ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/51503403

Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting

ARTICLE *in* ANALYTICAL CHEMISTRY · AUGUST 2011

Impact Factor: 5.64 · DOI: 10.1021/ac200567u · Source: PubMed

CITATIONS
37

READS
34

2 AUTHORS, INCLUDING:

Liwen Wang
Case Western Reserve University
14 PUBLICATIONS **353** CITATIONS

SEE PROFILE

Available from: Liwen Wang
Retrieved on: 29 January 2016

NIH Public Access
Author Manuscript
Anal Chem. Author manuscript; available in PMC 2012 October 1.

Published in final edited form as:
Anal Chem. 2011 October 1; 83(19): 7234–7241. doi:10.1021/ac200567u.

Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting

Liwen Wang[1] and **Mark R. Chance**[1,2,*]

[1]Center for Proteomics & Bioinformatics, School of Medicine, Case Western Reserve University, Cleveland, OH, 44106, USA

[2]Center for Synchrotron Biosciences, Case Western Reserve University, Brookhaven National Laboratory, Upton, NY 11973

Abstract

Structural MS is a rapidly growing field with many applications in basic research and pharmaceutical drug development. In this feature article the overall technology is described and several examples of how hydroxyl radical based footprinting MS can be used to map interfaces, evaluate protein structure, and identify ligand dependent conformational changes in proteins are described.

Driven by noncovalent interactions, polypeptide chains fold into a myriad of shapes that correspond to functional domains of biological significance. Therefore, assessment of protein structure is critical for understanding the functions of proteins at a molecular level. By comparing protein amino acid sequences and structures, scientists can classify proteins into family groups and deduce potential structure–function relationships by homology with proteins of known structure. Techniques such as X-ray crystallography and NMR spectroscopy can determine protein structures at atomic-level resolution. These experimentally-determined structures of proteins, nucleic acids, and complex assemblies are collected into a database called the Protein Data Bank (PDB) [1]. Technical limitations of atomic resolution techniques have limited the direct determination of protein 3D structure when compared to the number of protein sequences available. Swiss-Prot [2], a protein sequence database, currently contains 500,000 unique sequences, whereas only 51,535 protein structures were registered in PDB as of 2009, and many of these PDB structures are highly redundant [3]. Thus, experimental structures for the majority of known proteins do not exist. To fill this gap, computational modeling strategies based on homology [4] and de novo prediction methods are necessary to provide structure–function predictions for systems in which the structure is poorly characterized [5].

Structural MS is a powerful technology that can provide direct assessment of protein structure, both to answer specific questions based on structure and as a confirmation of results of structural modeling. In general, structural MS includes three overall approaches: covalent labeling of macromolecules, chemical cross linking, and hydrogen deuterium exchange (HDX). Covalent labeling approaches are based on footprinting technologies, which are historically well-established methods to probe protein and nucleic acid macromolecular structure in solution[6,7].

*Address correspondence to Dr. Mark R Chance, Case Western Reserve University, Case Center for Proteomics & Bioinformatics, Department of Physiology & Biophysics, Cleveland, OH 44106, USA. Phone (216) 368-4406, Fax (216) 368-3812, mark.chance@case.edu.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Footprinting is especially powerful for mapping the interfaces between the domains of macromolecules and mapping conformational changes due to ligand binding. In footprinting experiments, chemical reagents probe the solvent accessible surface of the macromolecule after modification or cleavage of macromolecules [6–8]. One of the most popular reagents for footprinting (and structural MS) is the hydroxyl radical, which has a Van der Waals surface similar to water and thus can provide very high resolution structural information. Hydroxyl radicals can react with side-chain sites in proteins that are solvent accessible; the resultant mass shift marks the site of modification and can be used to track the accessibility of that site.

Alternatively, chemical cross-linking uses reagents to form a covalent bond between neighboring domains in inter-or intramolecular protein contacts. Detection of the cross-linked sites by MS can provide valuable information on protein tertiary and quaternary structure [9]. However, the potential complexity of the cross-linked species and the low abundance of cross-linked peptides are challenges for this method.

HDX is a valuable structural probe of proteins [10] in which the protein of interest is exposed to deuterium in D_2O. Carefully controlled exchange of amide protons causes mass shifts for peptides of interest if the peptide regions are solvent accessible and not involved in stable secondary structure. To trap the exchange competent species, pepsin digestion at pH ~2.5 and low temperature are applied to minimize the back-exchange prior to MS analysis. All three methods are quite valuable and somewhat complementary. Because of back-exchange, HDX has some size limitations in the species that can be examined; also, much of the information is at the peptide level, although it typically spans the entire protein sequence. However, its ability to probe backbone secondary structure conformation in solution is rivaled only by NMR, which has other limitations, particularly in the amounts of material required. In contrast to HDX, hydroxyl radical mediated protein footprinting (PF) primarily labels side chains of proteins in solution, providing no information on the backbone structure. However, the labeled products are very stable, and very large macromolecular complexes can be easily examined using a variety of enzymes or solution conditions to fragment the protein before MS detection [11–13]. However, as discussed below, the relative reactivity of the side chains varies by up to 1000×; thus, the ability to probe a specific site of interest may be limited by the amino acids present.

Many methods of generating hydroxyl radicals for PF experiments have been developed, such as chemical methods [14], radiolysis of water [15–17], or photolysis of hydrogen peroxide [18]. Each method has advantages and disadvantages. For instance, the most popular chemical method to generate hydroxyl radical is Fenton's reagent [14], which decomposes hydrogen peroxide (H_2O_2) to hydroxyl radicals via catalysis of transition metals such as ferrous ion (II). It is a low cost experiment with easy to acquire reagents. But addition of chemicals such as EDTA, used in the Fenton system to increase the solubility of the metal ions, may perturb the biological conditions of the samples or bias the reactivity by interaction with macromolecules. Also, Fenton reactions may require minutes of exposure, which can permit unwanted secondary reactions.

Radiolysis of water by electron pulse radiolysis [15], gamma rays [16], or X-rays [17] requires no chemical additions, but samples must be in buffered solution to resist the change of pH caused by water decomposition. Also, an obvious disadvantage is the requirement of ionizing radiation. However, radiolytic footprinting can be performed in vivo[19]. Gamma ray sources suffer from similar disadvantages to Fenton's reagent, e.g., long exposure times. However, electron and synchrotron beam methods have the advantage of rapid exposure of the sample (microseconds to milliseconds) that can minimize the influence of secondary reactions. This is also an advantage of methods that rely on photolysis of peroxide because

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the hydroxyl radicals are generated quickly, and very rapid bi-molecular recombination of radicals and fast reactions with macromolecules or scavengers cause short lifetimes [20]. A potential disadvantage of the photolysis approach is that samples must be mixed with peroxide prior to photolysis, which can oxidize sulfur containing residues. Also, the laser photolysis apparatus is somewhat complex and expensive to build and maintain.

The methods that generate radicals quickly (ionizing radiation and photolysis) provide, overall, a very similar pattern of oxidation and thus probe similar sites within proteins. Thus, this article focuses on the downstream analytical steps after oxidation. It outlines significant results in the field that illustrate protein structural analysis, both static and time resolved, as well as the application of the technology to assist computational modeling to interpret protein structures.

HYDROXYL RADICAL REACTIONS WITH PROTEINS

Because they are similar in size to water molecules, hydroxyl radicals can easily oxidize side chains of surface reactive amino acid residues in aqueous solution. These hydroxyl radicals react with protein functional groups depending on their specific chemistry and solvent accessibility. Peptide amide bonds form the backbone of a protein from which the 20 different amino acid side chains extend. The backbone C-α and amide bonds are much less reactive targets of radical attack compared to the preferred side chain sites. The relative reactivity of the twenty side chains follows the order: Cys > Met > Trp > Tyr > Phe > Cystine > His > Leu ~ Ile > Arg ~ Lys~ Val > Ser ~ Thr ~ Pro > Gln ~ Glu > Asp ~ Asn > Ala > Gly [8]. Lower reactivity residues such as Asp, Asn, Ala, Gly, and residues whose oxidation products are difficult to detect, such as Ser and Thr, often do not provide much information for protein structure analysis. However, fourteen out of the twenty amino acids, encompassing ~65% of the sequence of a typical protein, provide protein structure information with good resolution [8]. The irreversibly modified protein products can be quantified and detected by LC/MS, typically after proteolysis, providing measures of surface accessibility for specific side chains in the protein structure. Figure 1 displays a flowchart of hydroxyl radical mediated PF structural analysis.

MS-BASED PROTEOMICS APPROCHES

MS is important for specific discovery of proteins and characterization of protein modifications. Identification of protein modifications by MS-based proteomic analysis usually requires proteolytic cleavage of proteins to peptides by enzymes such as trypsin, Glu-C, or Asp-N that cleave protein sequences at specific amino acid residues. For example, trypsin, the most frequently used enzyme, cleaves proteins at the C terminus of Lys and Arg. Glu-C cleaves proteins at the C terminus of Glu and Asp. Asp-N cleaves proteins at the N terminus of Asp. A protein can be digested to a mixture of several to hundreds of unique peptides based on the protein size and sequence and the proteolytic enzyme specificity. The experimental peptide masses are compared with the predicted proteolyzed peptide masses generated with data from a gene or protein database using bioinformatics tools.

In hydroxyl radical mediated PF, the radicals modify proteins via hydrogen abstraction from saturated carbon sites or hydroxyl addition to unsaturated carbon–carbon double bonds found on aromatic rings. Mass shifts of +16 Da resulting from formation of a hydroxyl group in the side chain and +14 Da from formation of a carbonyl group are most common for many reactive residues (e.g., Leu, Ile, Arg, Val, Pro, Lys) (Figure 2A [21]). Further reactions may occur on some reactive residues depending on the structure and radical environment such that different mass shifts occur on specific residues. For example, mass shifts of +48 Da, +32 Da, or –16 Da on Cys are based on formation of sulfonic acid, sulfinic acid, or Ser (replacing sulfur with oxygen atom), respectively (Figure 2B [21]). Radiolysis

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

decarboxylates the C-terminus of proteins on Glu and Asp, which generate –30 Da products. Oxidation of His can produce a series of characteristic products with mass shift of +16 Da, –22 Da, –23 Da, +5 Da, and –10 Da. Oxidation of Arg can generate a –43 Da product in addition to +16 and +14 Da products [21].

Verification of Protein Modification Sites by Tandem Mass Spectrometry

The identification of the specific site of modification, which represents the specific structural probe in the experiment, is critical for the PF approach. Tandem mass spectrometry, which includes multiple stages of mass spectrometric analysis (MS_n, where n = number of analyses, most often two), is one of the most valuable techniques for identifying proteins and annotating their modifications. Usually, the first mass analysis stage isolates a peptide precursor ion of interest from others based on its mass to charge ratio (m/z). The second mass analysis stage identifies the peptide product ions produced upon peptide fragmentation that occurs between the two analysis stages.

Tandem mass spectra are composed of product ions dissociated from peptide precursor ions, i.e. a/x ions, b/y ions, and c/z ions. In the mass spectrometer, peptides fragment predominantly at their peptide bonds. Depending on the location of bond cleavage, the fragments are denoted as a-, b-, or c-ions if the charge is retained on the N-terminus and x-, y-, or z-ions if the charge is retained on C-terminus. The differences in masses between fragment ions of a given series can be used to calculate the mass of an amino acid residue and identify the peptide's sequence [22] (Figure 1B). In addition, the modification site of a peptide can be located when a mass shift (e.g., +16 Da for oxidation of Met in Figure 1B) is observed in the product ions. Information on the chemistry of modification of specific side chains along with the tandem MS data can be used to easily identify the site of modification in a footprinting experiment.

Reverse Phase LC coupled with MS (RPLC/MS)

RPLC is a separation technique commonly used to improve the sensitivity of MS for performing protein identification and screening for modifications. Subsequent to exposure of samples to radicals and protein proteolysis, a mixture of peptides is subjected to RPLC/MS. RPLC fractionates peptides based on their hydrophobicity differences. An analytical column with C_{18} (octadecylsilyl) as the stationary phase and a solvent mixture such as water–methanol or water–acetonitrile as a mobile phase is typically used to separate peptide mixtures. After injection into the C_{18} column, peptides distribute between the two phases and separate as the mobile phase is systematically altered, sequentially releasing peptides based on differences in the hydrophobic interactions of various peptides within the two phases.

LC is often interfaced to the mass spectrometer by electrospray ionization (ESI) [23] for online analysis of biomolecules. Briefly, biomolecule samples are delivered to a capillary tip by RPLC solvent and ionized by the high voltage (~2 kV) of the capillary tip. Ionized biomolecules are then sprayed into the mass spectrometer concurrently with evaporation of solvent. As a "soft" ionization method, ESI overcomes the tendency of traditional "hard" ionization techniques, such as EI (electron impact), FAB (fast atom bombardment), to produce unwanted fragmentation of biomolecules.

BIOINFORMATICS TOOLS FOR MS DATA ANALYSIS

MS data analysis is the final and in some cases the most labor-intensive stage in the proteomic workflow. Several software packages typically aid analysis of mass spectral data sets. The most common software tools make comparisons between the experimental data and protein sequences stored in databases to identify proteins and isolate sites of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

modification. Manual processing of MS data is impossible because one experiment can generate thousands of mass spectra. Additionally, the enormous number of possible combinations for peptide matches complicates the interpretation of the data and could lead to false spectral matches. Processing MS data of unknown protein mixtures is even more daunting because a single database can generate millions of potential peptides matches.

Comparing peptide tandem mass spectra data against a known protein database is widely used for peptide/protein identification. Several software applications have been developed that automate this process, the most common of which are Sequest [24], Mascot [25], X! Tandem [26], MassMatrix [21], and OMSSA [27]. Generally, these software packages generate theoretical tandem mass spectra based on enzyme cleavage and peptide fragmentation of the target protein sequences and protein modifications. The experimental tandem mass spectra are then searched against the theoretical database and assigned a score; the highest score is generally the assignment most likely to be correct. Many functional modifications frequently observed in proteins, such as phosphorylation (+80 Da), acetylation (+42 Da), methylation (+14 Da), or modifications that are an artifact of the proteomics workflows, such as Met oxidation (+16 Da) or deamidation (+1 Da), are included as software features. Additional modifications of interest can be added manually by annotating their molecular information with respect to particular side chain sites. It is necessary to manually include the specific oxidations known in hydroxyl radical-mediated modification because of their unique qualities [21]. Table 1 lists the composition change and mass shifts corresponding to the well characterized and frequently observed modifications. The variety of oxidation reactions mediated by hydroxyl radicals complicate MS data from the proteolyzed peptide mixtures. Modifications of peptides on specific residues need to be built in for database searching; in addition, extraction of the intensities of modified and unmodified species (Figure 1C) is time consuming. To speed this process, ProtMapMS software [28] was developed for automated identification and quantification of the extent of modifications.

The oxidation reaction in radiolysis follows (pseudo) first-order kinetics, that is, the decrease of the fraction of unmodified peptide with exposure time fits a first-order kinetic equation:

$$y(t) = e^{-kt}$$

[29] where y is fraction of unmodified peptide, t is exposure time in seconds, and k is the oxidation rate constant. A dose response curve for each oxidized peptide can be plotted by calculating y versus t, and k can be derived by fitting the curve to the above equation (Figure 1D). When comparing *different* peptides, modification rates vary because of intrinsic reactivity, but for a *particular* oxidized peptide, increases in k provide direct evidence that the oxidized residues in the peptide are more solvent accessible. When comparing a specific peptide within a protein under different conditions (e.g., +/− ligand), changes in the rate of modification are directly correlated with changes in solvent accessibility, i.e., relative reactivity precisely tracks the relative solvent accessibility. However, the absolute reactivity can be instructive as well. For example if highly reactive residues (e.g., Met, Phe) are *not* modified in a PF experiment, this provides significant evidence that they have very low solvent accessibility. On the other hand if modestly reactive residues are modified (e.g., Leu, Pro), this provides significant evidence that they are quite solvent accessible.

Prior to automation of the data analysis by ProtMapMS, calculating this rate data was very labor intensive because it involved extraction of peak areas of the chromatogram for each target peptide (Figure 1C). ProtMapMS first identifies both modified and unmodified forms of the peptide and then extracts their corresponding chromatograms for the target peptide

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

precursor ions based on the window defined by a user around the mass value of interest (in ppm). Finally, it returns a dose response curve and oxidation rate of peptides.

PROTEIN FOOTPRINTING COMBINED WITH STRUCTURE MODELING TO SOLVE STRUCTURE RIDDLES

Computational modeling is in demand to fill the gap between known protein sequences and structures. Template-based structural modeling, that is, homology modeling (e.g., Swiss-Model [30] and MODPIPE [31]), was developed for modeling protein structures from similar sequences of proteins with known structures, typically from a protein template found in the PDB. When sequence similarity is not detectable or variable regions are substantial (such as for variable loops), de novo modeling (e.g., Rosetta [32]) can predict protein structures based on the global energy minimum of the native protein states.

Footprinting data combined with structural modeling can optimize and test predictive models. High-accuracy homology modeling [1 Å RMSD (root mean square deviation)] can be achieved with sequence identities >50% [33]. In such a case, experimental footprinting data are typically compared with the solvent accessible surface areas of side chains (SASA) calculated from the models. Several software are available for calculating SASA, such as the MSMS V 2.5 software [34], CCP4 program [35], and VADAR program [36,37]. Agreement between SASA value and oxidation behavior of the accessible and reactive side chains is a prerequisite for accurate modeling of a protein structure.

For instance, antigens based on HIV-1 envelope glycoprotein gp120 have been generated in order to drive vaccine development. The structure of the outer domain of gp120 (gp120-OD8), which contains the immunologically important V3 loop, was assessed through a combination of PF and comparative modeling [38]. A model of gp120-OD8 was generated by Swiss-Model server using a template of HIV-1 JR-FL gp120 core protein containing the V3 loop; in this structure gp120 is in a complex with human CD4 and the X5 antibody (PDB ID: 2B4C). The template shares 76.4% sequence identity with the sequence of gp120-OD8. In order to better compare the model and experimental PF data, the ligands were removed in the modeling (Figure 3). SASA values ($Å^2$) calculated by MSMS V 2.5 software and rate constants of oxidized peptides were compared and revealed that oxidation behavior of reactive residues of gp120-OD8 are generally consistent with those predicted from the homology model, except that rate constants of four residues in the V3 base were much higher than other adjacent residues of the V3 loop, though their predicted SASA values are similar (Figure 3). This discrepancy supports a flexible V3 loop structure of free gp120-OD8 in which the V3 tip may provide contacts with the rest of the protein while residues in the V3 base remain solvent accessible [38]. This combination of modeling and footprinting provided unique insight into the V3 loop structure.

PF data have been used to detect dynamic excursions of some residues that are buried within the protein in crystallographic data [39]. For example, PF data of serpin showed that Met 374 and Met 385 were solvent accessible residues. But the SASA data deduced from the crystal structure showed very low accessibility of the side chains overall, with zero solvent accessibility for the reactive sulfur atoms in Met 374 and Met 385. SASA data were calculated from a set of molecular dynamics (MD) trajectories of serpin [39]. The SASA values over a 6.8 ns simulation revealed dynamic excursions of the reactive sulfur atom of the Met residues, which is consistent with the PF data. Because the predication of SASA from MD simulations is believed to generate a more realistic ensemble of side chain orientations in solution [40], this method was also applied to PF of galectin-1, for which no crystal structure is available [41]. Using a homology model in the MD simulation, the data predicted solvent accessibility of the residues in the interface region that correlated with the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

PF data. Thus, a combination of computational results and PF generated a realistic picture of galectin-1. Besides these cases in which PF data confirm or reject structure models, the data can also be used as experimental constraints to define structure and interactions for docking simulations. PF data generated information on the residues participating in the binding surface and/or those involved in conformational reorganization in actin complexes analyzed using the ClusPro docking server [42]. Overall, footprinting data can provide attract or block constraints at specific molecular sites based on the observed accessibility changes to guide the final molecular complexes generated by docking [43].

MAPPING CONFORMATIONAL CHANGES OF PROTEIN INTERACTION BY PROTEIN FOOTPRINTING

PF is a powerful technique for detecting interfaces and conformational changes due to protein interactions, which can be achieved through the comparison of oxidation behavior of individual peptides of the target protein with and without the interaction. For example, PF detected the conformational change of the acton binding protein gelsolin after activation by Ca^{2+}[12]. Gelsolin is involved in remodeling the actin cytoskeleton during numerous cellular processes. Upon activation by Ca^{2+}, gelsolin undergoes essential structural rearrangements that are related directly to actin filament and actin monomer binding [44]. PF analysis confirmed conformational changes of gelsolin under Ca^{2+} activation and localized the specific domains involved in these changes. Figure 4 represents a dose response curve of gelsolin peptides in an inactive state (Ca^{2+} concentrations <1 nM, blue) and Ca^{2+} activated states (Ca^{2+} concentration = 0.2 mM, red) [12]. Some peptides (amino acid residue ranges: 49–72, 121–135, 162–166, 431–454, 722–748) in the activated state showed apparent increases of oxidation rate, indicating that these residues were more exposed after activation, while a dramatic decrease of oxidation rate in some peptides (amino acid residue ranges: 276–300, 652–686) indicates that the regions became more buried after Ca^{2+} induced activation.

Conformational changes of macromolecules in the time domain are also of interest. Rapid mixing of ligands and macromolecules to observe the kinetics of ligand dependent activation monitored by footprinting has been conducted with hydroxyl radicals to probe the folding and dynamics of nucleic acids [45–47]. Recently, hydroxyl radical mediated PF has also been used to examine protein folding [48] and protein dynamics [49]. For example, the dynamic conformation of gelsolin (as opposed to the examination of equilibrium states above) was investigated under a constant X-ray exposure time of 50 ms while solvent accessibility changes after Ca^{2+} addition were analyzed from 10 ms to 3s [49]. Similarities in the dynamics for the change in accessibility of peptides in the S1-S1 latch region (peptide 49–72), S6-C-terminal helix (peptide 722–748), and S4-inter subunit β-sheet (peptide 431–454) at 100 μM Ca^{2+} (final concentration) demonstrated that conformational changes (e.g., release of “S6 latch” and the disruption of the S4/S6 β-sheet) in these regions are concerted. Protein-folding dynamics have also been studied by time-dependent fast photochemical oxidation of proteins (FPOP) [48]. Conformational changes of barstar during folding were monitored within 1 ms (the time between the unfolded state to first observed intermediate state) by determining the oxidation extent decreases as the protein folds, consistent with burial of solvent accessible surface residues as a function of folding. The distribution of modified states (+16, +32, +48 Da, etc) in the unfolded state showed a similar pattern to that of a denatured protein. With the increase of the time between the heating pulse (which drives folding) and the FPOP probe, the amount of modified species decreases so that the centroid mass shifted with time. Comparability of rate constant (1.5 ms^{-1}) calculated from fitting the plot of shifted centroid mass versus delay time with the rate (3.1 ms^{-1}) for the transition from the unfolded state to the first intermediate as examined by other techniques illustrates

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the potential application of this method to study the protein-folding kinetics at the global protein level.

FUTURE PROSPECTS

The article demonstrates the state of art in hydroxyl radical mediated PF analysis describing its origin, various methodologies, and its diverse applications. The robustness, reproducibility, and flexibility of the method make it a highly promising structural analysis technique to complement traditional techniques such as crystallography and NMR in structural biology. Many groups are currently using hydroxyl radical mediated structural MS approaches in their research. A hydrogen exchange and covalent labeling (including PF) MS interest group was introduced by the American Society for Mass Spectrometry and has enrolled over 700 members. This group serves to share developments and applications of these techniques, including an annual short course, to accelerate progress in the field.

In terms of future useful developments, the low reactivity of residues such as Gly, Ala, Asp, Asn, Arg, Glu, and Gln and undetectablity of oxidized products of Ser and Thr make these probe sites not useful for PF. If there is no probe site in a peptide, no structural information can be provided, and this method could be improved by use of targeted MS techniques. For example, single reaction/multiple reaction monitoring (SRM/MRM) experiments can quantify the oxidation extent of residues with low reactivity (and thus low product yield) to improve the sensitivity of this method. In SRM, peak areas of selected fragment ions from unoxidized peptides and oxidized peptides have been used to quantify the oxidation extent of peptides [49]. Improved sensitivity of low abundant oxidation products (Arg, His) was demonstrated by combining PF with SRM approaches for angiotensin II peptide oxidation analysis.

Time-resolved hydroxyl radical footprinting was introduced to probe the folding of RNA [45,47]. The solvent accessibility of the RNA backbone was measured by hydroxyl radical ribose oxidation and resulting strand cleavage. The principles have been extended in recent studies of dynamic changes of protein states and protein folding and will provide impetus for greater applications of PF in examining protein dynamics in the future [48, 49].

At present, PF is a very powerful technique for identification of interfaces and conformational changes in protein complexes and protein dynamics. In the future, detecting structures of some proteins that are challenging to X-ray crystallography and NMR, such as membrane proteins, could be the focus of research interest for PF coupled with computational modeling. Most membrane proteins are important therapeutic targets because they are associated with many important functions including signaling and intercellular communication, vesicle trafficking, ion transport, and protein translocation. However, structure–function analysis of membrane proteins is limited because only 262 atomic-level structures of unique membrane proteins are publicly available (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). The task is also challenging for predictive modeling because compared to globular proteins, membrane proteins display the opposite arrangement with hydrophobic residues outside contacting the bilayer and a mix of hydrophobic and hydrophilic residues in the protein core. The potential for PF analysis to provide structural information of a membrane protein, rhodopsin, suggests that it can be a promising strategy for assessing membrane protein structures and their ligand induced conformational changes in the future [50].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Acknowledgments

This work was supported in part by grants from the National Institutes of Health (National Institute for Biomedical Imaging and Bioengineering): P30-EB-09998 and R01-EB-09688. The National Synchrotron Light Source is supported by the Department of Energy.

Biographies

Liwen Wang is Research Associate at Center for Proteomics and Bioinformatics (CPB) in Case Western Reserve University (CWRU). Dr. Wang received her Ph.D in Analytical Chemistry from the Ohio State University. She has more than 10 years experience with instrument analysis. Her current research focuses on characterization of protein structure and modifications using mass spectrometry based-structural biology and proteomics strategies.

Mark R. Chance is the Director of the CPB, Interim Chair and Professor of Genetics and Professor of General Medical Sciences at Case Western Reserve University. He also Directs the Center for Synchrotron Biosciences' at Brookhaven National Laboratory. The research in his lab mainly focuses on the important areas of structural and cellular proteomics by use of cutting-edge techniques such as X-ray crystallography, synchrotron radiolysis, mass spectrometry, and bioinformatics. Dr. Chance received his Ph.D. in Biophysics from University of Pennsylvania.

REFERENCE

1. Berman H, Henrick K, Nakamura H, Markley JL. Nucleic Acids Res. 2007; 35:D301–D303. [PubMed: 17142228]
2. Bairoch A, Boeckmann B, Ferro S, Gasteiger E. Brief Bioinform. 2004; 5:39–55. [PubMed: 15153305]
3. Westbrook J, Feng Z, Jain S, Bhat TN, Thanki N, Ravichandran V, Gilliland GL, Bluhm W, Weissig H, Greer DS, Bourne PE, Berman HM. Nucleic Acids Res. 2002; 30:245–248. [PubMed: 11752306]
4. Schwede T, Kopp J, Guex N, Peitsch MC. Nucleic Acids Res. 2003; 31:3381–3385. [PubMed: 12824332]
5. Rohl CA, Strauss CE, Misura KM, Baker D. Methods Enzymol. 2004; 383:66–93. [PubMed: 15063647]
6. Brenowitz M, Senear DF, Shea MA, Ackers GK. Proc Natl Acad Sci U S A. 1986; 83:8462–8466. [PubMed: 3464963]
7. Brenowitz M, Chance MR, Dhavan G, Takamoto K. Curr Opin Struct Biol. 2002; 12:648–653. [PubMed: 12464318]
8. Takamoto K, Chance MR. Annu Rev Biophys Biomol Struct. 2006; 35:251–276. [PubMed: 16689636]
9. Ye X, O'Neil PK, Foster AN, Gajda MJ, Kosinski J, Kurowski MA, Bujnicki JM, Friedman AM, Bailey-Kellogg C. Protein Sci. 2004; 13:3298–3313. [PubMed: 15557270]
10. Konermann L, Pan J, Liu Y-H. Chemical Society Reviews. 2011; 40:1224–1234. [PubMed: 21173980]
11. Kiselar JG, Maleknia SD, Sullivan M, Downard KM, Chance MR. Int J Radiat Biol. 2002; 78:101–114. [PubMed: 11779360]
12. Kiselar JG, Janmey PA, Almo SC, Chance MR. Mol Cell Proteomics. 2003; 2:1120–1132. [PubMed: 12966145]
13. Xu G, Chance MR. Anal Chem. 2005; 77:4549–4955. [PubMed: 16013872]
14. Tullius TD, Dombroski BA. Proc Natl Acad Sci U S A. 1986; 83:5469–5473. [PubMed: 3090544]
15. Asmus KD. Methods Enzymol. 1984; 105:167–168. [PubMed: 6328176]
16. Hayes JJ, Kam L, Tullius TD. Methods Enzymol. 1990; 186:545–549. [PubMed: 2172714]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

17. Sclavi B, Woodson S, Sullivan M, Chance M, Brenowitz M. Methods Enzymol. 1998; 295:379–402. [PubMed: 9750229]
18. Sharp JS, Becker JM, Hettich RL. Anal Chem. 2004; 76:672–683. [PubMed: 14750862]
19. Adilakshmi T, Lease RA, Woodson SA. Nucleic Acids Research. 34:e64. [PubMed: 16682443]
20. Ralston CY, Sclavi B, Sullivan M, Deras ML, Woodson SA, Chance MR, Brenowitz M. Methods Enzymol. 2000; 317:353–368. [PubMed: 10829290]
21. Xu H, Freitas MA. BMC Bioinformatics. 2007; 8:133. [PubMed: 17448237]
22. Mann M, Wilm M. Anal Chem. 1994; 66:4390–4399. [PubMed: 7847635]
23. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Science. 1989; 246:64–71. [PubMed: 2675315]
24. Eng JK, McCormack AL, Yates Iii JR. Journal of the American Society for Mass Spectrometry. 1994; 5:976–989.
25. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551–3567. [PubMed: 10612281]
26. Craig R, Beavis RC. Bioinformatics. 2004; 20:1466–1467. [PubMed: 14976030]
27. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. J Proteome Res. 2004; 3:958–964. [PubMed: 15473683]
28. Kaur P, Kiselar JG, Chance MR. Anal Chem. 2009; 81:8141–8149. [PubMed: 19788317]
29. Maleknia SD, Brenowitz M, Chance MR. Anal Chem. 1999; 71:3965–3973. [PubMed: 10500483]
30. Peitsch MC, Jongeneel CV. Int Immunol. 1993; 5:233–238. [PubMed: 8095800]
31. Sali A, Blundell TL. J Mol Biol. 1993; 234:779–815. [PubMed: 8254673]
32. Simons KT, Bonneau R, Ruczinski I, Baker D. Proteins. 1999 Suppl 3:171–176. [PubMed: 10526365]
33. Baker D, Sali A. Science. 2001; 294:93–96. [PubMed: 11588250]
34. Sanner MF, Olson AJ, Spehner JC. Biopolymers. 1996; 38:305–320. [PubMed: 8906967]
35. Acta Crystallogr D Biol Crystallogr. 1994; 50:760–763. [PubMed: 15299374]
36. Richards FM. J Mol Biol. 1974; 82:1–14. [PubMed: 4818482]
37. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS. Nucleic Acids Res. 2003; 31:3316–3319. [PubMed: 12824316]
38. Wang L, Qin Y, Ilchenko S, Bohon J, Shi W, Cho MW, Takamoto K, Chance MR. Biochemistry. 2010; 49:9032–9045. [PubMed: 20825246]
39. Zheng X, Wintrode PL, Chance MR. Structure. 2008; 16:38–51. [PubMed: 18184582]
40. Dykjaer, J.; Woods, R. NMR Spectroscopy and Computer Modeling of Carbohydrates. Vol. Vol. 930. American Chemical Society; 2006. p. 18
41. Charvatova O, Foley BL, Bern MW, Sharp JS, Orlando R, Woods RJ. J Am Soc Mass Spectrom. 2008; 19:1692–1705. [PubMed: 18707901]
42. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. Nucleic Acids Res. 2004; 32:W96–W99. [PubMed: 15215358]
43. Kamal JKA, Chance MR. Protein Science. 2008; 17:79–94. [PubMed: 18042684]
44. Pope BJ, Gooch JT, Weeds AG. Biochemistry. 1997; 36:15848–15855. [PubMed: 9398317]
45. Sclavi B, Woodson S, Sullivan M, Chance MR, Brenowitz M. J Mol Biol. 1997; 266:144–159. [PubMed: 9054977]
46. Sclavi B, Sullivan M, Chance MR, Brenowitz M, Woodson SA. Science. 1998; 279:1940–1943. [PubMed: 9506944]
47. Chance MR, Sclavi B, Woodson SA, Brenowitz M. Structure. 1997; 5:865–869. [PubMed: 9261085]
48. Chen J, Rempel DL, Gross ML. Journal of the American Chemical Society. 2010; 132:15502–15504. [PubMed: 20958033]
49. Kiselar JG, Chance MR. J Mass Spectrom. 2010; 45:1373–1382. [PubMed: 20812376]
50. Angel TE, Gupta S, Jastrzebska B, Palczewski K, Chance MR. Proc Natl Acad Sci U S A. 2009; 106:14367–14372. [PubMed: 19706523]

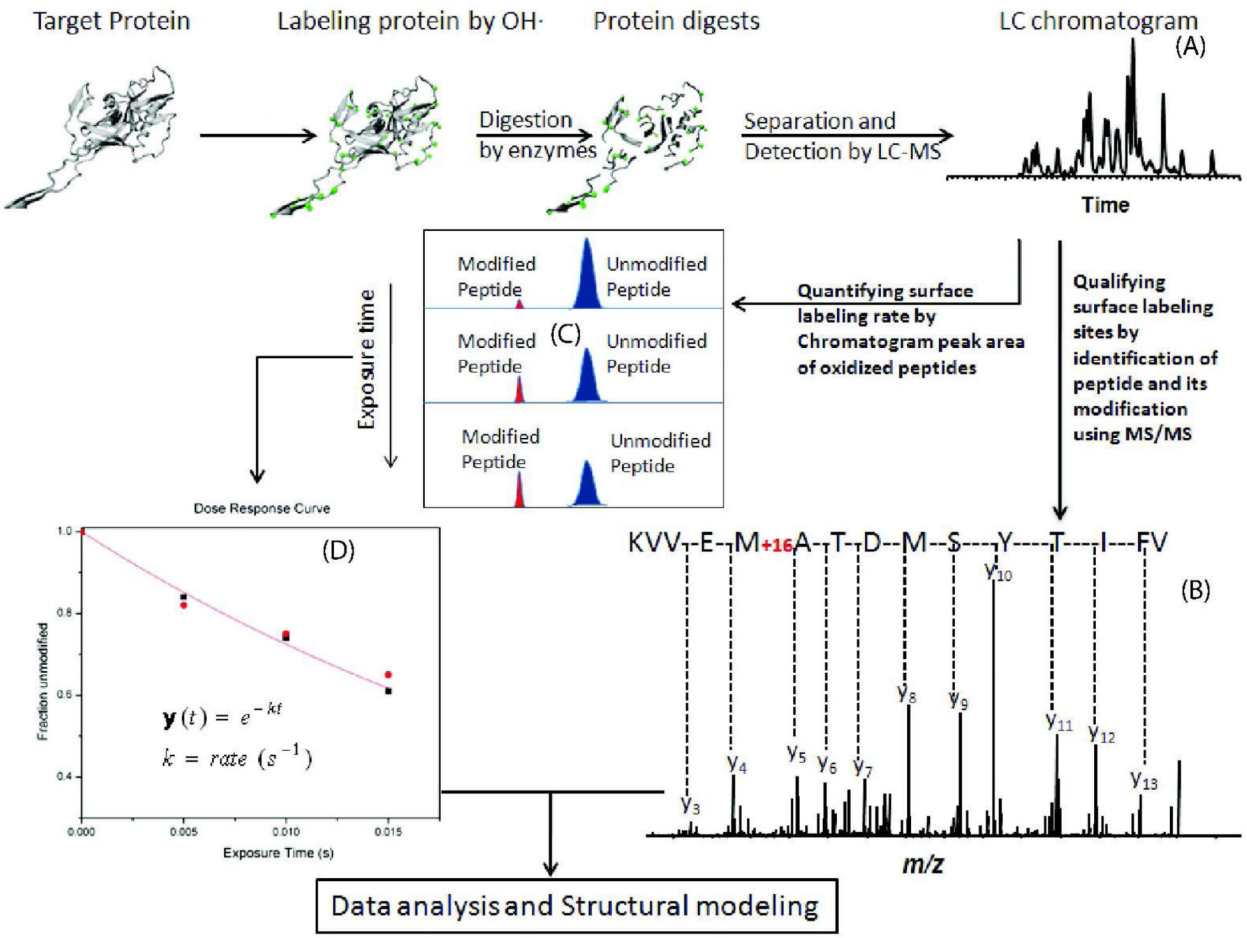

Figure 1.

Flowchart of hydroxyl radical mediated PF. A) chromatogram of the separation of protein digests using LC; B) Localization of oxidation sites by tandem MS at a precursor with sequence “VFITYS*Mox*DTAMEVVK” (*Mox* = ionized Met); C) Diagram of extraction of peaks from target oxidized peptides and their unmodified counterparts in the same LC trace with increasing exposure time to hydroxyl radicals fromtop to bottom; D) Calculation of oxidation rate by plotting the dose response curve of unmodified peptide fraction versus exposure time of hydroxyl radicals.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

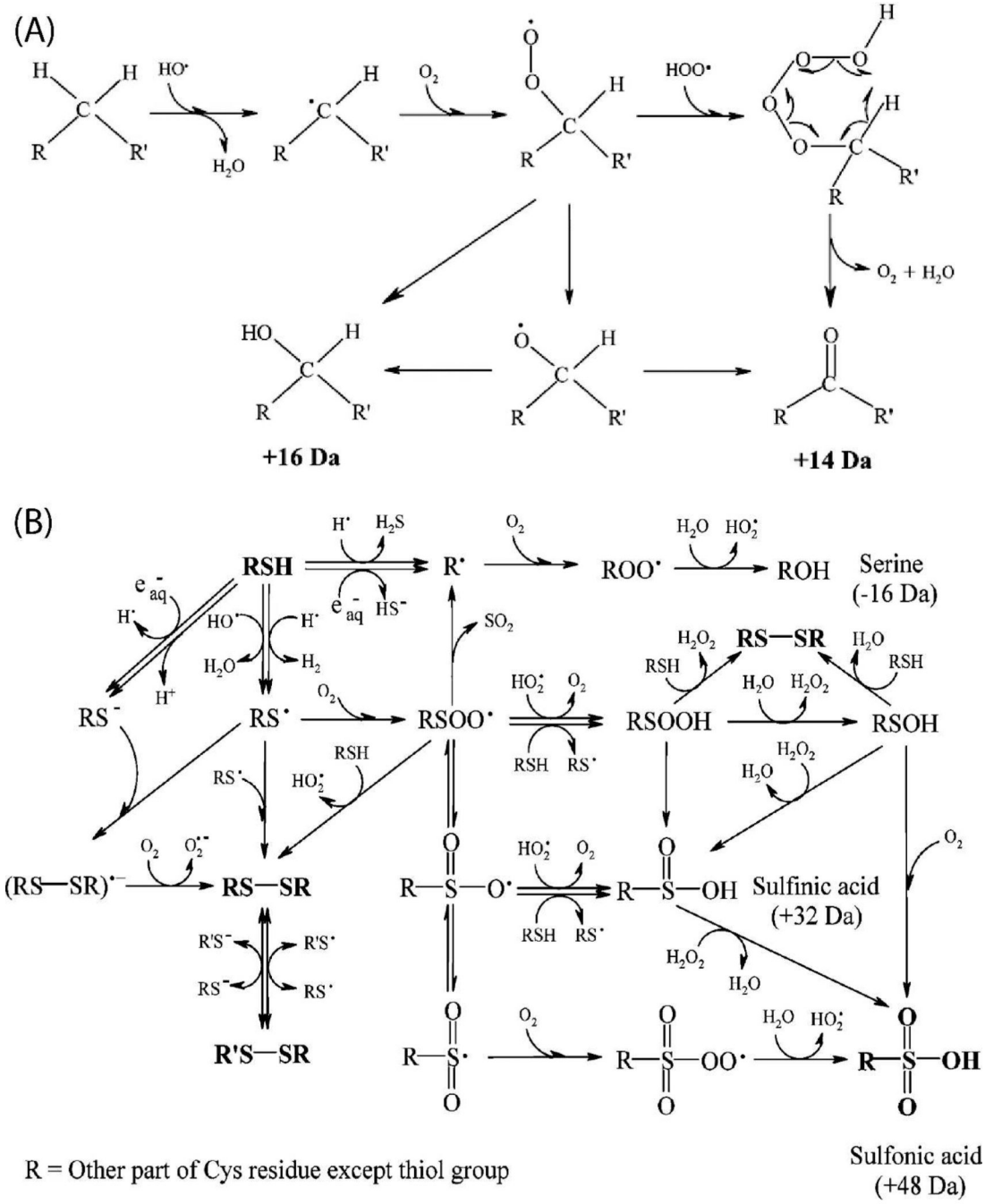

Figure 2.
Reactions involved in oxidation of A) hydrocarbon side chains and B) Cys [21].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

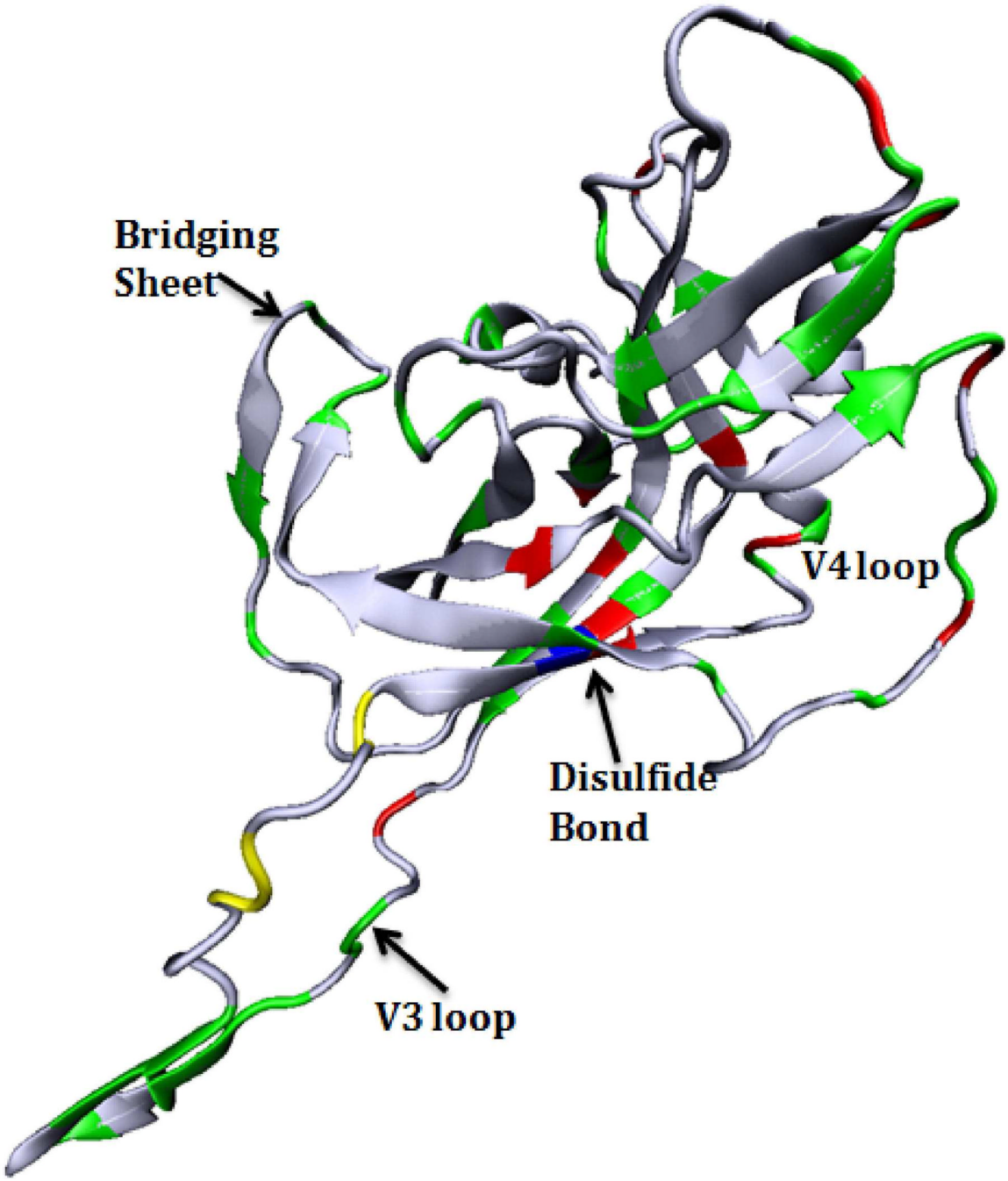

Figure 3.
Homology Model of the gp120-OD8 including V3, V4 loops and the bridging sheet. Glycosylation sites are identified experimentally by MS. Red: high mannose glycosylation sites. Green: oxidized residues in synchrotron radiolysis. Yellow: oxidized residues in N-terminal of V3 loop in synchrotron radiolysis which should be more exposed than other residues including those on V3 tip according to footprinting data. Blue: disulfide bond in V3 loop.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

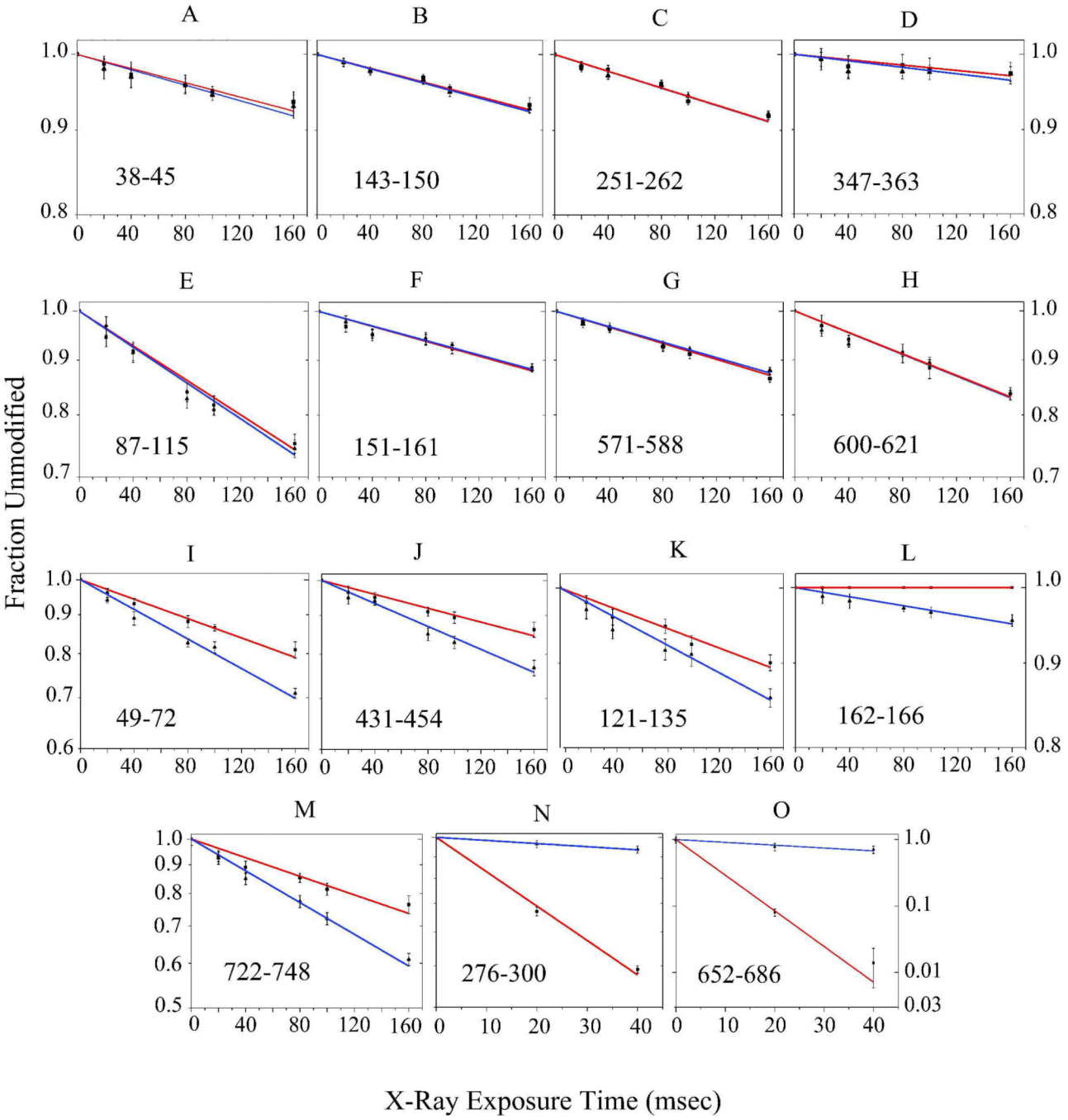

Figure 4.

Dose response curves of gelsolin peptides for Ca^{2+} bound (active) and Ca^{2+} free (inactive) states. Parts A–O present data for 15 oxidized peptides. Parts I–M: five peptides showed an increase in oxidation rate in the Ca^{2+} bound (active) state. Parts N–O: two Met-containing peptides showed dramatic decreases in oxidation rate. Parts A–H: no change was observed for these peptides. Reproduced with permission from ref. 12. (Copyright 2003, the American Society for Biochemistry and Molecular Biology.)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Composition change and mass shift of the most common hydroxyl radical oxidation products for various amino acid side chains

Amino acid residues	Composition change	Mass shift (Da)
W, M, Y, F, H, L, I, R, V, T, P, K	+ O	+15.9949
L, I, R, V, P, K	+O, −2H	+13.9793
M, W, Y, F, C	+2O	+31.9898
M	+O, −S, −C, −4H	−31.9898
C	+O, −S	−15.9772
C	+3O	+47.9847
H	+O, −2C, −N, −H	−23.0160
H	+2O, −2C, −2N, −2H	−22.0320
H	+2O, −C, −2N, −2H	−10.0320
R	+O, −3N, −C, −5H	−43.0534
D, E	−C, −2H, −O	−30.0106