1. Introduction

Proteins are essential components of living organisms and participate in nearly every biochemical process within cells. Examples of the processes include enzyme catalysis, cell signalling, host defense, metabolism, etc. These large, complex bio-molecules are connected by long chains of amino acids, that fold in very intricate patterns giving rise to a unique three-dimensional conformation. The biological function and physicochemical properties of a protein are determined by this higher order structure Ecroyd & Carver (2008); Hegyi & Gerstein (1999); Sadowski & Jones (2009). Most proteins tend to achieve the lowest possible free energy of the polypeptide chain and the surrounding solvent forming a native structure under physiological conditions Anfinsen (1973). This tightly folded conformation typically represents the biologically active state necessary for performing the required biochemical task. However, these macromolecules also have a temporal behavior leading to significant flexibility and dynamic motion because of the fluctuations in the surrounding electrostatic forces and hydrogen bonds that are important for maintaining conformations Henzler-Wildman & Kern (2007); Teilm et al. (2009). These temporal variations are important for certain functions such as protein-protein interactions and protein stability Kamerzellan & Middaugh (2008); Travaglini-Allocatelli et al. (2009); van der Kamp et al. (2010). Thus, it is the presence of both spatial and temporal characteristics that allows for an ensemble of various molecular conformations to exist in solution. Change in environment of proteins, such as solvent acidity, urea concentration, temperature fluctuations, can change the folding pattern of the protein. Studying these partially or fully denatured states provides insights for understanding a variety of in vivo processes such as structural changes associated with aggregation, signal transduction, and transportation across membranes. Certain biological conditions can cause misfolding and aggregation of proteins, often causing severe disorders such as Alzheimer’s disease, spongiform encephalopathies, and certain forms of diabetes Dobson (2003). Many genetic diseases are caused by protein-folding disorders, because an altered gene results in a modified protein sequence which is not able to undergo native folding and results in the disease phenotype Dobson (2001). Proteins have the ability to interact with one another, and can also bind to smaller ligands, which forms the basis of signaling and regulatory processes, playing a critical role in the mechanisms of drug activity. Owing to the
critical importance of structure-function paradigm, the pioneering scientists, Anfinsen and Stein and Moore, who established the relationship between protein structure and function, were awarded the Nobel Prize in Chemistry in 1972 Anfinsen (1973); Moore & Stein (1973). The static protein structures commonly seen in X-ray images from the literature depict only one of the many possible conformations that the protein assumes at a particular instant in time. In fact, conformational dynamics are essential for mediating multifaceted functional roles performed by many proteins Fenimore et al. (2002); Frauenfelder et al. (1991); Huang & Montelione (2005). One particular example is the case of enzymes that require induced-fit binding behavior for proper operation Falke (2002); Schulz (1992). This suggests that the structural rearrangements of such proteins represents a well balanced compromise between a highly ordered core conformation to ensure specificity, and a relatively flexible and dynamic state that maintains diverse functionality. However, there are certain proteins that remain disordered, and without any associated characteristic structure under physiological conditions, that fold specifically only while binding to another target Gunasekaran et al. (2003); Sugase et al. (2007); Wright & Dyson (1999); Yi et al. (2007). Nevertheless, most proteins behave according to the function-dictated-by-structure principle.

The above discussion illustrates the close interplay between the processes of protein folding, dynamics, conformation, intermolecular interaction, and function. In this Chapter, we will explore these structure function concepts as they relate to the design of novel pharmaceutical products. For example, these structure-dynamics-function aspects outlined above dictate action mechanisms of protein drugs (also called therapeutic protein, protein pharmaceutical, protein biopharmaceutical, or just biopharmaceutical) including their appropriate design and development to treat disease. Second, technologies like x-ray crystallography have been more difficult to apply to membrane proteins, likely the most important targets for small molecule drugs, and where information on the structural consequence of ligand binding are critical to drug development. This provides a strong interest in the development and application of reliable, sophisticated analytical techniques for thorough structural examination of therapeutic and membrane proteins in order to ensure and/or understand appropriate functionality and safety of the both small and large molecule drug development. This chapter introduces effective techniques that help realize this goal and demonstrates their application across monoclonal antibodies and membrane proteins. Monoclonal antibodies are designed to bind to specific protein targets in the cell blocking function of the targets; while membrane proteins perform essential processes in the cell, such as controlling the flow of information and materials between cells and mediating activities like nerve impulses and hormone action.

2. Significance

The function and efficacy of protein drug and biologic therapies is determined by the structure of the protein and its ability to interact with the surrounding partners. The interrogation and verification of the three dimensional conformation becomes critical in order to demonstrate the consistency of structure and function in biologics development. This necessitates the deployment of reliable, sensitive, and high-resolution techniques capable of examining higher order structure of such biomolecules in detail.

Biopharmaceutical manufacturers are required to demonstrate the consistency of the conformational complexity to the regulatory agencies. Traditional biophysical techniques used for this purpose include circular dichroism (CD), fluorescence, ultraviolet (UV), differential scanning calorimetry (DSC), isothermal titration calorimetry (ITC), analytical ultracentrifugation (AUC), and Fourier transform infrared spectroscopy (FTIR) Pain (2000).
These techniques provide information on a global state of the molecule, and typically output the average of overall information across the whole protein. These methods can in some cases successfully detect small changes between highly similar proteins, but the differential signal will not specifically identify the defect across the global readout. Some methods may reveal signal from only a limited number of residues (e.g., aromatic amino acid residues) in the protein and the absence of a such residues in the area of interest may lead to loss of information. Alternative sensitive, sophisticated tools for structure determination, nuclear magnetic resonance spectroscopy (NMR) and X-ray crystallography, allow both sensitive and specific probes of small structural changes Drenth (1999); Ramsey & Purcell (1952). NMR spectroscopy provides detailed structural information on proteins in solution, which is based upon distance constraints obtained from nuclear Overhauser effect Wuthrich (1990). However, these methods require high protein concentrations and can face significant challenges in the case of large proteins (in the case of NMR), and some proteins are not amenable to X-ray crystallography due to their inability to form crystals. In addition, these techniques tend to be fairly complex, and the conformation of the protein observed in the crystal is just a particular (although high resolution) conformation pulsed by the crystal lattice. This provides a strong interest to employ alternative practical methods to detect small, local differences within proteins in solution for both large and small proteins and at a range of protein concentrations. Mass spectrometry (MS) has established itself as a crucial technique in the biochemist’s repository of tools over the past two decades, and many different flavors of protein MS are available with a variety of choices for sample preparation, molecular ionization, detection, and instrumentation. Structural proteomics techniques such as covalent labeling Maleknia et al. (2001); Suckau et al. (1992), hydrogen/deuterium exchange (H/DX) Wales & Engen (2006), and chemical cross-linking Back et al. (2003); when coupled with highly sensitive mass spectrometry instruments; alleviate many of the above limitations and have shown promising results in the past decade. This chapter focuses on the fundamentals of these techniques, discusses challenges and limitations experienced by each method, and concludes with the successful application examples of these analytical tools.

3. Techniques

The following techniques can be used to determine conformational change, binding stoichiometry, and affinity for protein-ligand interactions.

3.1 Hydrogen/deuterium exchange mass spectrometry (H/DX-MS)

H/DX methods were introduced in 1990s and have now powerfully established themselves for probing the biomolecular structure Bai et al. (1995); Englander & Kallenbach (1983); Hvidt & Nielsen (1966); Krishna et al. (2004); Wales & Engen (2006); Woodward et al. (1982). The principle behind the technique is that protein backbone amide hydrogens are exchangeable with deuterium atoms from the solvent surrounding the protein at specific exchange rates that can be measured experimentally. The amide hydrogens at the surface exchange very rapidly, while those buried in the core have much slower exchange rates. The backbone amide hydrogens participating in the formation of hydrogen bonds will also have relatively slower exchange rates. Hence, the rate of exchange of hydrogens provides valuable insights into the bio-molecular backbone, secondary structure and structural stability. Fig 1 shows the overall schematics of an H/DX experiment Wales & Engen (2006). The protein under consideration is subjected to a deuterium rich environment that labels surface accessible residues, followed by quenching of the reaction. The incorporation of deuterium into the biomolecule under consideration results from the natural process of hydrogen exchange with
deuterium from the surrounding environment. There are a variety of methods available for the introduction of deuterium into a peptide or protein; and various experimental strategies are used for investigating the biomolecular exchange as seen in Fig 1. The protein can either be studied intact (for global exchange analysis), or can be digested by proteolysis (for local exchange analysis), and intact proteins or peptide fragments can be analyzed using mass spectrometry, which is able to measure the increase in mass as hydrogen atoms are exchanged for deuterium. Specific solvent accessible residues in the protein show an increased mass in the mass spectrometer readout. The experiment is repeated multiple times, each time increasing the duration of the deuterium pulse exposure to the protein, allowing for the study of deuterium exchange rate kinetics.

Although H/DX technique has shown promise for utility in biopharmaceutical studies, greater automation, seamless coupling of high performance separation, and sophisticated software for automated data interpretation is required for a more routine implementation of H/DX-MS into commercial experiments Houde et al. (2011); Wales et al. (2008). Various approaches have been developed recently towards the automated data analysis for data acquisition and post-processing Chalmers et al. (2006); Kazazic et al. (2010); Pascal et al. (2009). The H/DX-MS method has been successfully applied to study structural changes introduced by kinase activation, compare isoform-specific differences in binding to a common ligand, and to map the epitopes of monoclonal antibodies Houde et al. (2009); Lee et al. (2004); Stokasimov & Rubenstein (2009).

3.2 Hydroxyl-radical mediated covalent labeling mass spectrometry or protein footprinting

Another popular method for investigating macromolecular conformation in solution is protein footprinting, also called Covalent Labeling, which was invented initially to characterize the sites of DNA-protein interaction Brenowitz et al. (1986); Galas & Schmitz (1978); Humayun et al. (1977); Schmitz & Galas (1980). The technique was further extended to protein structure examination by subjecting them to limited proteolysis in conjunction with separation using SDS polyacrylamide gel electrophoresis, with the first report of protein footprinting appearing in 1988 Sheshberadaran & Payne (1988). The advent of sophisticated analytical tools such as mass spectrometry for examining cleaved fragments of proteins, significantly improved the spatial resolution of the technique. A key distinction from H/DX-MS is that most labeling reagents target side-chains, while HDX-MS specifically examines the bio-molecular backbone and protein secondary structure.

The basic principle behind hydroxyl radical mediated protein footprinting approaches for probing solvent accessible residues is similar to that of H/DX-MS technique. The overall schematic for experimental setup for a covalent labeling experiment is shown in Fig 2. The protein solution is exposed to hydroxyl radicals, generated by multiple methods, which leads to stable, covalent oxidative modifications on the surface accessible residues Hambly & Gross (2005); Maleknia et al. (2001); Sharp et al. (2003; 2004); Takamomo & Chance (2006). The chemistry of amino acid and peptide oxidation using MS revealed that in dilute aqueous solution, oxidative modification of side chains is observed in a much more predominant form as compared with backbone cleavage or cross-linking. These stable side chain modifications result in mass shifts, which can be easily revealed by isolating protein fragment and then comparing the masses to unmodified forms of the protein. Thus, labeling is followed by subjecting the protein to proteolysis and high pressure liquid chromatography coupled with mass spectrometry as in the case of H/DX-MS method. Tandem mass spectrometry (MS/MS) methods have been found to be particularly suited for further identifying and localizing the specific sites of oxidation Chance (2001); Kiselar et al. (2002); Maleknia et al. (1999). Thus, the structural resolution of covalent labeling is very high, and at the single side chain level. In the
Fig. 1. Overall scheme for hydrogen exchange mass spectrometry experiments. A: Pulse labeling. After a protein has been exposed to a perturbant (chemical denaturant, heat, pH, binding, complex formation, pressure, etc.), unfolded regions (gray) become labeled with deuterium (red) during a quick pulse of D2O (typically 10 s). Deuterium exchange is quenched by reducing the pH and temperature. B: Continuous labeling. D2O buffer is added to a protein (in H2O buffer) such that the final D concentration is >95%. After a set period of time, an aliquot of the labeled protein is removed from the original tube and mixed with quench buffer to reduce the pH and temperature. Aliquot removal is repeated for subsequent labeling times. The protein concentration and solution volume are controlled such that all the aliquots are identical upon quench except for the amount of time the protein was exposed to D2O. C: Localized exchange information. Quenched samples (from part A, part B, or both) are digested with pepsin or another acid protease. The resulting peptides are analyzed with online HPLC-ESI-MS or with MALDI-MS. The resulting data analysis provides information on deuterium exchange in short fragments of the peptide backbone. D: Global exchange information. Quenched samples (from part A, part B, or both) are directly analyzed with HPLC-ESI-MS or MALDI-MS. The data provide a global picture of how the protein behaves in D2O. Reprinted with permissions from Wales & Engen (2006) Copyright 2006 John Wiley and Sons.
typical workflow, series of samples are exposed to variable doses, and a dose-response curve is generated for observed peptides individually in order to provide relative quantitation of oxidation as a function of hydroxyl radical exposure time. The generation of stable, covalent modifications allows a wide range of samples and proteases to be employed under broad solution conditions and pH values. The reactivity of side chains to hydroxyl radical attack initially and the attenuation of this reactivity as a result of structural perturbation such as ligand binding, unfolding, or macromolecular interactions provides insights into the change in surface accessibility at particular sites under consideration. Since the side chains get modified during the procedure, specific probe sites can be investigated using tandem mass spectrometry, while, in the case of H/DX-MS, the conformational changes may be attributed only to a specific peptide fragment. However, the two approaches are complementary to each other since H/DX-MS characterizes backbone secondary structure and stability while protein footprinting probes the side chains of residues.

The interpretation of high volumes of data resulting from covalent labeling experiments used to pose as the biggest bottleneck for the overall experiment, thus, limiting their potential. A typical hydroxyl radical-mediated covalent labeling experiment leads to multiple oxidation states of various amino acid side chains Takamoto & Chance (2006); Xu & Chance (2007), leading to a challenging task for data analysis. This bottleneck has now been eliminated with the advent of ProtMapMS, a computational analytical tool, that is specifically tailored to meet the needs of covalent labelling experiments Kaur et al. (2009). Figure 3 illustrates typical liquid chromatographic elution profile results automatically generated in a covalent labeling experiment using ProtMapMS. The four plots in Fig 3 represent the chromatographic elution plots from a doubly charged insulin B-chain peptide 23-29 for an X-ray exposure time of 0, 8, 15, and 20 ms successively. The unoxidized form of the peptide is indicated by cyan (m/z = 430.22) color, while the different oxidative forms are blue, green, and red. Interestingly, five green peaks labeled A-E in Fig 3 represent the oxidatively labeled products for the peptide incorporating one oxygen atom represent five unique isomeric forms of the peptide molecule, differing in the position of the attached oxygen atom within the same peptide. Fig 3 shows that the relative intensities of the modified forms increase as the amount of X-ray exposure time to the protein increases. This behavior is expected since the protein molecules have increased opportunity to react with hydroxyl radicals. The oxidative forms of peptide are seen to elute at a slightly different time (although in close proximity) than their unoxidized counterpart. Improvements in two specific areas will help in making the covalent labeling experiments more routine. More accurate quantitative relationship of solvent accessibility and the side chain reactivity is highly desirable to add quantitative rigor to the structural characterization. This will provide specific constraints that can be used in computational modeling approaches for a more comprehensive analysis. Flexible computational modeling approaches should be developed that allow for including surface accessibility constraints in a quantitative manner for more accurate results. Such improvements will lead the way for oxidative footprinting and other covalent labeling approaches using MS to be utilized for understanding the conformation and dynamics of very complex macromolecular assemblies. Footprinting technique has been very successfully applied in the past for RNA structure analysis Sclavi et al. (1998). There is a great potential for similar progress for protein structure prediction by incorporating the technique into a wider utilization.

3.3 Chemical cross-linking

Covalent cross-linking is another important technique for characterizing the connectivity of solution-phase complexes, and for obtaining new intramolecular or intermolecular distance constraints between biomolecules Sinz (2003); Vasilescu et al. (2004); Wine et al. (2002).
Fig. 2. Hydroxyl radical footprinting: data collection and data analysis. Top panel: Protein is exposed to hydroxyl radical and modified covalently. The resulting protein sample is then digested by protease or chemical cleavage to fragments that are suitable in size for mass spectrometry. The experiment is carried out for each individual protein and for the protein complex. In a tight binding interface, some regions are protected from hydroxyl radical attack. Middle panel: Peptides are separated by liquid chromatography and introduced into a mass analyzer. The selected ion chromatograms (SIC) are constructed for each ion (with particular mass) as a function of retention time. By monitoring the mass and time, we know what species appears at what retention time. By integrating peak areas in SIC, we can calculate the total indicated ion abundance. Bottom panel: The determinations of modification rates are performed by calculating the loss of intact peptide in order to maximize the interrogation of intact material. Reprinted with permissions from Takamoto & Chance (2006) Copyright 2006 Annual Reviews.
It involves covalently attaching two specific functional groups of the protein(s) under investigation by means of a special reagent called cross-linker. The location and the identity of the created cross-links imposes a distance constraint on the location of the respective side chain sites and provides important clues on the three-dimensional conformation of the protein or a protein complex. Coupling chemical cross-linking with sensitive mass spectrometric analysis allows to characterize the position(s) of the introduced cross links for generating distance constraints. The wide variety of crosslinking reagents allow for varied specificities towards numerous functional groups such as primary amines, sulfhydryls, or carboxylic acids; and the wide range of spacer lengths offered by different cross-linking reagents allow the possibility to address a broad range of scientific questions. However, owing to the inherent complexity of the reaction mixtures, the identification of the cross-linked products can be quite tedious. The greatest challenge of utilizing chemical cross-linking and MS analysis is the lack of computational tools that can effectively interpret the enormous complexity of the reaction mixtures. There is a significant overhead of labor intensive manual processing involved in the data analysis since all the existing programs exhibit their specific limitations Sinz (2006). Some of the limiting bottlenecks have been eliminated with the advent of specialized search programs such as GPMAW, xQuest, searchXlinks, VIRTUALMSLAB, and ASAP de Koning et al. (2006); El-Shafey et al. (2006); Peri et al. (2001); Rinner et al. (2008); Wefing et al. (2006). Further progress into an integrated suite of algorithms addressing comprehensive needs for chemical cross-linking combined with mass spectrometry would greatly facilitate making it a generally applicable technique for rapid protein structure characterization for biopharmaceutical experiments.

Cross-linking and covalent labeling methods are complementary to each other - cross-linking methods provide distance constraints for two amino acids whereas covalent labeling approaches derive information about protein surface mapping since the reactions are typically controlled by the accessibility of surface amino acids to the covalent labeling reagent.
3.4 Future trends
Experimental data from structural proteomics experiments can be used together with computational structural modeling techniques such as comparative modeling and threading. The stand alone theoretical models without support from experimental data lack reliability, especially in the case of ab-initio modeling where suitable templates may not be available. Hybrid approaches resulting from a combination of theoretical modeling and experimental methods such as hydrogen-deuterium exchange and covalent labelling are gaining increasing popularity, allowing to combine the merits from both the methods Pantazatos et al. (2004); Zhu et al. (2003). The results from experimental analysis specifically provide explicit constraints such as distance constraints in the case of chemical cross-linking, that reflect the surface accessibility or burial of particular sites, which can be included for refining computational structure prediction models, hence greatly reducing the model space to be considered while increasing the reliability from complementary approaches.

4. Applications
4.1 Study of membrane proteins using protein footprinting
Proteins embedded in membranes assist in water or ion transport in signaling processes across the biological membrane. Typically, transmembrane proteins are comprised of hydrophobic cores with ionizable or charged residues at specific locations that are crucial for their appropriate functionality Muller et al. (2008). G protein-coupled receptors (GPCRs) comprise a large protein family of transmembrane receptors that sense molecules outside the cell and activate the signal transduction pathways inside and regulate cellular responses Rosenbaum et al. (2009). The presence of ordered, structural waters are likely to be important factors to impart structural plasticity required for agonist-induced signal transmission for allosteric activation of the G protein-coupled receptors (GPCRs) Rosenbaum et al. (2007). The functionality of these ordered water molecules is not clearly known. They may provide structural stabilization, mediate conformational changes in signaling, neutralize charged residues, or carry out a combination of all these functions. Structural investigation of GPCR superfamily members using radiolytic footprinting revealed the presence of conserved embedded water molecules likely to be important for GPCR function Angel, Gupta, Jastrzebska, Palczewski & Chance (2009).

The behavior of soluble proteins with hydroxyl radical footprinting is well-characterized, such that the intrinsic reactivity and the solvent accessibility of the side chains govern their observed reactivity Chance et al. (1997); Kiselar et al. (2002); Takamoto & Chance (2006). However, these approaches have not been investigated for membrane proteins, factors influencing labeling or the overall scavenging effects of detergents or lipids have not been well understood. Recently, in order to gain insights into membrane proteins, radiolytic protein footprinting was used to interrogate the structural dynamics of ground state (rhodopsin), photoactivated (Meta II), and inactive ligand-free receptor (opsin) and native membranes Angel, Gupta, Jastrzebska, Palczewski & Chance (2009). In contrast to the previous literature on soluble proteins, oxidative modifications were found on residues located in both solvent-accessible and solvent-inaccessible regions. The oxidized residues within the transmembrane domain were labeled, and their reactivity was found to be varying as a function of rhodopsin activation state. Using radiolytic hydroxyl radical labeling in conjunction with H$_2$O$^{18}$ solvent mixing, it was discovered that labeling within the transmembrane region is highly influenced by the tightly bound waters and that regions undergoing local conformational alterations and water reorganization experience changes in the oxidation status.
Fig 4 shows the findings of the study, indicating alterations in rates of oxidation introduced going from ground state to activated receptor, reflecting local structural changes upon the formation of both Meta II and opsin. No exchange of the structural waters was observed with the surrounding solvent in either the ground state or for the Meta II or opsin states. However, oxidative labeling of selected side chain residues within the transmembrane helices was observed and activation-induced changes in local structural constraints likely mediated by dynamics of both water and protein were revealed. This work suggests a possible general mechanism for water-dependent communication in family A GPCRs, and illustrates the role of radiolytic footprinting for characterizing the structure and dynamics of the transmembrane region, including dynamics of water in membrane proteins, and has the potential to define allosteric channels for other transmembrane signalling proteins, and ion channels.

The implications for these results are considerable for the design of drugs to target important Type A GPCRs, e.g. serotonin, adenosine, or β2-adrenergic receptor. If this model for rhodopsin functional activation is correct, it means that effective drugs mediate specific and local changes in water/side-chain interactions within the transmembrane region and that this rearrangement mediates the correct and efficient signaling across the membrane. In addition it focuses attention on water molecules and their potential rearrangements in recent crystallographic data of Type A family GPCRs Angel, Chance & Palczewski (2009). Recent extensions of the covalent labeling methodology to ion channel structure-function studies Gupta et al. (2010), where the movement of water coupled to the rearrangement of specific side chains was specifically tracked in the mechanism of channel opening, show the emerging power of the method to reveal important structure function considerations relevant for drug development.

4.2 Characterizing monoclonal antibody IgG1 using hydrogen-deuterium exchange

Monoclonal antibodies (mAb) are used both in fundamental research and in clinical settings as highly specific therapeutic agents for treating an array of different diseases. Currently, recombinant immunoglobulin gamma (IgG) mAbs comprise the largest percentage of molecules in the biopharmaceutical development pipeline. This provides a strong motivation for utilizing new or improved analytical methods and tools for mAb characterization. Presently, there is very limited information available on crystal structures of entire IgGs. Moreover, such cases provide information only of a very stable structure sampled by the protein; while lacking information on the conformational dynamics or in-solution motion.

In a recent study, H/DX-MS has been used to study both global and local conformational behavior of a recombinant monoclonal IgG1 antibody to obtain detailed conformational dynamics Houde et al. (2009). It demonstrates the capabilities of H/DX-MS as a powerful analytical tool to study large protein biopharmaceuticals such as mAbs. The conformational features of an intact glycosylated IgG1 are compared against its deglycosylated form. This assists in drawing conclusions to determine how glycosylation affects the IgG1 conformation. First, deuterium exchange into the intact form of IgG1 was measured. This is useful for providing information about the overall solvent accessibility of the protein and, also indicates whether the protein is amenable to H/DX-MS experiments so that further investigation can be performed. The next step performed the analysis of exchange into isolated Fab/Fc fragments. Next, the intact protein was labeled and digested (after quenching the deuterium labeling reaction) using pepsin as a protease. Digestion protocols were performed for both glycosylated and deglycosylated versions of the IgG1, followed by analysis using mass spectrometry. Five independent experiments were performed, each containing an undeuterated sample and five different labeling times.
Fig. 4. Pictorial summary of modification rate constants. Radiolytic modification rate constants were determined for many residues in rhodopsin (Left), Meta II (Center), and opsin (Right). Residues with rate constants $>0.1$ s$^{-1}$ are rendered as spheres colored by rate constant ranges: 0.5-1.2 s$^{-1}$, light blue; 1.3-3.9 s$^{-1}$, light green; 4.0-5.9 s$^{-1}$, green; 6.0-7.9 s$^{-1}$, light-yellow; 8.0-9.9 s$^{-1}$, yellow; 10-14.9 s$^{-1}$, light-orange; 15-25 s$^{-1}$, orange; $>200$ s$^{-1}$, red.

Following photoactivation, modification rates increased for M86, C140, M143, the pair of residues in helix IV I154 and M155, M163, and M288. Residues exhibiting decreased modification rates were Y301, P303, and Y306 in helix VII. There also was a reduced modification rate of M86 and F116 in opsin as compared with the two other states. The mixed modification of peptide 137-146, comprising part of the C-II loop, showed a large increase in the rates of detectable modification for opsin relative to ground state and activated rhodopsin, whereas M183 in the E-II loop exhibited no change in modification rate as a function of receptor activation state. The carboxyl terminal peptide did not show a marked difference in modification rates between the three states of the receptor. Changes in rates of oxidation observed when comparing ground state and activated receptor reflect local structural changes upon formation of both Meta II and opsin. Reprinted with permissions from Angel, Gupta, Jastrzebska, Palczewski & Chance (2009) Copyright 2009 National Academy of Sciences.

Figure 5 depicts the oxidative changes detected in IgG1 in the presence and absence of glycosylation. Analysis of H/D exchange pattern into the intact, glycosylated IgG1 indicated that the molecule was folded, very stable, and could be analyzed with very high sensitivity. Since the approach can detect subtle, localized changes within the protein, H/D exchange could be localized to very specific regions of the antibody. Deglycosylation resulted in changes in the IgG1 conformation, and were characterized by comparing H/D exchange rates of the glycosylated and deglycosylated forms of the antibody. Two specific regions of the IgG1 (residue positions 236-253 and 292-308) were found to have experienced change in H/D exchange properties upon deglycosylation. These results are consistent with previous findings using X-Ray crystallography and NMR techniques associating the role of glycosylation in the interaction of IgG1 with Fc receptors. Overall, H/DX-MS showed that changes in
Fig. 5. Comparison of deuterium levels in IgG1 with and without glycosylation. (A) The model structure of IgG1, with the glycosylation indicated in black sticks. Parts colored blue indicate regions where the deglycosylated form had, over all time points, less deuterium (more protection from exchange). Parts colored red indicate regions where the deglycosylated form had, over all time points, more deuterium (less protection from exchange). Note that although blue regions appear to be more surface accessible than the red regions, the conformational disturbances introduced during the deglycosylation process result in greater protection in blue areas than their red counterparts. (B) Representative deuterium incorporation profiles comparing exchange in heavy chain residues 292-308 (PREEQYNSTYRVSVLT), 236-242 (LGGPSVF), and 242-253 (FLFPPKPDKDLTM). The solid black line represents data from the glycosylated form, and the dotted line represents data from the deglycosylated form. Reprinted with permission from Houde et al. (2009). Copyright 2009 American Chemical Society.
conformation as a result of deglycosylation were in areas critical for Fc receptor binding. The data illustrate the utility of H/DX-MS to provide valuable information on the higher order structure of antibodies and characterizing conformational changes that these molecules may experience upon modifications such as glycosylation, possibly affecting the functionality of the protein.

Recently, the application of covalent labeling to the analysis of glycoprotein structure was demonstrated Wang et al. (2010). This shows that covalent labeling can also be successfully applied to monoclonal antibodies to determine the structure-function relationships relevant to biopharmaceutical development.

5. Conclusions

There is a close inter-relation between the functionality of a protein and the processes of protein folding, dynamics, three-dimensional structure, and intermolecular interaction. These aspects further dictate action mechanisms of a protein drug and its effectiveness in the treatment of a disease. The examination of primary sequence, specific protein modifications, and the three dimensional conformation are important variables used to demonstrate structural equivalency of clinically relevant formulations for biologics development. An evolving regulatory climate is likely to require highly accurate data on "higher-order structure" of biologics for the approval of biosimilars in the US. This provides strong impetus for developing reliable, sensitive, and high-resolution analytical technologies along with associated computational methods for detailed primary and secondary structural interrogation of biomolecules. MS based techniques are proving to be indispensable tools for monitoring protein folding, structure, and dynamics. Rapid progress in this technology has allowed the possibility of highly sophisticated experiments for addressing complex biological questions.

Many different configurations of protein MS are now available with a wide range of choices for sample preparation, molecular ionization, detection, and instrumentation. This chapter introduced MS based structural proteomics techniques such as H/D exchange, oxidative covalent labeling, and chemical cross-linking. These techniques, when used in conjunction with highly sensitive mass spectrometry instruments; overcome many of the limitations experienced by traditional biophysical methods, and have gained wide popularity with promising results in the past decade. In the H/D exchange process, a covalently bonded hydrogen atom from the protein backbone is replaced by a deuterium atom from the surrounding environment, or vice versa. It provides information about the solvent accessibility of various parts of the molecule. The rate of H/D exchange imparts understanding of protein backbone and the secondary structure. Covalent labeling is based on a similar principle, except that specific labeling reagents can be used that lead to stable, covalent modifications on the solvent accessible residues. Protein footprinting is a popular covalent technique in which hydroxyl radicals are generated that create specific oxidative modifications on the surface accessible residues, which helps in mapping protein surfaces. Most covalent labeling reagents target side chains, while H/DX-MS specifically probes protein backbone and tertiary structure. Chemical cross linking involves covalently attaching two specific inter- or intra-molecular functional groups of side chains by means of a cross-linking reagent. The cross-linking agent imposes a distance constraint on the respective functional groups, providing valuable information on the three dimensional structure of a macromolecule or macromolecular assembly.

Greater automation of experimental workflow along with reliable and sophisticated computer software will pave the way for a more routine incorporation of structural mass spectrometry
into commercial experiments. The future trends lie in the use of structural proteomics experimental data coupled with computational modeling techniques such as comparative modeling and threading. Hybrid approaches combining the theoretical models and experimental methods will allow to combine the merits from both approaches. Experimental results can provide specific, explicit constraints such as distance limitations or surface accessibility information, which can be included for refining computational structure prediction models for more robust and reliable results.

The chapter concludes with the discussion of successful application examples of MS based structural proteomics tools in the context of the design of novel pharmaceutical products. Structural examination of GPCR superfamily members using hydroxyl radical mediated footprinting coupled with H$_2$O$_{18}$ labeling showed the presence of conserved embedded water molecules that are likely to be important for GPCR function. These structural waters were not found to exchange with the surrounding solvent in either the ground state or for the Meta II or opsin states. On the other hand, oxidative modification of selected side chain residues within the transmembrane helices was detected and activation-induced changes in local structural constraints were revealed, likely to be mediated by the dynamics of both water and protein. The results suggest the possibility of a general mechanism for water-dependent communication in family A GPCRs. This example illustrates the importance and potential of radiolytic footprinting for characterizing the structure and dynamics of the transmembrane region, including dynamics of water in membrane proteins.

H/DX-MS has been used to characterize both global and local structural behavior of a recombinant monoclonal IgG1 antibody to study detailed conformational dynamics. The intact, glycosylated form of IgG1 was found to be folded and stable, while the deglycosylation resulted in changes in the IgG1 conformation. This was evident by comparing H/D exchange rates of the glycosylated and deglycosylated forms of the antibody. Two specific regions of the IgG1 (residue positions 236-253 and 292-308) were found to have experienced change in exchange properties upon deglycosylation. The data illustrate the utility of H/DX-MS to provide insights into the higher order structure of antibodies and characterizing conformational changes that these molecules may experience upon modifications such as glycosylation, possibly affecting the functionality of the protein.

Covalent labeling has also shown successful deployment towards monoclonal antibodies characterization to establish the structure-function relationships in context of biopharmaceutical development. Such studies are currently ongoing at several pharmaceutical companies and are likely to have a significant impact on the development of both new drugs and biosimilars in the near future.

6. References

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains, *Science* 181: 223–230.

Angel, T. E., Chance, M. R. & Palczewski, K. (2009). Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors, *Proc. Natl. Acad. Sci. U.S.A.* 106: 8555–8560.

Angel, T. E., Gupta, S., Jastrzebska, B., Palczewski, K. & Chance, M. R. (2009). Structural waters define a functional channel mediating activation of the GPCR, rhodopsin, *Proc. Natl. Acad. Sci. U.S.A.* 106: 14367–14372.

Back, J. W., de Jong, L., Muïjsers, A. O. & de Koster, C. G. (2003). Chemical cross-linking and mass spectrometry for protein structural modeling., *J Mol Biol* 331(2): 303–313.
Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange, *Science* 269: 192–197.

Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986). "Footprint" titrations yield valid thermodynamic isotherms, *Proc. Natl. Acad. Sci. U.S.A.* 83: 8462–8466.

Chalmers, M. J., Busby, S. A., Pascal, B. D., Ye, H., Hendrickson, C. L., Marshall, A. G. & Griffin, P. R. (2006). Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry, *Anal. Chem.* 78: 1005–1014.

Chance, M. R. (2001). Unfolding of apomyoglobin examined by synchrotron footprinting, *Biochem. Biophys. Res. Commun.* 287: 614–621.

Chance, M. R., Sclavi, B., Woodson, S. A. & Brenowitz, M. (1997). Examining the conformational dynamics of macromolecules with time-resolved synchrotron x-ray ‘footprinting’, *Structure* 5(7): 865–869.

de Koning, L. J., Kasper, P. T., Back, J. W., Nessen, M. A., Vanrobaey, F., Van Beeumen, J., Gherardi, E., de Koster, C. G. & de Jong, L. (2006). Computer-assisted mass spectrometric analysis of naturally occurring and artificially introduced cross-links in proteins and protein complexes, *FEBS J.* 273: 281–291.

Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 356: 133–145.

Dobson, C. M. (2003). Protein folding and misfolding, *Nature* 426: 884–890.

Drenth, J. (1999). *Principles of protein x-ray crystallography*, Springer-Verlag.

Ecroyd, H. & Carver, J. A. (2008). Unraveling the mysteries of protein folding and misfolding, *IUBMB Life* 60: 769–774.

El-Shafey, A., Tolic, N., Young, M. M., Sale, K., Smith, R. D. & Kery, V. (2006). "Zero-length" cross-linking in solid state as an approach for analysis of protein-protein interactions, *Protein Sci.* 15: 429–440.

Englander, S. W. & Kallenbach, N. R. (1983). Hydrogen exchange and structural dynamics of proteins and nucleic acids, *Q. Rev. Biophys.* 16: 521–655.

Falke, J. J. (2002). Enzymology. A moving story, *Science* 295: 1480–1481.

Fenimore, P. W., Frauenfelder, H., McMahon, B. H. & Parak, F. G. (2002). Slaving: solvent fluctuations dominate protein dynamics and functions, *Proc. Natl. Acad. Sci. U.S.A.* 99: 16047–16051.

Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. (1991). The energy landscapes and motions of proteins, *Science* 254: 1598–1603.

Galas, D. J. & Schmitz, A. (1978). DNase footprinting: a simple method for the detection of protein-DNA binding specificity, *Nucleic Acids Res.* 5: 3157–3170.

Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D. & Nussinov, R. (2003). Extended disordered proteins: targeting function with less scaffold, *Trends Bichem. Sci.* 28: 81–85.

Gupta, S., Bavro, V. N., D’Mello, R., Tucker, S. J., Venien-Bryan, C. & Chance, M. R. (2010). Conformational changes during the gating of a potassium channel revealed by structural mass spectrometry, *Structure* 18: 839–846.

Hambly, D. M. & Gross, M. L. (2005). Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale, *Journal of the American Society for Mass Spectrometry* 16(12): 2057 – 2063.

Hegyi, H. & Gerstein, M. (1999). The relationship between protein structure and function: a comprehensive survey with application to the yeast genome, *J. Mol. Biol.* 288: 147–164.

Henzler-Wildman, K. & Kern, D. (2007). Dynamic personalities of proteins, *Nature* 450: 964–972.
Houde, D., Arndt, J., Domeier, W., Berkowitz, S. & Engen, J. R. (2009). Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry, *Anal. Chem.* 81: 5966.

Houde, D., Berkowitz, S. A. & Engen, J. R. (2011). The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies, *J Pharm Sci* 100: 2071–2086.

Huang, Y. J. & Montelione, G. T. (2005). Structural biology: proteins flex to function, *Nature* 438: 36–37.

Humayun, Z., Kleid, D. & Ptashne, M. (1977). Sites of contact between lambda operators and lambda repressor, *Nucleic Acids Res.* 4: 1595–1607.

Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins, *Adv. Protein Chem.* 21: 287–386.

Kamerzell, T. J. & Middaugh, C. R. (2008). The complex inter-relationships between protein flexibility and stability, *J Pharm Sci* 97: 3494–3517.

Kaur, P., Kiselar, J. G. & Chance, M. R. (2009). Integrated algorithms for high-throughput examination of covalently labeled biomolecules by structural mass spectrometry, *Anal. Chem.* 81: 8141–8149.

Kazazic, S., Zhang, H. M., Schaub, T. M., Emmett, M. R., Hendrickson, C. L., Blakney, G. T. & Marshall, A. G. (2010). Automated data reduction for hydrogen/deuterium exchange experiments, enabled by high-resolution Fourier transform ion cyclotron resonance mass spectrometry, *J. Am. Soc. Mass Spectrom.* 21: 550–558.

Kiselar, J. G., Maleknia, S. D., Sullivan, M., Downard, K. M. & Chance, M. R. (2002). Hydroxyl radical probe of protein surfaces using synchrotron x-ray radiolysis and mass spectrometry., *Int J Radiat Biol* 78(2): 101–114. URL: [http://dx.doi.org/10.1080/09553000110094805](http://dx.doi.org/10.1080/09553000110094805)

Krishna, M. M., Hoang, L., Lin, Y. & Englander, S. W. (2004). Hydrogen exchange methods to study protein folding, *Methods* 34: 51–64.

Lee, T., Hoofnagle, A. N., Kabuyama, Y., Stroud, J., Min, X., Goldsmith, E. J., Chen, L., Resing, K. A. & Ahn, N. G. (2004). Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry, *Mol. Cell* 14: 43–55.

Maleknia, S. D., Brenowitz, M. & Chance, M. R. (1999). Millisecond radiolytic modification of peptides by synchrotron x-rays identified by mass spectrometry., *Anal Chem* 71(18): 3965–3973.

Maleknia, S. D., Ralston, C. Y., Brenowitz, M. D., Downard, K. M. & Chance, M. R. (2001). Determination of macromolecular folding and structure by synchrotron x-ray radiolysis techniques., *Anal Biochem* 289(2): 103–115. URL: [http://dx.doi.org/10.1006/abio.2000.4910](http://dx.doi.org/10.1006/abio.2000.4910)

Moore, S. & Stein, W. H. (1973). Chemical structures of pancreatic ribonuclease and deoxyribonuclease, *Science* 180: 458–464.

Muller, D. J., Wu, N. & Palczewski, K. (2008). Vertebrate membrane proteins: structure, function, and insights from biophysical approaches, *Pharmacol. Rev.* 60: 43–78.

Pain, R. H. (2000). *Mechanisms of protein folding*, Oxford University Press.

Pantazatos, D., Kim, J. S., Klock, H. E., Stevens, R. C., Wilson, I. A., Lesley, S. A. & Woods, V. L. (2004). Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS, *Proceedings of the National Academy of Sciences of the United States of America* 101(3): 751–756. URL: [http://www.pnas.org/content/101/3/751.abstract](http://www.pnas.org/content/101/3/751.abstract)
Pascal, B. D., Chalmers, M. J., Busby, S. A. & Griffin, P. R. (2009). HD desktop: an integrated platform for the analysis and visualization of H/D exchange data, J. Am. Soc. Mass Spectrom. 20: 601–610.
Peri, S., Steen, H. & Pandey, A. (2001). GPMMAW—a software tool for analyzing proteins and peptides, Trends Biochem. Sci. 26: 687–689.
Ramsey, N. F. & Purcell, E. M. (1952). Interactions between nuclear spins in molecules, Physical Review 85(1): 143–144.
Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L. N., Beck, M., Schmidt, A., Mueller, M. & Aebersold, R. (2008). Identification of cross-linked peptides from large sequence databases, Nat. Methods 5: 315–318.
Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobylka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C. & Kobylka, B. K. (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function, Science 318: 1266–1273.
Sadowski, M. I. & Jones, D. T. (2009). The sequence-structure relationship and protein function prediction, Curr. Opin. Struct. Biol. 19: 357–362.
Schmitz, A. & Galas, D. J. (1980). Sequence-specific interactions of the tight-binding I12-X86 lac repressor with non-operator DNA, Nucleic Acids Res. 8: 487–506.
Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M. & Woodson, S. A. (1998). RNA Folding at Millisecond Intervals by Synchrotron Hydroxyl Radical Footprinting, Science 279(5358): 1940–1943.
Sharp, J. S., Becker, J. M. & Hettich, R. L. (2003). Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry., Anal Biochem 313(2): 216–225.
Sharp, J. S., Becker, J. M. & Hettich, R. L. (2004). Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry., Anal Chem 76(3): 672–683.
Sugase, K., Dyson, H. J. & Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein, Nature 447: 1021–1025.
Takamoto, K. & 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.
Teilum, K., Olsen, J. G. & Kragelund, B. B. (2009). Functional aspects of protein flexibility, *Cell. Mol. Life Sci.* 66: 2231–2247.

Travaglini-Allocatelli, C., Ivarsson, Y., Jemth, P. & Gianni, S. (2009). Folding and stability of globular proteins and implications for function, *Curr. Opin. Struct. Biol.* 19: 3–7.

van der Kamp, M. W., Schaeffer, R. D., Jonsson, A. L., Scouras, A. D., Simms, A. M., Toofanny, R. D., Benson, N. C., Anderson, P. C., Merkley, E. D., Rysavy, S., Bromley, D., Beck, D. A. & Daggett, V. (2010). Dynameomics: a comprehensive database of protein dynamics, *Structure* 18: 423–435.

Vasilescu, J., Guo, X. & Kast, J. (2004). Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry, *Proteomics* 4: 3845–3854.

Wales, T. E. & Engen, J. R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics, *Mass Spectrom. Rev.* 25: 158–70.

Wales, T. E., Fadgen, K. E., Gerhardt, G. C. & Engen, J. R. (2008). High-speed and high-resolution UPLC separation at zero degrees Celsius, *Anal. Chem.* 80: 6815–6820.

Wang, L., Qin, Y., Ilchenko, S., Bohon, J., Shi, W., Cho, M. W., Takamoto, K. & Chance, M. R. (2010). Structural analysis of a highly glycosylated and unliganded gp120-based antigen using mass spectrometry, *Biochemistry* 49: 9032–9045.

Wefing, S., Schnaible, V. & Hoffmann, D. (2006). SearchXLinks. A program for the identification of disulfide bonds in proteins from mass spectra, *Anal. Chem.* 78: 1235–1241.

Wine, R. N., Dial, J. M., Tomer, K. B. & Borchers, C. H. (2002). Identification of components of protein complexes using a fluorescent photo-cross-linker and mass spectrometry, *Anal. Chem.* 74: 1939–1945.

Woodward, C., Simon, I. & Tuchsen, E. (1982). Hydrogen exchange and the dynamic structure of proteins, *Mol. Cell. Biochem.* 48: 135–160.

Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm, *J. Mol. Biol.* 293: 321–331.

Wuthrich, K. (1990). Protein structure determination in solution by NMR spectroscopy, *The journal of biological chemistry* 265(36): 22059–22062.

Xu, G. & Chance, M. R. (2007). Hydroxyl radical-mediated modification of proteins as probes for structural proteomics, *Chemical Reviews* 107(8): 3514–3543.

Yi, S., Boys, B. L., Brickenden, A., Konermann, L. & Choy, W. Y. (2007). Effects of zinc binding on the structure and dynamics of the intrinsically disordered protein prothymosin alpha: evidence for metalation as an entropic switch, *Biochemistry* 46: 13120–13130.

Zhu, M. M., Rempel, D. L., Du, Z. & Gross, M. L. (2003). Quantification of protein-ligand interactions by mass spectrometry, titration, and H/D exchange: PLIMSTEX, *Journal of the American Chemical Society* 125(18): 5252–5253.

URL: http://pubs.acs.org/doi/abs/10.1021/ja029460d
Proteomics was thought to be a natural extension after the field of genomics has deposited significant amount of data. However, simply taking a straight verbatim approach to catalog all proteins in all tissues of different organisms is not viable. Researchers may need to focus on the perspectives of proteomics that are essential to the functional outcome of the cells. In Integrative Proteomics, expert researchers contribute both historical perspectives, new developments in sample preparation, gel-based and non-gel-based protein separation and identification using mass spectrometry. Substantial chapters are describing studies of the sub-proteomes such as phosphoproteome or glycoproteomes which are directly related to functional outcomes of the cells.

Structural proteomics related to pharmaceutics development is also a perspective of the essence. Bioinformatics tools that can mine proteomics data and lead to pathway analyses become an integral part of proteomics. Integrative proteomics covers both look-backs and look-outs of proteomics. It is an ideal reference for students, new researchers, and experienced scientists who want to get an overview or insights into new development of the proteomics field.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Parminder Kaur and Mark R. Chance (2012). The Utility of Mass Spectrometry Based Structural Proteomics in Biopharmaceutical Biologics Development, Integrative Proteomics, Dr. Hon-Chiu Leung (Ed.), ISBN: 978-953-51-0070-6, InTech, Available from: http://www.intechopen.com/books/integrative-proteomics/the-utility-of-mass-spectrometry-based-structural-proteomics-in-biopharmaceutical-biologics-developm