CHAPTER 1

Overview of Mass Spectrometry Technologies for Examining Protein Structure: Current and Future Directions

SHANNON M. SWIATKOWSKI and MARK R. CHANCE

Center for Proteomics, Case Western Reserve University, Cleveland, OH, USA

1.1 INTRODUCTION

Understanding the molecular structure and dynamics of macromolecules at high resolution and with high throughput is a topic of great importance in biology. Nuclear magnetic resonance (NMR) and crystallographic approaches are the foundation of rapid progress in this area. Access to genome sequences and cloning resources from an ever-increasing number of organisms and allied high-throughput structure and modeling studies are likely to enable resolution of the structure of most protein domains in the near future (Chance et al., 2004). However, the machinery of eukaryotic cell biology involves multidomain proteins that interact in large complexes as molecular machines (Sali et al., 2003; Russell et al., 2004). Understanding how these domains interact is crucial in understanding their function. As this "database" of structural information evolves and develops, examination of the structure–function relationships of a wide range of proteins becomes possible. In addition, many biological questions of interest invoke questions of protein dynamics, ligand binding, complex formation, or the structural effects of posttranslational modifications. Many of these experiments are beyond the range of classical structural biology approaches (see below) and structural mass spectrometry (MS) methods have been very successful in filling this technological gap. The fundamental contributions of mass spectrometry to structural biology studies have grown dramatically due to increases in instrument sensitivity and resolution that have accrued over the past 10 years. This has advanced our ability to reliably sequence and identify protein

Mass Spectrometry Analysis for Protein–Protein Interactions and Dynamics, Edited by Mark Chance

fragments and their modified products, a feature on which structural mass spectrometry fundamentally relies. This book catalogs the state of the art in these approaches and provides a perspective on the future prospects for the field. The three main technologies of structural mass spectrometry that have rapidly evolved and grown, include covalent labeling strategies, hydrogen–deuterium (H/D) exchange, and chemical cross-linking.

Although the technologies have a great many differences in their sample preparation, instrumentation requirements, and other details of the approaches, their similarities must not be overlooked. First, they all rely on detailed identification and sequencing of peptide fragments generated by specific or nonspecific cleavage of intact and (generally) purified protein species (or complexes). Second, they infer structural information based on a mass shift of these peptide species after exposure to the labeling reagents of choice. The target atoms that are labeled must be solvent accessible, at least transiently. Third, the value of the structural information is greatly enhanced by having a structural model of the protein or proteins. It is, in fact, very clear that the advancement of these approaches will be significantly accelerated by a union of these experimental technologies with computational modeling approaches in the context of the rapidly expanding structure databases (Chance et al., 1997; Guan et al., 2004; Kamal and Chance, 2008; Takamoto and Chance, 2006).

Structural models for most protein domains, providing a foundation for structural mass spectrometry, are accumulating rapidly (Eswar et al., 2007). Advances in protein structure determination and computational modeling mediated by structural genomic initiatives throughout the world promise to correlate sequence and structure for most protein domains within the next 5 years (Burley et al., 1999; Chance et al., 2002, 2004). Coincident with progress toward this milestone is the realization of the importance of macromolecular interactions and even the fundamental significance of large macromolecular complexes mediating most normal and aberrant biological functions (Gavin, 2005). Solving the structure and connecting it to function for these large complexes are two of the most important challenges in structural biology today. Unlike solving the structure of protein domains or short nucleic acids that contain tertiary structure, this effort is far from high throughput and likely involves a combination of computational and experimental approaches, tailored specifically to the problem at hand.

The barriers to determining the structure and dynamics of proteins and their complexes include known limitations in crystallography and NMR technologies. Issues such as complex size, crystallizability, solubility, and amounts of materials are well known. In recent years, electron microscopy (EM) and tomography techniques, particularly at low temperatures, have substantially improved and are making important contributions to determining the structure of complexes (Sali et al., 2003). These approaches have resolution limitations for many samples and are better for larger complexes or cells due to sample dose tissues. This leaves a gap in technological progress for the "medium" size complexes, particularly medium-sized binary complexes (50–200 kDa). This has spurred the development of a host of computational methods that can fill in the gap and contribute to understanding the relationship between protein structure and function.

1.2 HYDROGEN/DEUTERIUM EXCHANGE MASS SPECTROMETRY

Deuterium exchange is a very powerful technique in the repertoire of structural mass spectrometry. H/D exchange MS methods were developed in the early 1990s, inspired by related NMR methods. The practicality of the method lies in the fact that amide hydrogens are sensitive probes for solvent accessibility, protein lability, and protein secondary structure. The H/D exchange method is shown in Fig. 1.1 (and compared with hydroxyl radical footprinting, a covalent labeling approach). The protein backbone amide hydrogens are exchangeable with deuterium atoms from the solvent surrounding the protein at different measurable exchange rates. The amide hydrogens at the surface of proteins exchange very rapidly, while amide hydrogens that are buried or are participating in stable hydrogen bonds have much slower exchange rates (Busenlehner and Armstrong, 2004). Thus, H/D exchange rates can be measured along the entire length of the protein backbone, providing a comprehensive measure of protein structure and solvent accessibility. Since backbone amide hydrogens are also involved in the formation of hydrogen bonds in protein secondary structures, their exchange rates are also a reflection of secondary structure and structural stability.

The protein of interest is subjected to a pulse of deuterium intended to label structural regions that are solvent accessible and to monitor changes in accessibility in response to the binding of a ligand (Katta and Chait, 1993; Zhang and Smith, 1993). After solvent labeling, the reaction is quenched, the protein is fragmented by proteolysis, the peptide fragments are separated by high-pressure liquid chromatography (HPLC), and mass spectrometry analysis is performed. Peptide fragments with increased mass relative to control experiments without addition of deuterium indicate *specific segments* that were solvent accessible and exchange competent during the deuterium pulse. The method used to minimize back exchange during the analytical steps is lowering the pH to $\sim$2.5; thus, only proteases with activity at acid pH (e.g., pepsin) can be used to fragment the protein.

Chapter 2 describes the fundamental concepts that govern the hydrogen exchange (HX) reaction beginning with the chemistry of hydrogen exchange and continuing through to discuss HX mechanisms in proteins and how they can be assessed with mass spectrometry. These concepts build a foundation for a discussion of basic HX MS methodology and how it can be applied to various biological problems in subsequent chapters of this book. Similarly, Chapter 12 outlines PLIMSTEX (protein–ligand interactions in solution by mass spectrometry, titration, and H/D exchange). This strategy can be used to determine the conformational change, binding stoichiometry, and affinity for a variety of protein–ligand interactions. Chapters 5–7 provide a more comprehensive look at protein complex structure through several case studies of complex formation. In Chapter 5, HX MS data are used in combination with docking, biochemical, and genetic data to better understand the biophysics of protein–protein interactions in protein kinase A (PKA) and nuclear factor kappa B (NF-κB). Hydrogen/deuterium exchange (HDX) studies of protein complex formation can also be extended to viruses as explained in Chapter 6, where HX MS data of viral capsid structure and protein dynamics studies are presented

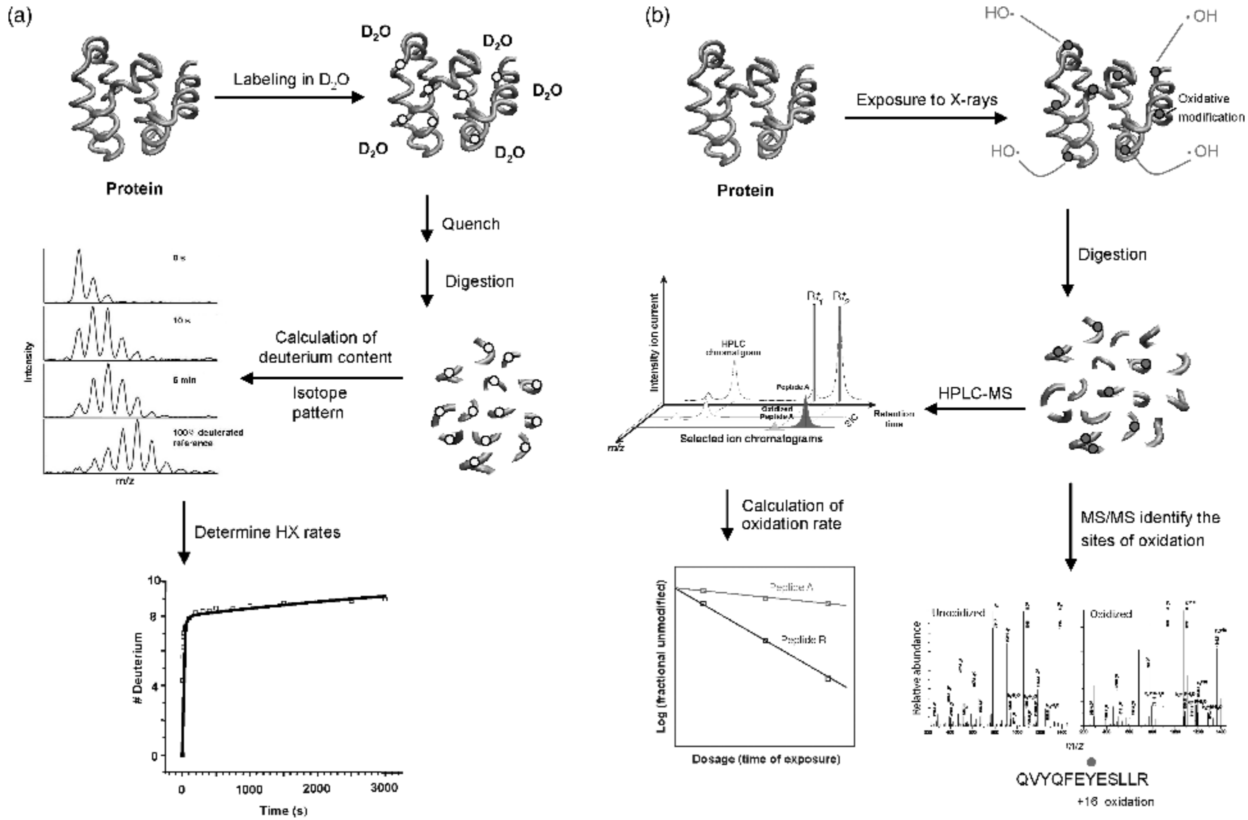

FIGURE 1.1 Schematic representations of synchrotron footprinting and H/D exchange mass spectrometry techniques. (a) By changing the solvent from H_2O to D_2O, amide protons in the backbone of protein exchange with the protons from the solvent. The exchange reaction is quenched by quickly changing solution pH to approximately 2.4 and usually combining with rapid freezing. Typically, the protein is digested with pepsin. The kinetics of amide H/D exchange is measured by mass spectrometry to provide essential dynamic information for protein. (b) When a protein is exposed to synchrotron X-rays, the hydroxyl radicals generated from water will modify side chains of the protein. After the X-ray exposure, the protein samples are digested by proteases and solvent accessibility information is provided by MS. The particular modification sites are determined by tandem MS and the side chain reactivity is accurately measured by quantitative liquid chromatography-coupled mass spectrometry. Figure reproduced, with permission from Elsevier Ltd., from Zheng et al., 2008.

along with several technical aspects of performing HX MS experiments on viruses. Chapter 7 highlights four examples of the use of enhanced peptide amide hydrogen/deuterium exchange coupled with proteolysis, liquid chromatography, and mass spectrometry (DXMS) to probe protein–protein interactions. The DXMS method can be used to complement thermodynamic information of a protein–protein interaction, monitor conformational changes associated with kinase activation, compare isoform-specific differences in binding of a common ligand, and map the epitopes of monoclonal antibodies. Chapter 4 provides a specific comparison of hydrogen/deuterium exchange and covalent labeling on the same protein, the first such direct comparison. It highlights the complementarity of the methods. Finally, Chapter 11 concludes the discussion of this method by explaining the many advantages of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) MS compared to other means of monitoring the progress of hydrogen/deuterium exchange reactions.

1.3 HYDROXYL-RADICAL-MEDIATED PROTEIN FOOTPRINTING

The development of the hydroxyl radical as a modification reagent for footprinting and its application in conjunction with mass spectrometry were directly inspired by the development of deuterium exchange mass spectrometry methods. Hydroxyl-radical-mediated protein footprinting is similar. The overall method is outlined in Fig. 1.2. In this case, the protein solution is exposed to ionizing radiation and the hydroxyl radicals covalently react with surface-accessible residues, primarily side chain groups (Maleknia et al., 2001). As in the deuterium exchange methods, the protein is subjected to proteolysis. However, in contrast to deuterium exchange, the production of stable modifications through hydroxyl radical exposure allows a wide range of samples as well as proteases to be used to fragment the protein under a wide range of solution conditions and pH values. Also, the stable modification of side chains allows a specific probe site to be identified using tandem mass spectrometry methods, while, for deuterium exchange, typically the conformational change can only be localized to the specific peptide fragment. The drawback is that if a reactive side chain is not present in a particular peptide segment, there are no probes. However, the examination of side chains is complementary to the deuterium exchange method that examines backbone structure.

To generate the limited dose required for footprinting and to quantitatively examine the reactivity of the specific peptides in question, a series of samples are exposed to variable doses, the samples are digested, and the individual peptides are analyzed by HPLC and mass spectrometry. Thus, a dose–response curve is generated for each peptide of interest; this generates a quantitative biophysical measure (based on the observed rate of modification) of the relative reactivity of the sites in the different peptides. Consistent and reliable quantitation, which is essential to footprinting, is provided by measuring the relative amounts of the modified and unmodified peptide products in the same experiment. Since the modifications are stable, it is relatively straightforward to use tandem mass

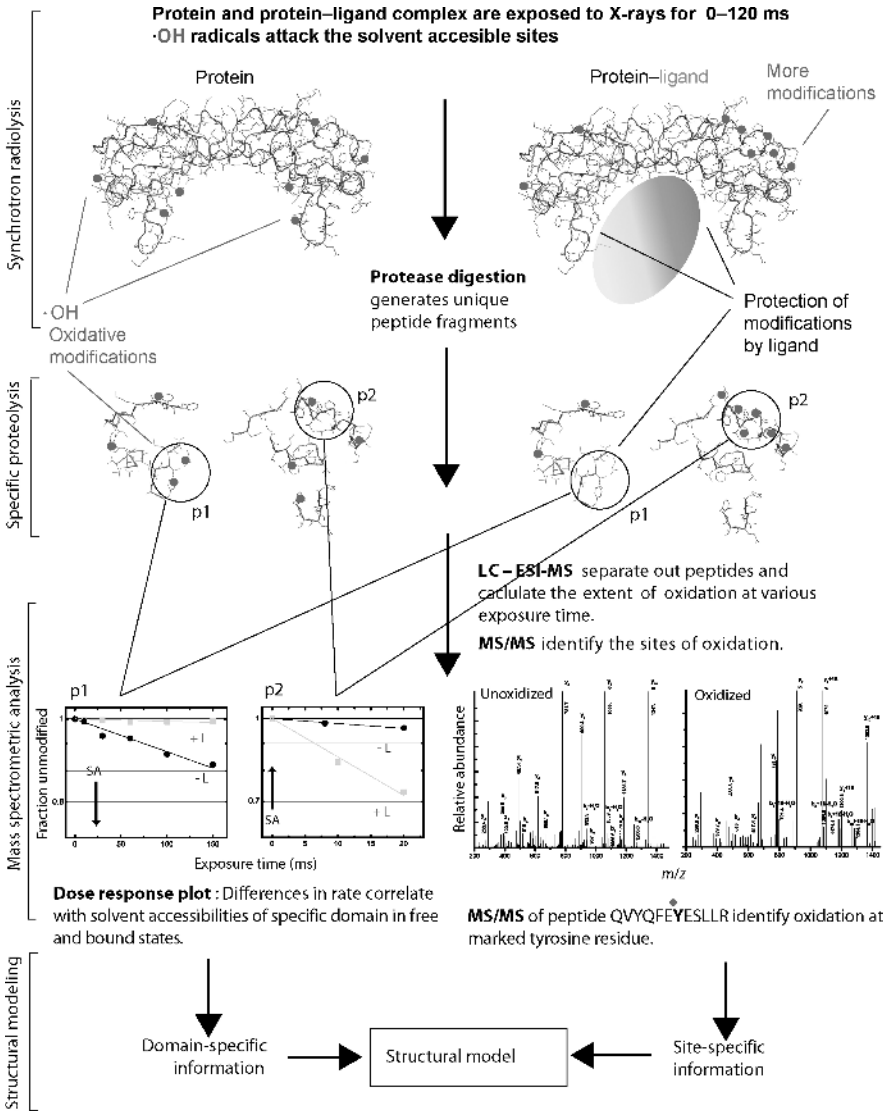

FIGURE 1.2 Schematic representation of protein footprinting using synchrotron radiolysis and mass spectrometry. The examples emphasize the protection formed in the interface of a protein– ligand complex as well as allosteric conformation changes that can result in increases in reactivity upon ligand binding; however, the comparison could be for any two (or multiple) functional states of the protein of interest. Two sets of samples, one free protein and the other a protein–ligand complex, are exposed to X-rays for different time intervals. The exposed samples are digested with specific digestion enzymes. The digested fragments are analyzed by ESI-MS to quantitate the extent of modification products and determine the fraction “unmodified” for a specific exposure time. A plot of fraction unmodified versus exposure time, known as the dose response plot, fit to a first-order function providing the rate of modification for the specific peptide. Comparisons of the dose response of the same peptide under different conditions provide structural information about ligand binding. MS/MS is used to determine the specific

spectrometric methods to specifically identify the amino acid positions of the protein that have been modified; these represent the probe sites for the analysis (Guan et al., 2002; Kiselar et al., 2002).

On the right-hand side of Fig. 1.2, where the reaction is carried out in the presence of a protein–protein complex, modification is suppressed at the site corresponding to the interacting surface and unchanged at sites distal to the contact. The decrease of reactivity of peptides in the interface is quantitatively measured using the dose–response curves; a sample dose–response for a peptide whose reactivity is suppressed in the presence of the ligand is shown at the bottom left of the figure. It is critical for the method that the dose–response curves indicate a linear regime extrapolated to zero fraction modified; this ensures the reactivities of particular sites are not changing due to the oxidation process itself. Also, the method of examining the loss of the unmodified fraction emphasizes the interrogation of intact material. Although these particular refinements are unique to the hydroxyl-radical-mediated protein footprinting approach, they are derived from a detailed knowledge and respect for safeguards that have evolved throughout the long history of development of "footprinting" research (Brenowitz et al., 2002; Takamoto and Chance, 2004); such safeguards ensure that the structural and biochemical information provided by footprinting methods is reliable. It must be emphasized, however, that footprinting (along with H/D exchange) provides only "local" information about the reactivity of the side chain probes. Allosteric changes in conformation induced by ligand binding can also give rise to either protections (decreases in side chain reactivity) or enhancements (increases in reactivity) depending on the induced conformational changes. This must be carefully borne in mind when interpreting data from these experiments, and various examples are mentioned in this review.

Chapter 3 reviews the chemistry of covalent labeling approaches for structural mass spectrometry and provides details of the various methods used to generate reactive species and define their reactions with proteins. Specifically discussed is the hydroxyl radical footprinting methodology and mass spectrometry approaches for quantitative footprinting. The complementarity of the hydroxyl-radical-mediated footprinting and hydrogen/deuterium exchange mass spectrometry techniques is discussed in Chapter 4. This chapter compares the structural results of the two methods on the trypsin inhibitor and also explores the ability of these methods to probe protein conformational dynamics. Chapter 10 provides a thorough discussion on the subject of using computational approaches to examine protein–protein interactions through the combination of computational and experimental data. This chapter specifically describes how a combination of radiolytic footprinting coupled to mass spectrometry analysis and docking with the ClusPro server can be used to derive a structure for the actin/cofilin binary complex. To obtain a "snapshot" of a protein that is uninterrupted by oxidation-induced protein unfolding, Chapter 12

modification site within the peptide and provide side-chain-specific structural resolution. Synchrotron protein footprinting data are often used as one of the constraints in model building for complexes. Figure reproduced, with permission from the International Union of Crystallography, from Gupta et al., 2007.

describes a fast radical footprinting technique termed fast photochemical oxidation of proteins (FPOP). This approach ensures that there is sufficient protein to be oxidized and then analyzed by employing a flow system coupled with a pulsed laser to produce OH radicals from hydrogen peroxide.

1.4 CHEMICAL CROSS-LINKING

Chemical cross-linking is also a valuable tool for examining higher-order structure of proteins and protein complexes. Many types of reagents are used in cross-linking experiments and can be used to cross-link multiple kinds of residues. Strategies to insert specific susceptible residues in sites of interest are also valuable. Cross-linking methodology is able to provide or confirm low-resolution structures when some model structure is available (Novak and Giannakopulos, 2007). Variations in the method (top-down versus bottom-up) are outlined in the book, and further potential applications of chemical cross-linking of proteins, as well as combinations with other techniques such as hydrogen/deuterium exchange and molecular modeling, are suggested.

Chapter 8 introduces the cross-linking methodology along with several strategies of the technique and how it can be used in conjunction with mass spectrometry to obtain structural information regarding the organization and function of protein complexes that may be otherwise impossible to obtain. Chapter 9 provides a specific example of how cross-linking can be used as a tool to probe actin structure and its interactions with actin binding proteins. It focuses on several examples of actin complexes with cytoskeletal proteins and considers some approaches that facilitate the mapping of cross-linked peptides by mass spectrometry. Finally, the book outlines HDX of intact proteins, an excellent tool for examining protein dynamics (Chapter 11) and methods for deriving biophysical data on ligand association (Chapter 12).

REFERENCES

Busenlehner, L. S. and Armstrong, R. N., 2004. Insights into enzyme structure and dynamics elucidated by amide H/D exchange mass spectrometry. *Arch Biochem Biophys* 433, 34–46.

Burley, S. K., Almo, S. C., Bonanno, J. B., Capel, M., Chance, M. R., Gasterland, T., Lin, D., Sali, A., Studier, F. W., and Swaminathan, S., 1999. Structural genomics: beyond the human genome project. *Nat Genet* 23, 151–157.

Brenowitz, M., Chance, M. R., Dhavan, G., and Takamoto, K., 2002. Probing the structural dynamics of nucleic acids by quantitative time-resolved and equilibrium hydroxyl radical ‘footprinting’. *Curr Opin Struct Biol* 12, 648.

Chance, M. R., Sclavi, B., Woodson, S. A., and Brenowitz, M., 1997. Examining the conformational dynamics of macromolecules with time-resolved synchrotron X-ray ‘footprinting’. *Structure* 5, 865–869.

Chance, M. R., Bresnick, A. R., Burley, S. K., Jiang, J. S., Lima, C. D., Sali, A., Almo, S. C., Bonanno, J. B., Buglino, J. A., Boulton, S., Chen, H., Eswar, N., He, G., Huang, R., Ilyin, V., McMahan, L., Pieper, U., Ray, S., Vidal, M., and Wang, L. K., 2002. Structural genomics: a pipeline for providing structures for the biologist. *Protein Sci* 11, 723–738.

Chance, M. R., Fiser, A., Sali, A., Pieper, U., Eswar, N., Xu, G., Fajardo, J. E., Radhakannan, T., and Marinkovic, N., 2004. High-throughput computational and experimental techniques in structural genomics. *Genome Res* 14, 2145–2154.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U., and Sali, A., 2007. Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci* Chapter 2, Unit 2.9.

Gavin, A. C., 2005. Keystone symposia: proteomics and bioinformatics and systems and biology. *Expert Rev Proteomics* 2, 291–293.

Guan, J. Q., Vorobiev, S., Almo, S. C., and Chance, M. R., 2002. Mapping the G-actin binding surface of cofilin using synchrotron protein footprinting. *Biochemistry* 41, 5765.

Guan, J. Q., Almo, S. C., and Chance, M. R., 2004. Synchrotron radiolysis and mass spectrometry: a new approach to research on the actin cytoskeleton. *Acc Chem Res* 37, 221–229.

Gupta, S., Sullivan, M., Toomey, J., Kiselar, J., and Chance, M.R., 2007. The Beamline X28C of the Center for Synchrotron Biosciences: a national resource for biomolecular structure and dynamics experiments using synchrotron footprinting. *J Synchrotron Radiat* 14, 233–243.

Kamal, J. K. and Chance, M. R., 2008. Modeling of protein binary complexes using structural mass spectrometry data. *Protein Sci* 17, 79–94.

Katta, V. and Chait, B. T., 1993. Hydrogen/deuterium exchange electrospray ionization mass spectrometry: a method for probing protein conformational changes in solution. *J Am Chem Soc* 115, 6317.

Kiselar, J. G., Maleknia, S. D., Sullivan, M., Downard, K. M., and Chance, M. R., 2002. Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. *Int J Radiat Biol* 78, 101.

Maleknia, S. D., Ralston, C. Y., Brenowitz, M. D., Downard, K. M., and Chance, M. R., 2001. Determination of macromolecular folding and structure by synchrotron X-ray radiolysis techniques. *Anal Biochem* 289, 103.

Novak, P. and Giannakopulos, A., 2007. Chemical cross-linking and mass spectrometry as structure determination tools. *Eur J Mass Spectrom* 13 (2), 105–113.

Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M., and Sali, A., 2004. A structural perspective on protein–protein interactions. *Curr Opin Struct Biol* 14, 313–324.

Sali, A., Glaeser, R., Earnest, T., and Baumeister, W., 2003. From words to literature in structural proteomics. *Nature* 422, 216–225.

Takamoto, K. and Chance, M. R. In: Myers, R. A., editor. *Encyclopedia of Molecular Cell Biology and Molecular Medicine.* 2nd ed. Weinheim: Wiley VCH; 2004, 521–548.

Takamoto, K. and Chance, M. R., 2006. Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annu Rev Biophys Biomol Struct* 35, 251–276.

Zhang, Z. and Smith, D. L., 1993. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* 2, 522.

Zheng, X., Wintrode, P. L., and Chance, M. R., 2008. Complementary structural mass spectrometry techniques reveal local dynamics in functionally important regions of a metastable serpin. *Structure* 16, 38–51.