analysis
Online biophysics for process control	Discrete sampling, off-line and often time-consuming analysis	Online sampling and analysis during fermentation and purification, the ability to make changes based on result obtained on the fly

of more samples to better understand the variability of the methods. The vision of the future for the application of biophysics during protein development also includes being able to perform these tests on actual process samples and to obtain results in real-time so that decisions can be made based on the identity, conformation of the protein, and the state of aggregation. Automation and/or use of easy-to-operate instrumentation for these techniques in a manufacturing environment is another important goal for which to strive. Different phases of drug product development have different specific needs as well that could result in the evolution of different instrumentation and applications in the future. This issue is discussed below and briefly summarized in Table 12.1.

During the selection of the unique protein that will become the product candidate, in addition to biological activity, the stability of the candidates under consideration to the conditions used for manufacturing and storage is assessed. Characteristics to be considered include stability to low pH, agitation, mixing of the air–liquid interface, and temperature. The protein therapeutic also needs to withstand storage in solution at 4–8 °C for two years, often at protein concentrations above 100 mg/mL [27]. Screening for this type of stability usually involves predictive assays that rely on subjecting the protein to harsher conditions than it would encounter normally, in order to predict what may happen with time under milder conditions. This requires an understanding of potential pathways of degradation, in order to ensure that the response of the protein to the conditions used are truly predictive of long-term stability during the actual process. After stressing the material, the impact of the conditions on the integrity of the protein, with particular emphasis on protein aggregation and irreversible unfolding of the native three dimensional structures, is assessed. Assays with minimal material requirements and high throughput are especially valuable at these early stages. Qualitative results that allow comparison of the relative degree of change so that candidates can be categorized as passing or failing are an acceptable output.

The ability to assess multiple different protein characteristics on a single sample after each stress, rather than having to remove aliquots followed by sample manipulation in order to be compatible with the different analyses, would be hugely

valuable at this stage. Primary attributes that should be assessed are conformational and colloidal stability, the propensity of the protein to aggregate, and chemical modification of the amino acid residues in the primary structure. The ability to measure, either directly or indirectly, the primary, secondary, tertiary, and quaternary structure of a protein, and the size of any aggregated species generated, all in a high throughput format, is the ultimate goal. The more candidates that can be assessed, the greater the chance of identifying one that has the desired properties, and so the availability of automated methods is also important. In the future one can envision a robotic system that subjects samples to different stresses such as elevated temperature, extremes of pH and ionic strength, mechanical shaking or stirring, exposure to light, etc., and then runs the multi-well plates through sequential biophysical analyses, ultimately providing a relative ranking of the candidates based on a multivariate analysis of the matrix of data generated. It is worth noting that in some cases in addition to the traditional biophysical techniques (such as MS, different types of spectroscopy) other methods, such as chromatography, can often be used as part of this assessment. For example, ion exchange chromatography can detect changes in chemical modification, hydrophobic interaction chromatography can be used to detect changes in conformation, and size exclusion chromatography can follow loss of monomer, or formation of smaller oligomers such as dimers and tetramers [28].

An important aspect of this early stage of development is the feedback between protein engineering, modeling, and the results of the predictive assays. There is an iterative process as the correlation between the predicted behavior, the actual behavior as the protein moves through process development, and the structure of the modeled protein becomes available. Collecting these data into usable databases allows constant improvement in the sequence-based predictive algorithms, such that more and more of the potential "hot spots" for modification or self-association can be eliminated before the protein is ever included in the panel of potential candidates to be screened.

The focus changes to developing the production process and formulation to be used for the commercial product once the specific molecule that will be developed as a therapeutic has been chosen. At this stage, material availability is no longer rate limiting, and more rigorous techniques that compare the higher order structure of the actual material obtained during the different processing steps can be used to ensure that the final product was not irreversibly damaged by the conditions being used for its manufacture. The ability to get real-time, high resolution information on the primary, secondary, and tertiary structure, and especially of the aggregation state, of samples as they are generated by the cells, and passed through the purification process, would allow decisions about sample collection to be made based on the quality of the material as it was being processed. This requires online instruments that are robust enough to withstand the conditions of a protein manufacturing plant, and are also rapid enough to provide results in time to be used to make process conditions. Online light scattering analysis to assess aggregation; Raman spectroscopy to assess the secondary and tertiary structure of the protein; mass spectrometry to determine primary structure including amino acid sequence, carbohydrate content, and chemical modification; and morphological analysis to assess types of aggregate are some of the potential process analytical technologies that are currently being explored.

During formulation development, the stability of the target protein is assessed in different buffer compositions, pH, storage conditions, and delivery devices. These studies typically involve the generation of many samples that must be analyzed in order to arrive at the optimal formulation conditions, and thus many of the principles that apply during candidate selection apply here as well and some of the same assays can be used. The primary difference is that at this point in the development lifecycle the amount of material is no longer rate limiting and so formats other than the 96 (or more) well plates can be considered. However, an instrument that uses robotics to stress and test multiple samples for several attributes simultaneously is still the goal. Special attention should be paid to the aggregation state and the integrity of the primary sequence of the protein. Ideally the analyses would occur under the actual solution and storage conditions that would be used, including protein concentration. The majority of the protein therapeutics under development will be administered at high concentration, and so the ability to determine these properties without dilution is an important consideration.

As the product moves into production, the emphasis switches from developing and optimizing conditions to maintaining process/product control; biophysical techniques to follow the protein higher order structure are important elements of comparability studies, and are required for obtaining licensure of the drug. In this case the methods must be shown to be fit for the purpose and the sensitivity of the assays to detect changes in the product must be determined. Another important aspect of preclinical and clinical development is the monitoring of stability samples, stored both under accelerated and recommended conditions, for comparability. Biophysical techniques are also used as tools to help ensure that changes in device, concentration, and formulation made as different indications or patient populations are added do not affect the conformation of the biotherapeutic. Techniques that can give reproducible and accurate results, and where the readout is understood, are most commonly used at this stage of development, rather than the high throughput tests that were employed in the beginning of the product development lifecycle. These analyses must be sensitive to changes in the protein conformation that can occur if the protein is exposed to slightly different process or storage conditions, as demonstrated by samples exposed to conditions outside the normal parameters. At this stage future directions lie in the ability to carry out multiple biophysical tests on the same sample in the commercial formulation, removed directly from the commercial delivery device. This capability would allow for testing of a statistically relevant number of samples, and direct comparison of the results. One difficulty with the current tools available for biophysical characterization of proteins is that most of these lack the sensitivity necessary to detect changes of less than 5 %. They also provide information on the average of the molecular population. Thus even when a difference is detected, it is not possible to determine whether 6 % of the population has lost all signals in that assay, or 100 % of the population has lost 6 % of the signal. Evolution of single molecule methods to the point where they are applicable for quick and reproducible analyses with very little variability would be a huge step forward in our ability to interrogate samples and truly understand if they are comparable or not. The ability to apply high resolution methods such as NMR and MS

to gain better understanding of protein higher order structure down to a single residue is being explored as one avenue to obtain this type of information. The ability to track and characterize minor species in the structural ensemble that is present at any given time in a protein solution would also contribute to this. Finally, proteins are not static species, but are truly dynamic and can sample multiple folding structures as part of the natural thermodynamic equilibrium of the states possible in solution. Techniques to provide quick assessments of the dynamics of any given protein solution would also be helpful for this stage of development.

Finally, during commercial production batches occasionally fail the various lot release assays; biophysical techniques can be used to help identify the root cause and contribute to the safety assessment of the different lots of protein produced under supposedly equivalent conditions. In this case very often a single sample is being tested, with a single visible aggregate being studied, and so the methods must have the sensitivity to detect and analyze a very small amount of protein and provide a positive identification of the material if possible. For this application ideal future biophysical tools would include analysis by mass spectrometry for molecule identification and determination of any chemical modification, as well as analysis of the conformation of the protein, and the aggregation state. This analysis should occur *in situ* in the glass vial, syringe, or other device used to administer the drug to the patient, and all the analyses must be performed on the same particle or other species that resulted in the lot release failure. While throughput is important, the ability to obtain very reliable results from such a small sample set is far more important than throughput at this stage.

As illustrated in Table 12.1, and from the discussion above, there are many gaps between the current state of protein biophysical characterization during biotherapeutic development and the desired future state. While challenging, much progress has been made in recent years. The evolution of computational and material sciences is resulting in miniaturization of instrumentation to the point where the "lab on a microchip" will become feasible. The development of high throughput, automated, instruments that assess more than one attribute on these chips, coupled with sophisticated statistical calculations and multivariate analysis of the information does not seem outside the realm of possibility in the relatively near future.

12.4 Conclusions

This chapter has provided perspectives on future directions in three major areas of molecular biophysics, as examples of what the future holds. While many advances will be specific to a particular field, there are a several overarching themes that are common to the three topics discussed here as well as many other areas of biology. Throughout biophysics, the focus is moving from isolated components to entire systems. At the same time single molecule analyses will increasingly enable us to visualize and characterize minority species against the background of the entire molecular population, including transitory states. The evolution of high throughput

methods will result in an increase in throughput and a decrease in the amount of material required. Increasing use of automation will make biophysical approaches accessible to a wider group of users, and application to a larger variety of systems. Massive databases will allow comparison of results across different samples, systems, and even laboratories, while increasingly powerful computational approaches will enable large systems to be modeled at high resolution. Finally, the emerging field of synthetic biology will enable biophysics to extend beyond natural systems to novel synthetic systems.

References

1. Bilokapic S, Schwartz TU (2012) 3D ultrastructure of the nuclear pore complex. Curr Opin Cell Biol 24:86–91
2. Corbett KD, Harrison SC (2012) Molecular architecture of the yeast monopolin complex. Cell Rep 1:583–589
3. Boutet S, Lomb L, Williams GJ, Barends TR, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, Kassemeyer S, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, van Bakel N, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang M, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang D, Zatsepin NA, Schafer D, Defever J, Neutze R, Fromme P, Spence JC, Chapman HN, Schlichting I (2012) High-resolution protein structure determination by serial femtosecond crystallography. Science 337:362–364
4. Petoukhov MV, Svergun DI (2013) Applications of small-angle X-ray scattering to biomacromolecular solutions. Int J Biochem Cell Biol 45(2):429–437
5. Annesley TM (2003) Ion suppression in mass spectrometry. Clin Chem 49:1041–1044
6. Lengqvist J, Svensson R, Evergren E, Morgenstern R, Griffiths WJ (2004) Observation of an intact non-covalent homotrimer of detergent-solubilised rat microsomal glutathione transferase 1 by electrospray mass spectrometry. J Biol Chem 279(14):13311–13316, M310958200
7. Barrera NP, Di Bartolo N, Booth PJ, Robinson CV (2008) Micelles protect membrane complexes from solution to vacuum. Science 321:243–246
8. Pan Y, Stocks BB, Brown L, Konermann L (2009) Structural characterization of an integral membrane protein in its natural lipid environment by oxidative methionine labeling and mass spectrometry. Anal Chem 81:28–35
9. Wen JZ, Zhang H, Gross ML, Blankenship RE (2009) Membrane orientation of the FMO antenna protein from Chlorobaculum tepidum as determined by mass spectrometry-based footprinting. Proc Natl Acad Sci U S A 106:6134–6139
10. Sobott F, McCammon MG, Hernandez H, Robinson CV (2005) The flight of macromolecular complexes in a mass spectrometer. Philos Trans A Math Phys Eng Sci 363:379–389, discussion 389–391
11. Heck AJR (2008) Native mass spectrometry: a bridge between interactomics and structural biology. Nat Methods 5:927–933
12. Abzalimov RR, Kaltashov IA (2010) Electrospray ionization mass spectrometry of highly heterogeneous protein systems: protein ion charge state assignment via incomplete charge reduction. Anal Chem 82:7523–7526
13. Bohrer BC, Mererbloom SI, Koeniger SL, Hilderbrand AE, Clemmer DE (2008) Biomolecule analysis by ion mobility spectrometry. Annu Rev Anal Chem 1:293–327

14. Damen C, Chen W, Chakraborty A, van Oosterhout M, Mazzeo J, Gebler J, Schellens J, Rosing H, Beijnen J (2009) Electrospray ionization quadrupole ion-mobility time-of-flight mass spectrometry as a tool to distinguish the lot-to-lot heterogeneity in N-glycosylation profile of the therapeutic monoclonal antibody trastuzumab. J Am Soc Mass Spectrom 20:2021–2033
15. Bagal D, Zhang H, Schnier PD (2008) Gas-phase proton-transfer chemistry coupled with TOF mass spectrometry and ion mobility-MS for the facile analysis of poly(ethylene glycols) and PEGylated polypeptide conjugates. Anal Chem 80:2408–2418
16. Collins MO, Choudhary JS (2008) Mapping multiprotein complexes by affinity purification and mass spectrometry. Curr Opin Biotechnol 19:324–330
17. Monti M, Cozzolino M, Cozzolino F, Vitiello G, Tedesco R, Flagiello A, Pucci P (2009) Puzzle of protein complexes in vivo: a present and future challenge for functional proteomics. Expert Rev Proteomics 6:159–169
18. Terentiev AA, Moldogazieva NT, Shaitan KV (2009) Dynamic proteomics in modeling of the living cell. Protein–protein interactions. Biochemistry (Mosc) 74:1586–1607
19. Malik R, Dulla K, Nigg EA, Korner R (2010) From proteome lists to biological impact-tools and strategies for the analysis of large MS data sets. Proteomics 10:1270–1283
20. Zhou M, Robinson CV (2010) When proteomics meets structural biology. Trends Biochem Sci 35:522–529
21. Gavin AC, Maeda K, Kuhner S (2011) Recent advances in charting protein–protein interaction: mass spectrometry-based approaches. Curr Opin Biotechnol 22:42–49
22. Sardiu ME, Washburn MP (2011) Building protein–protein interaction networks with proteomics and informatics tools. J Biol Chem 286:23645–23651
23. Sinz A (2010) Investigation of protein–protein interactions in living cells by chemical crosslinking and mass spectrometry. Anal Bioanal Chem 397:3433–3440
24. Zhu Y, Guo TN, Park JE, Li X, Meng W, Datta A, Bern M, Lim SK, Sze SK (2009) Elucidating in vivo structural dynamics in integral membrane protein by hydroxyl radical footprinting. Mol Cell Proteomics 8:1999–2010
25. Xie J, Schultz PG (2006) A chemical toolkit for proteins—an expanded genetic code. Nat Rev Mol Cell Biol 7:775–782
26. Narhi L (2012) In: Narhi L (ed) (2013) Biophysical characterization during protein therapeutic development. Springer
27. Mahler H-C, Friess W, Grauschopf U, Kiese S (2009) Protein aggregation: pathways, induction factors and analysis. J Pharm Sci 98:2909–2934
28. Chen S, Lau H, Brodsky Y, Kleemann GR, Latypov RF (2010) The use of native cation-exchange chromatography to study aggregation and phase separation of monoclonal antibodies. Protein Sci 19:1191–1204