To the Graduate Council:

I am submitting herewith a dissertation written by Carlee Suzanne Patterson McClintock entitled "Development of an Electrochemical Technique for Oxidative Surface Mapping to Investigate Solution-Phase Protein Dynamics using High Performance Mass Spectrometry and Advanced Informatics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

Robert Hettich, Major Professor

We have read this dissertation
and recommend its acceptance:

Cynthia Peterson

Jeffrey Becker

Hong Guo

Dean Myles

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Development of an Electrochemical Technique for Oxidative Surface Mapping to Investigate Solution-Phase Protein Dynamics with High Performance Mass Spectrometry and Advanced Informatics

A Dissertation
Presented for the
Doctor of Philosophy
Degree

The University of Tennessee, Knoxville

Carlee Suzanne Patterson McClintock
May 2010

Copyright © 2010 by Carlee Suzanne Patterson McClintock
All rights reserved.

This dissertation is dedicated to four individuals:

To my mother Laura,
whose unconditional love
continues to make anything possible.

To my husband David,
whose supportive presence by my side
helps me see my own potential.

To my sister Jessica,
whose lifelong friendship has been the source
of many shared lessons, inspirations, and celebrations.

To my son James,
whose smile lights up my world and
gives me the strength to carry on.

Acknowledgments

Special thanks are expressed to Bob Hettich, whose guidance has played a pivotal role in cultivating my scientific and professional development. The foundation for this growth was provided in large part by the early training provided by the Genome Science and Technology (GST) graduate program, as well as the opportunities supported by GST for traveling to national conferences and networking with fellow students and faculty members.

Appreciation is extended to Dr. Gary Van Berkel for the use of lab equipment, as well as to Dr. Vilmos Kertesz for contributing electrochemical expertise whenever needed. I would also like to extend my gratitude to the entire Organic and Biological Mass Spectrometry (OBMS) Group at Oak Ridge National Laboratory (ORNL) for maintaining an atmosphere conducive to graduate student fellowship.

Molecular dynamics simulations for two of the three proteins were conducted by Jerry Parks with critical contributions to interpreting the results and refreshing enthusiasm for the collaborative project. Extensive work was performed by Marshall Bern of Palo Alto Research Center to generate ByOnic search results, with greatly appreciated attention to detail and insightful personal communications. Support for InsPecT software was provided by Sam Payne and Natalie Castellana of UCSD.

This work was funded by the National Institutes of Health under the General Medicine grant 1R01-GM070754. ORNL is managed by UT-Battelle, LLC under Contract DE-AC05-00OR22725 for the U.S. Department of Energy.

Abstract

Oxidative protein surface mapping has gained popularity over recent years within the mass spectrometry (MS) community for gleaning information about the solvent accessibility of folded protein structures. The hydroxyl radical targets a wide breadth of reactive amino acids with a stable mass tag that withstands subsequent MS analysis. A variety of techniques exist for generating hydroxyl radicals, with most requiring sources of radiation or caustic oxidizing reagents. The purpose of this research was to evaluate the novel use of electrochemistry for accomplishing a comparable probe of protein structure with a more accessible tool. Two different working electrode types were tested across a range of experimental parameters, including voltage, flow rate, and solution electrolyte composition, to affect the extent of oxidation on intact proteins. Results indicated that the boron-doped diamond electrode was most valuable for protein research due to its capacity to produce hydroxyl radicals and its relatively low adsorption profile. Oxidized proteins were collected from the electrochemical cell for intact protein and peptide level MS analysis. Peptide mass spectral data were searched by two different "hybrid" software packages that incorporate *de novo* elements into a database search to accommodate the challenge of searching for more than forty possible oxidative mass shifts. Preliminary data showed reasonable agreement between amino acid solvent accessibility and the resulting oxidation status of these residues in aqueous solution, while more buried residues were found to be oxidized in "non-native" solution. Later experiments utilized higher flow rates to reduce protein residence time inside the electrochemical flow chamber, along with a different cell activation approach to improve controllability of the intact protein oxidation yield. A

multidimensional chromatographic strategy was employed to improve dynamic range for detecting oxidation of lower reactivity residues. Along with increased levels of oxidation around "reactive hotspot" sites, the enhanced sensitivity of these measurements uncovered a significant level of background oxidation in control proteins. While further work is needed to determine the full utility that BDD electrochemistry can lend protein structural studies, the experimental refinements reported here pave the way for improvements that could lead to a high-throughput structural pipeline complementary to predictive modeling efforts.

Table of Contents

1. Introduction to Oxidative Protein Surface Mapping 1
   1.1 Current State of Protein Structure Research 1
   1.2 Computational Tools for Advancing Experimental Protein Structure Research 2
   1.3 Molecular Dynamics Simulations for Evaluating Solvent Accessibility Probe Data 4
   1.4 Mass Spectrometry Analyses of Solvent Accessibility Labeling Techniques 5
   1.5 Oxidative Mapping of Proteins with Hydroxyl Radicals 6
   1.6 Motivation for Development of Electrochemical Protein Oxidation Method 8
   1.7 Identification of Oxidized Sites from Mass Spectral Data 9
   1.8 Scope of Dissertation Research Project 10
2. Experimental Platform for Oxidative Mapping with Electrochemistry 13
   2.1 Materials 13
      2.1.1 Secondary Oxidation Quenchers 13
   2.2 Electrochemistry 14
      2.2.1 Electrode Descriptions 15
      2.2.2 Operation of Electrochemical Cell for Protein Oxidation 15
      2.2.3 Cyclic Voltammetry Experiments 17
      2.2.4 Open Circuit Potential Readings 17
      2.2.5 Instrument Configuration 18
   2.3 Liquid Chromatography and Mass Spectrometry 19
      2.3.1 Intact Protein Mass Spectrometry 22
      2.3.2 Proteolytic Digestion 23
      2.3.3 Peptide LC-MS/MS Methods 24
   2.4 Informatics and Data Mining 25
      2.4.1 DBDigger Software 26
      2.4.2 InsPecT Software 27
      2.4.3 ByOnic Software 29
   2.5 Metrics of MS Search Software Packages 30

2.6 Protein Structure Evaluation 31
2.6.1 Circular Dichroism Spectroscopy 31
2.6.2 Calculation of SASA from Crystal Structures and MD Simulation Results 32
2.6.3 Evaluation of Oxidation Site Solvent Accessibility by Molecular Dynamics 33
3. Development of an Experimental Technique for Probing Higher Order Protein Structure with Electrochemical Oxidation and Mass Spectrometry 35
3.1 Introduction 36
3.2 Results 37
3.2.1 Results of Porous Flow-Through (PFT) Graphitic Carbon Electrode Experiments 39
3.2.1.1 Oxidation of Intact Proteins 39
3.2.1.2 Testing of Physiologically Relevant Buffers 40
3.2.1.3 Protein Adsorption and Electrode Regeneration 43
3.2.1.4 Peptide MS/MS Analysis 46
3.2.2 Results of Boron-Doped Diamond (BDD) Electrode Experiments 50
3.2.2.1 Oxidation of Peptide and Intact Proteins 50
3.2.2.2 Oxidation in Physiologically Relevant Buffer 53
3.2.2.3 Peptide MS/MS Analysis 53
3.2.3 Interpreting Oxidation Search Results 56
3.3 Discussion 59
3.4 Conclusions 62
4. An Electrochemistry-Based Pipeline to Probe Solution-Phase Protein Dynamics with Rapid Oxidation Exposure and Multidimensional Peptide Separations 63
4.1 Introduction 63
4.2 Experimental Refinement of Electrochemical Parameters 65
4.3 Protein-Electrode Interaction Studies 67
4.4 Mass Spectrometry Analysis of Intact Protein Oxidation 74
4.5 Peptide Separation Strategies 80
4.6 Entry Points for Bias in Data Acquisition 83
4.6.1 Liquid Chromatography Strategy 83

4.6.2 Mass Spectrometry Approach **85**
4.7 Importance of Control Measurements **86**
4.7.1 Negative Control to Identify Background Oxidation **87**
4.7.2 Positive Control to Obtain List of All Oxidation-Sensitive Sites **91**
4.8 Survey of Modified Sites in Control and Oxidized Proteins **92**
4.9 Conclusions **95**
5. Mining Oxidative Modifications from Multidimensional LC-MS Data for Increased Dynamic Range in Electrochemical Mapping of Protein Dynamics **97**
5.1 Introduction **97**
5.2 Interpreting Tandem Mass Spectra of Native and Oxidized Peptides **98**
5.3 Potential Sources of Information Loss **99**
5.3.1 Multiply Oxidized Peptides **99**
5.3.2 Unanticipated Oxidation Events **100**
5.3.3 Coeluting Oxidized Peptide Isoforms **102**
5.4 Peptides Identified from Oxidized Proteins by Two Software Packages **103**
5.4.1 Ubiquitin **106**
5.4.2 Lysozyme **113**
5.4.3 Apomyoglobin **114**
5.5 Conclusions **117**
6. Conclusions **119**
6.1 Next-Generation Optimization of Electrochemical Surface Mapping **121**
6.2 Specific Research Targets for Optimization **123**
6.3 Outlook **125**
6.4 Perspective **127**
References **128**
Appendix **138**
Vita **146**

List of Figures

Figure 2-1: Boron-doped diamond working electrode configuration inside electrochemical flow-by cell. 16
Figure 2-2: Fourier transform ion cyclotron resonance (FTICR) mass spectrometer outfitted with a 9.4-Tesla actively shielded magnet and equipped with Z-spray electrospray ionization (ESI) source. 21
Figure 2-3: Software currently available for searching MS/MS data is displayed on a continuum from total reliance on a database for spectral interpretation to pure *de novo* sequencing that does not access a database 26
Figure 2-4: Chemical structures for the five possible histidine oxidation mass shifts highlighted in Table 2-1 28
Figure 3-1: Intact protein FTICR-MS of ubiquitin before and after electrochemical oxidation. 41
Figure 3-2: SIM of the ubiquitin (A) and lysozyme (B) 9+ charge state, showing unoxidized (red) and mono-oxidized (blue) signals over time. 42
Figure 3-3: Representative MS/MS spectra show two different modifications on the same histidine residue. 48
Figure 3-4: Mass spectra from three separate experiments showing increasing abundance of mono-oxidized ubiquitin relative to the unoxidized protein. 52
Figure 3-5: Mass spectra of proteins electrochemically oxidized in physiologically relevant solution. 54
Figure 4-1: Electrochemical detection of oxidation-reduction potential for potassium sulfate solution. 68
Figure 4-2: Workflow for monitoring protein concentration with the UV detector to detect any change during electrode activation. 71
Figure 4-3: Monitoring electrode changes with CVs performed before and after protein oxidation. 72
Figure 4-4: Monitoring change in protein structure and/or concentration with CD data for unoxidized and electrochemically oxidized apomyoglobin. 75
Figure 4-5: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. 77
Figure 4-6: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. 78
Figure 4-7: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. 79
Figure 4-8: Lysozyme FTICR-MS from freshly prepared, commercially-purified stock protein with satellite peaks showing oxidation signature. 89
Figure 5-1: Triply oxidized ubiquitin peptide 1-M(+16)QIFVK(+14)TL(+14)TGK-11 identified by InsPecT. 101
Figure 5-2: Mixture "oxidisomer" tandem mass spectrum from electrochemically oxidized apomyoglobin 104

Figure 5-3: Plot of SASA values for ubiquitin Leu15 over MD simulation timeframe. Inset ubiquitin structure shows the Leu15 sidechain in the presence of water molecules during the second-highest exposure at 25.5 Å^2. Line roughly shows average SASA value of 2.9 Å^2. SASA factors: Average = 0.3, Max = 2.4. 109
Figure 5-4: SASA factors for the ubiquitin TITLE peptide, which illustrates a relatively level playing field due to its absence of high-reactivity residues..... 112

List of Tables

Table 2-1: Mass Shifts of Oxidative Modifications Searched with Two Hybrid Software Packages.....28
Table 3-1: Oxidative Modifications for MS/MS Searches and Protein-Specific Identifications.....45
Table 3-2: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Ubiquitin.....47
Table 3-3: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Apomyoglobin.....49
Table 3-4: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Lysozyme 57
Table 4-1: Physical features of model proteins subjected to electrochemical oxidation.....69

List of Attachments

File 1 Peptide Spectral Counts Mined from Mass Spectral Data.....Appendix-Tables.xls

Table A-1 *Oxidative Modifications Identified from Electrochemically-Oxidized Ubiquitin (Unabridged) – Chapter 3 results*

Table A-2 *Oxidative Modifications Identified from Electrochemically-Oxidized Apomyoglobin (Unabridged) – Chapter 3 results*

Table A-3 *Oxidative Modifications Identified from Electrochemically-Oxidized Lysozyme (Unabridged) – Chapter 3 results*

Table A-4 *Oxidative Modifications Identified from Electrochemically-Oxidized Ubiquitin by Two Software Searches – Chapter 4 results*

Table A-5 *Oxidative Modifications Identified from Electrochemically-Oxidized Lysozyme by InsPecT Searches – Chapter 4 results*

Table A-6 *Oxidative Modifications Identified from Electrochemically-Oxidized Ubiquitin by Two Software Searches – Chapter 5 results*

Table A-7 *Oxidative Modifications Identified from Electrochemically-Oxidized Lysozyme by Two Software Searches – Chapter 5 results*

Table A-8 *Oxidative Modifications Identified from Electrochemically-Oxidized Apomyoglobin by Two Software Searches – Chapter 5 results*

List of Symbols and Abbreviations

ACN	Acetonitrile
AU	Absorbance units
BDD	Boron-doped diamond
CD	Circular dichroism
CID	Collision-induced dissociation
CV	Cyclic voltammetry
DE	Dynamic exclusion
ESI	Electrospray ionization
FA	Formic acid
FTICR	Fourier transform ion cyclotron resonance
HDX	Hydrogen-deuterium exchange
LC	Liquid chromatography
LTQ	Linear trapping quadrupole
MD	Molecular dynamics
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multidimensional Protein Identification Technology
NMR	Nuclear magnetic resonance
OCV	Open circuit voltage
PDB	Protein Data Bank
PFT	Porous flow-through
PTM	Post-translational modification
RP	Reversed-phase
SCX	Strong cation exchange
SASA	Solvent-accessible surface area
UV	Ultraviolet
XRC	X-ray crystal

1. Introduction to Oxidative Protein Surface Mapping

1.1 Current State of Protein Structure Research

Protein topology evokes function, thus spurring the development of experimental techniques to collect data at the level of atomic resolution for protein structure elucidation. Two critical high resolution techniques are responsible for producing most of the detailed protein structures available: X-ray crystallography (XRC) and nuclear magnetic resonance (NMR) spectroscopy.[1] High resolution structures (< 5 Å) are typically obtained by XRC through analysis of X-ray diffraction caused by electrons of atoms in crystallized proteins.[1] Additional detail on hydrogen (^{2}H) atoms can be determined by neutron protein crystallography, though fewer structures are generated by this method due to limited beam line accessibility and the need for larger (~1 mm^3) protein crystals.[2] Solution phase NMR spectroscopy can reveal dynamic details of single proteins or interacting complexes, though its current limitation is that their molecular masses do not exceed roughly 50 kDa.[3] Cryogenic electron microscopy is also capable of producing structural details, although like crystallography techniques, high quality crystals must be produced.[4] While very powerful, all of these techniques require milligram quantities of highly purified proteins, and factors such as conformational flexibility, post-translational modification, and oligomerization often limit the amenability of proteins to these high resolution measurements.[5] Additionally, high-energy electromagnetic or particle radiation sources are required, which reduces the general accessibility for a typical laboratory researcher.

In addition to the aforementioned high resolution techniques, a number of lower (> 10 Å) resolution techniques can provide important structural information. These include small angle scattering (SAS) technologies employing either X-rays (SAXS) or neutrons (SANS), with the latter requiring deuterated proteins for analysis and both utilizing radiation sources.[6] More accessible techniques can measure differential exposure to solvent among conformational states, including limited proteolysis [7, 8] and hydrogen-deuterium exchange (HDX).[9] A few spectroscopic techniques for monitoring a change in protein conformation within a solution ensemble include fluorescence spectroscopy, which can detect aromatic (Trp) residues undergoing a change in their chemical environment,[10, 11] fluorescence resonance energy transfer (FRET), which gauges distance-dependent intensity changes between two spectrally overlapping fluorophores,[12] and circular dichroism (CD), which measures secondary structure content.[13, 14] All of these methods provide different types of information useful for evaluating protein structure.

1.2 Computational Tools for Advancing Experimental Protein Structure Research

While the atomic level resolution available through crystallographic snapshots or dynamic NMR models are generally regarded as the gold-standards for characterizing protein structures, computational approaches are being explored for proteins less amenable to experimental techniques.[15] To keep pace with the rapid production of proteomic data enabled by genome sequencing, predictive modeling efforts within the realm of computational biology are aimed to fill the gap between these high throughput datasets and the steadily growing Protein Data Bank (PDB).[16] Homology models can be built for unknown structures that share at least

30% sequence similarity with a resolved protein structure,[17] and protein threading utilizes sequence-based fold similarity to parse domains from existing structures in constructing a predictive model.[18] Additionally, *de novo* structural predictions are gaining significant traction, as evidenced by the results of the Critical Assessment of Structure Prediction (CASP) competition.[19] This contest invites researchers to submit structural predictions for a newly resolved protein structure that has not yet been publicly revealed, with predictions from the community showing increasing accuracy compared with the target structure. While the chosen method will be determined by the level of existing structural knowledge, the resulting predictive model(s) could be evaluated and likely refined by experimental data that reflects the accuracy of each prediction. One intriguing approach involves probing the solvent accessible surface area (SASA) of proteins with a low-selectivity reagent that fosters amino acid level resolution of exposure to solvent.

Concurrent with advances in computational modeling for protein structure prediction, there is a need for experimental approaches that provide data for evaluating the validity of structural models.[20] As discussed previously, a variety of prediction methods can be utilized to generate three-dimensional model structures *in silico* from a protein's primary amino acid sequence.[18] Among potentially thousands of predictive models, comparable α-carbon backbones differ in the manner of sidechain packing within the hydrophobic core.[21] The significance of this rotamer variation is that the total exposed surface area of a predicted protein structure may be influenced or even controlled by a different subset of residues among computational models. For example, a study performed in-house involved calculation of SASA using the probe radius of water (1.4 Å) to evaluate the solvent exposure of individual amino

acids across a series of models derived for a lysozyme NMR study (PDB 1E8L). This exercise revealed variable SASA values (see Appendix Table A-7) for all residues along with significant fluctuations for some regions. Similarly, conformational change within a protein structure may involve a SASA shift for various amino acids. One example of this phenomenon is the pH-sensitive exposure of a hydrophobic pocket in the case of β-lactoglobulins, which has been dynamically modeled *in silico*.[22] Experimental identification of solvent accessible amino acids would be advantageous in culling candidate structures potentially representing true conformations of protein structure. Paradoxically, the labeling accuracy of such experimental techniques should ideally first be verified on structurally characterized systems subjected to simulated solution-phase conditions.

1.3 Molecular Dynamics Simulations for Evaluating Solvent Accessibility Probe Data

Static crystal structures are used as initial coordinates in molecular dynamics (MD) simulations, wherein proteins are subjected to solution forces *in silico* over a period of time to reveal some insight into their true dynamic nature.[23] Simulation software programs such as CHARMM [24] and NAMD [25] employ force fields, which involve the use of empirical energy functions to approximate forces that occur among atoms within the macromolecule over time. These force field equations include parameters to describe forces acting upon bonded atoms such as spring constants and torsion, in addition to those that can occur between non-bonded atoms over longer distances, namely electrostatic and van der Waals interactions.[26] The limitations of this technique for simulating the dynamics of model proteins used in experimental demonstrations are highly dependent on the available computing power, but systems as large as

the ribosome complex (2.64 million atoms including solvent) have been modeled on supercomputers.[27]

The use of SASA statistics from MD simulations should provide a better metric than values calculated from static structures for evaluating amino acid solvent accessibility probe results. This may be particularly evident with the use of reagents that selectively target specific amino acids and thus require simpler data mining efforts relative to low-selectivity reagents. Average and maximum SASA values calculated for individual residues on the nanosecond timeframe in these relatively simple simulations may prove to be essential tools in evaluating oxidative mapping results. When paired with techniques that provide a reliable gauge of SASA values, MD simulations could be performed using predictive computational models (*ab initio*, domain parsing, threading) to increase throughput of structural proteomics.

1.4 Mass Spectrometry Analyses of Solvent Accessibility Labeling Techniques

Mass spectrometry (MS) is a powerful experimental tool for probing higher order protein structure, in that this technique is very sensitive and can be used with protein mixtures. One approach to harness this capacity for obtaining structural information is to employ covalent labeling of native proteins followed by high resolution MS characterization of modified amino acid residues. Covalent labeling techniques place a stable mass tag on reactive, solvent accessible residues within the native protein structure that can be later identified by MS characterization of mass-shifted proteolytic peptides. A variety of labeling reagents exist which vary in terms of their amino acid specificity, rate of reactivity, timescale for exposure to achieve sufficient labeling for detection, and some offer intrinsic opportunities for enrichment prior to

analysis.[28] Danger lies in perturbation of native structure that can occur with even a single modification depending on the structural sensitivity of the site labeled.[29, 30] Therefore, ideal conditions would include a relatively small tag and a gentle exposure to achieve only a few labeling events per molecule. Additionally, a rapid exposure is critical to avoid labeling of perturbed structures that may result due to oxidation-induced unfolding.[31] While there are a number of covalent labeling reagents that target specific types of amino acids,[28] the hydroxyl radical (•OH) is much less selective, smaller than a water molecule, highly reactive, and can be generated by a variety of methods. While all of the techniques discussed below share the same reactive species, they vary significantly in nature of •OH generation, the flux of •OH achievable, the reaction timescale, and the experimental equipment required.

1.5 Oxidative Mapping of Proteins with Hydroxyl Radicals

The hydroxyl radical that wreaks havoc *in vivo* can provide structural insight when harnessed *in vitro* for controlled oxidation of target proteins. The rapidly evolving field of oxidative surface mapping aims to combine computational protein structure prediction tools with experimental MS data, with the goal of gaining insight about the dynamic SASA values experienced by native solution-phase proteins.[32, 33] Hydroxyl radicals have been used to probe the SASA in both conformational and liganded variants of single proteins,[34, 35] folding intermediates,[36] and protein complexes.[37-39] The lifetime of the highly reactive •OH molecule is on the order of nanoseconds,[40] fostering a rapid reaction with susceptible sites within proteins. The reaction of electrogenerated •OH with proteins passing over the electrode surfacedepend on a few physical parameters: the flux or concentration of •OH generated, the

mean free path of •OH, and the rate of its diffusion away from the electrode surface into the bulk solution. While the •OH concentration might be indirectly measureable by a quenching or trapping assay, the latter two parameters are likely impacted by the rapid flow rate and solution turbulence inside the electrochemical flow chamber. The electrophilic •OH targets a variety of amino acids, with the electron-rich aromatic and sulfur-containing residues being most reactive, followed by aliphatic, charged, and alcohol sidechains.[29] These •OH-mediated modifications can occur by one of two mechanisms: the primary pathway involves electrophilic addition of •OH to an aromatic ring, while the less common reaction involves hydrogen abstraction from a sidechain by •OH, leaving a radical site that is effectively oxidized by molecular oxygen (O_2).[41, 42] A variety of methods exist for generating •OH from hydrogen peroxide, including photolysis by continuous ultraviolet (UV) light[43] or pulsed UV laser [44, 45], and iron-catalyzed generation with Fenton's reagent.[46] The latter has been used with electrochemistry to mimic drug metabolism [47] and for herbicidal degradation.[48] One caveat to these methods of •OH generation is that methionine has been shown to undergo secondary oxidation irrespective of solvent accessibility in the presence of peroxides.[46] Other methods include radiolysis of water by controlled exposure to X-rays [49, 50] or gamma rays [35, 51, 52] in the presence of catalase to decompose peroxide and reportedly reclaim the utility of oxidized methionine. Finally, the electrochemistry critical to electrospray ionization has been utilized at higher voltages and in the presence of an oxygen sheath gas to induce indirect oxidation of proteins as a "radical probe" of protein structure.[53-56] Oxidative surface mapping continues to emerge as a widely desirable technique for protein inquiry, and could benefit from an approach that improves accessibility to the average researcher without excessive cost or danger.

1.6 Motivation for Development of Electrochemical Protein Oxidation Method

Electrochemistry has been explored for decades as a tool for protein interrogation, since redox reactions can be achieved under physiologically relevant conditions that promote native structure. Early electrochemical studies on proteins confirmed the electroactive nature of cysteine, tyrosine, tryptophan, methionine, and histidine residues oxidized directly by electron transfer reactions at graphite electrodes,[57, 58] but were not focused on interrogation of higher order protein structure. Metal ions coordinated in the prosthetic groups of redox proteins, such as hemoglobin, catalase, or myriad cytochromes, are susceptible to direct oxidation.[59, 60] Metalloproteins can be interrogated for the location of their metal binding pockets by metal-catalyzed cleavage of peroxides to yield irreversible oxidation of nearby amino acid residues.[61] Similarly, electrochemical generation of reactive oxygen species generated during solution electrolysis can modify amino acids indirectly. For example, peroxynitrite can selectively nitrate tyrosine residues,[62-64] while •OH can damage a wide range of biomolecules.[65, 66] As mentioned previously, •OH has the capacity to attack proteins by addition to aromatic ring systems or abstraction of hydrogens from aliphatic sidechains.[41, 42] Progression down either reaction pathway generally results in a discrete mass shift due to irreversible oxidation. While electroactive residues can be oxidized both by direct (heterogeneous) electron transfer and indirectly mediated mechanisms, non-electroactive amino acids are presumably limited to indirect oxidation.

Electrochemical protein oxidation is accomplished exclusively via direct electron transfer with most electrodes, but diamond electrodes are gaining popularity due to their capacity for generating •OH from the oxidation of water.[67] Other advantages of these electrodes include a

wide potential range, with an inert surface and markedly reduced adsorption profile relative to other electrodes.[68] Production of •OH is facilitated by the semiconductor qualities obtained through doping of diamond film with a trivalent atom such as boron to increase positive charge carrying capacity.[68] Application of a sufficiently anodic voltage on a boron-doped diamond (BDD) working electrode causes electrons to be pulled from water to produce hydroxyl radicals by the following mechanism: $BDD + H_2O \rightarrow BDD(\bullet OH) + e^- + H^+$.[67] Therefore, the only reagent necessary for generating •OH with the commercially available BDD electrode is an inert electrolyte to facilitate the flow of current and achievement of electrolytic voltages. While the direct oxidation route may not be eliminated completely, the diffusion of •OH from the electrode surface into solution during electrolysis appears to be the primary oxidation pathway.[69] Interestingly, the potassium sulfate electrolyte originally chosen for its ionic strength, low ultraviolet absorbance, inorganic character and thus •OH inertness, may actually prove to be electrochemically reactive. If this latter aspect is true, then the electrolyte could contribute oxidizing species to the oxidant pool via its direct oxidation to a sulfate radical anion that can similarly oxidize amino acids.[70, 71] While studies on small molecules have yielded important results, determining the sites of oxidation within oxidized proteins is proving to be a more difficult task to unravel.

1.7 Identification of Oxidized Sites from Mass Spectral Data

Oxidized peptides can be formed in a staggering variety of combinations due to the sheer number of mass shifts possible for any given peptide. The hydroxyl radical is a low selectivity probe that targets a wide breadth of amino acids, some with up to five known oxidized

variants.[72] Therefore, even a single short peptide could theoretically become a complex mixture of peptides distributed from lightly to heavily oxidized if provided sufficient exposure to oxidizing conditions. While the mass spectrometer has the ability to measure each unique combination if given the opportunity, those spectra can only be matched with their oxidized peptide identities if they can be accurately interpreted by software capable of exploring the possible variants, which is the focus of Chapter 5.

Another confounding factor is that uniquely modified peptides may be present in remarkably low abundance, so dynamic range is a critical challenge to be surmounted. Most oxidative mapping experiments are followed by reversed-phase separation of tryptic peptides since oxidation (when it involves incorporation of an oxygen atom) effectively decreases hydrophobicity, thus predictably reducing peptide retention time [32] as well as ionization efficiency.[73] Because oxidation can also result in the loss of some charged functional groups (no incorporation of oxygen), net charge was exploited for increasing the overall peptide separation capacity. The use of an abbreviated MudPIT [74, 75] chromatographic strategy (detailed in Chapter 2) was undertaken with the goal of increasing dynamic range to thoroughly interrogate oxidized samples. Additionally, the extended analytical timeframe increases overall sampling and thus the potential for spectral redundancy to increase confidence in identifications.

1.8 Scope of Dissertation Research Project

In 2005, a new BDD electrode was being evaluated for a completely separate research project in our Mass Spectrometry group at Oak Ridge National Laboratory, namely electrochemical oxidation of small molecules (as performed by then post-doctoral researcher

Vilmos Kertesz). Based on the electrochemical attributes described above, the BDD electrode appeared to be an attractive candidate for the development of a novel oxidative protein surface mapping approach that negated the use of harsh reagents or hazardous equipment. While our initial work on protein surface mapping with hydroxyl radical chemistry had been focused on photolysis of hydrogen peroxide, we recognized the potential of electrochemical protein oxidation followed by MS detection at the intact protein and peptide levels. The results of this research formed the basis of this dissertation, and are presented here. Chapter 3 contains the demonstration of an electrochemical protein oxidation protocol published in 2008,[76] with a reporting of the superior BDD performance relative to a porous flow through (PFT) graphite electrode. The oxidation of three model proteins in both non-native and physiologically relevant conditions was reported along with recovery issues posed primarily by the PFT electrode. Chapter 4 focuses on improvements to the electrochemical oxidation protocol that facilitated more rapid exposure of the same model proteins to oxidizing conditions, with the goal of minimizing protein-electrode interactions that would allow direct oxidation and possibly surface adsorption. In these two chapters, InsPecT search software [77] was used to mine MS data for more than 40 possible oxidative modifications,[72] but over time, realization of the difficulties in identifying oxidation sites spurred the testing ByOnic [78] software with the goal of obtaining a set of peptide match results verified by two different software searches. An overview of how each software works to identify oxidized peptides is presented in Chapter 2, followed by a comparison of electrochemically oxidized sites identified by InsPecT and ByOnic software packages in Chapter 5.

Preliminary results showed sites oxidized in native conditions were in reasonable agreement with SASA values calculated from static X-ray or NMR structures. However, oxidation that had not been induced electrochemically became evident in more sensitive measurements of control samples, which presented a formidable challenge to interpretation of the search results. Nevertheless, solution-phase proteins are not static, but rather an ensemble of highly dynamic macromolecules that sample a wide variety of subtly different conformations. Comparison of covalently mass-tagged sites detected by MS to SASA values calculated from static structures could generate misleading conclusions, either about the accuracy of the structure or the effectiveness of the labeling strategy employed. Therefore, Chapter 5 includes some discussion of oxidative search results compared with SASA values obtained through MD simulations. Regardless of the final site-specific interpretation for these particular experimental conditions, development of this methodology has laid the foundation for future improvements that could feasibly result in a high-throughput pipeline for structural proteomics. Furthermore, this work imparts valuable technical detail about electrochemical protein oxidation to a research community that has been interested in such topics for more than 50 years.

2. Experimental Platform for Oxidative Mapping with Electrochemistry

The experimental protocols described below were devised with the goal of oxidizing model proteins with •OH generated by electrochemistry. This novel approach for oxidative mapping of proteins was developed to increase the accessibility and throughput of this technique. Adjustment of experimental parameters (current, flow rate, electrolyte and protein concentrations) allowed control over the resulting oxidation level as measured by intact protein MS. The methods described below were chosen to enhance an experimental pipeline for the rapid oxidation of proteins, followed by sensitive detection of oxidized peptides to maximize identification of solvent accessible amino acids within native solution-phase protein structures.

2.1 Materials

All proteins, peptides, acetic acid (AA), dithiothreitol (DTT), and ammonium acetate were obtained from Sigma-Aldrich (St. Louis, MO) at the highest purity available. All solvents (water, acetonitrile, methanol) were purchased from Burdick & Jackson (Muskegee, MI). Formic acid (FA) and potassium sulfate were obtained from EMD Chemicals (Darmstadt, Germany). Sequencing-grade trypsin from Promega (Madison, WI) was used for proteolytic digestions. Proteins and peptides were extracted from salt solutions with C_4 or C_{18} reverse-phase standard bed ZipTips, respectively, acquired from Millipore (Billerica, MA).

2.1.1 Secondary Oxidation Quenchers

Small molecules and redox proteins have been shown to help circumvent problematic oxidation events for established surface mapping techniques. The over-exposure to •OH after

the initial oxidation may cause partially unfolded conformations to be labeled during the experiment. Moderately reactive radical scavengers like free glutamate have been used to better control the oxidant exposure.[44] While seemingly sensible, this measure was not taken to avoid possible complications due to electrochemical adsorption. Also, •OH may self-quench, coalescing into hydrogen peroxide that can oxidize methionine irrespective of SASA value. Inclusion of catalase tetramer in reaction solutions to decompose peroxide has been shown to recoup methionine utility in oxidative studies.[35] This factor was also omitted not only to minimize a potential adsorption and/or oxidant flux problem, but also because catalase has a metal prosthetic group that could be directly oxidized. Peptidyl hydroperoxides formed during oxidation of aliphatic sidechains can also oxidize methionine, so methionine amide can be used to dodge this source of uncontrolled protein oxidation.[79] Again, concern over adsorption trumped these issues until the technique was established. Of more concern was the background oxidation discussed in Chapter 4, so the use of DTT as an antioxidant was tested in some control experiments to test whether this phenomenon might be due to unexpected oxidant production during the digestion process. Nitrogen gas was also used in these controls to displace oxygen in the head space atop the reaction solution.

2.2 Electrochemistry

Experiments were initially conducted with two different types of electrodes that utilized two different electrochemical cells. Later studies employed only the BDD electrode in a flow-by electrochemical cell, with minor changes to the experimental setup involving a new potentiostat and use of a UV detector. Electrochemical details are contained in the next five sections.

2.2.1 Electrode Descriptions

The porous flow-through (PFT) graphite electrode was housed inside a three-electrode electrochemical cell (Model 5030, ESA, Inc., Chelmsford, MA). As described previously by Van Berkel *et al*,[80] the PFT working electrode was composed of porous graphitic carbon 0.508 mm diameter x 0.36 mm thick (40% total porosity, 99% open porosity with a nominal pore size of 0.8 μm, and 29 nL electrode volume); the inner dimensions of the electrode assembly provided a nominal swept volume of 100 nL and a calculated surface area of 170 mm^2.

The boron-doped diamond (BDD) electrode is a flat disc with 220 mm^2 surface area, of which roughly 30 mm^2 was accessible through an oval window in a thin Teflon® gasket adjacent to the electrode surface (Fig. 2-1). The BDD working electrode was seated along with stainless steel counter/auxiliary and silver chloride (Ag/AgCl) reference electrodes inside a three-electrode electrochemical flow-by cell (Model 5041, ESA, Inc.) shown in Fig. 2-1A. Depending on which gasket thickness (25 or 50 μm) was used to separate the working and counter electrodes, the contact volume inside the flow cell was roughly 1-2 microliters.

2.2.2 Operation of Electrochemical Cell for Protein Oxidation

A CHI900 scanning electrochemical microscope (CH Instruments, Austin, TX) or a WaveNow potentiostat (Pine Instruments, Raleigh, NC) was used for anodic operation of the working electrode. Early experiments utilized an experimental mode of controlled voltage, or chronoamperometry, which measures the current (in microamperes, μA) applied over time to maintain a positive voltage (in volts, V). Later experiments utilized chronopotentiometry to measure the fluctuations in potential (voltage) that occurred during the sustained application of a

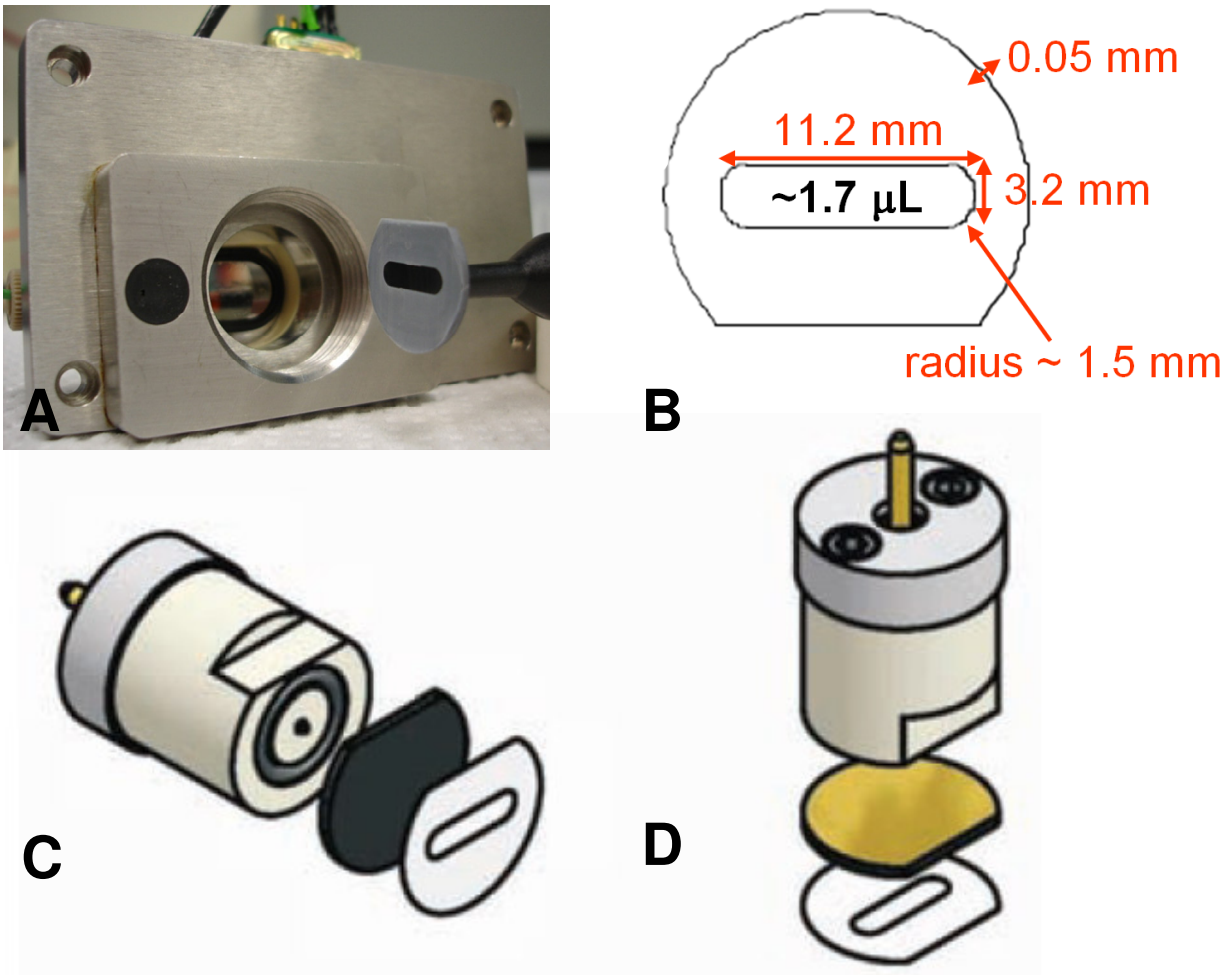


Figure 2-1: Boron-doped diamond working electrode configuration inside electrochemical flow-by cell. (A) Interior of electrochemical flow cell, with the BDD working electrode surface showing through the gasket window. (B) Schematic of the 50 µm gasket separating working and counter electrodes, with dimensions in red font used for calculating the flow chamber volume noted in black. (C and D) BDD contact assembly, showing small pin inside O-ring that contacts gold surface on back side of BDD electrode disc. The longer pin at the end of the assembly is connected to the potentiostat during use of the electrochemical cell (taken from ESA, Inc. materials)

particular current. The reason for this change is that •OH production is presumably governed by the current-dependent Nernst equation,[67] and that voltage readings can rise concomitant with insulation of the electrode surface for a set current. Therefore, maintaining a set voltage could result in a reduced •OH production over time.

2.2.3 Cyclic Voltammetry Experiments

Cyclic voltammetry (CV) is commonly used to query a particular analyte for the electrochemical potentials at which it becomes oxidized or reduced. In these experiments, the range of potentials from +1.0 V to -1.0 V was scanned at a rate of 100 mV/s, with three cycles used in each experiment. The resulting data plots current versus voltage, allowing visualization of the current level required to hit the set voltage endpoints. If a detectable change of the electrode surface had occurred during a protein oxidation experiment, comparison of the CV plots resulting from scans performed before (the pre-CV) and after (the post-CV) the protein experiment could provide two pieces of critical information. The first was whether the current level at the voltage endpoints had changed in the first cycle, and the second was whether the change was resolved by subsequent cycles. The first scenario suggests some level of protein adsorption had taken place. The second situation provided a clue about whether the adsorption was reversible or irreversible.

2.2.4 Open Circuit Potential Readings

The open circuit voltage (OCV) reading can be elicited from the electrochemical assembly once the potentiostat has been properly connected to the working, counter, and reference electrodes. The OCV reading, also known as the resting potential, measures the resting

voltage between the anode and cathode in the circuit without an electrical load from the potentiostat. This value can provide feedback on the resistance (or inversely the conductivity) between the electrodes. This resistance is due to a number of factors, including the electrode surface areas exposed, the distance between the electrodes, and the ionic strength of the solution (also dielectric constant of the solvent) separating them. An OCV value close to zero would imply low resistance and thus high conductivity (possibility of "short-circuiting" if electrode is activated), with the opposite true for large absolute values. An increase in the absolute value may indicate a degree of insulation, or a value closer to zero suggests presence of an electrolyte species contamination.

2.2.5 Instrument Configuration

For experiments conducted off-line from MS, a protein solution was loaded into a 250 µL or 1 mL loop connected to a six-port injection valve (Model 7125, Rheodyne, Rohnert Park, CA) placed in-line between a Switchos HPLC pump (Dionex, Sunnyvale, CA) and the flow-by electrochemical cell. All components were connected by 0.5 µm polyetheretherketone (PEEK) tubing (Upchurch Scientific, Oak Harbor, WA). Some experiments involved use of an Ultimate (Dionex, Sunnyvale, CA) UV detector placed after the electrochemical cell to monitor absorption at 205 nm and 280 nm wavelengths, and Xcalibur software was used to simultaneously record the signal from each wavelength. The 190-nL capacity of the ULT-UZ-M10 flow cell accommodated flow rates of 10-100 µL/min.

On-line experiments were performed using a similar configuration to that listed above, except that a stainless steel union was placed in-line between the electrochemical cell and the

floating electrospray ionization (ESI) source. This permitted grounding of the ESI source housing to prevent current feedback to the electrochemical electronics through the conductive solution. Conductivities for electrolyte solutions were calculated by an online tool available at http://www.lenntech.com/calculators.htm.

2.3 Liquid Chromatography and Mass Spectrometry

Liquid chromatography (LC) refers to the separation of solution-phase constituents dissolved in a mobile based on their differential affinity for a stationary phase material. The most common type of LC is reversed-phase (RP), wherein saturated hydrocarbon chains are tethered to resin beads that are packed inside a column to facilitate separation of components by their hydrophobic character. The high pressure utilized by pumps in high-performance LC (HPLC) is critical for performing separations "on-line" with mass spectrometry (MS) because it counters the high back-pressure that can accumulate with tightly packed columns. This direct coupling of LC-MS instrumentation is highly compatible because the LC mobile phase contains elements necessary for both electrospray ionization (ESI) (volatile solvents, acidic electrolyte) and MS detection (preformed ions in primarily salt-free solution). This technical union allows a more sensitive measurement relative to the simultaneous assessment that occurs with direct infusion, because MS detection is limited by dynamic range. This means that the ability to detect the most miniscule species is governed by its relation the most abundant ion signal also present at that point in time. Therefore, the advantage of LC is reduction in the peptide complexity at any given point over the LC timeframe, which effectively increases the dynamic range to allow detection of lower abundance species.[81]

Typically, peptide LC is performed on-line with a high-throughput instrument like a linear quadrupole [82] or three-dimensional (Paul) ion trap.[83] The high duty cycle of these instruments fosters a rich MS data set, usually with options for data-dependent acquisition of tandem mass spectra (MS/MS). These spectra provide peptide sequence information via collision-induced dissociation (CID) of an isolated parent ion within a narrow mass-to-charge (m/z) window. While only some of the peptide LC-MS experiments utilized an Orbitrap for scans of the full mass range (MS1) normally from 400 to 1700 m/z, all of them employed a linear trapping quadrupole (LTQ) for the MS/MS resulting from fragmentation of an isolated parent ion detected in the MS1. The Orbitrap mode for MS1 scanning requires more time (~1 sec) than the LTQ (~200 ms), with the advantage of exceptional mass accuracy (< 5 ppm) to foster delineation between species that would appear isobaric (same mass) on a lower performance instrument.[84] Another positive feature is its high resolution (30,000 full-width at half-maximum (30K FWHM) or better) to resolve peptides and notably proteins with higher charge states. In contrast, the resolution of the LTQ is usually limited such that peptides of charge 2+ and 3+ are not clearly distinguished, so many search software algorithms routinely consider both charge states when computing a neutral parent ion mass. This capacity for resolving highly charged biomolecules is shared by the Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (Fig. 2-2), whose typically lower throughput is offset by its advantages of excellent mass accuracy (< 5 ppm), large dynamic range (10,000:1), high resolving power (often more than 100K FWHM), and wide mass-to-charge range (from <100 to 2500 m/z).[85]

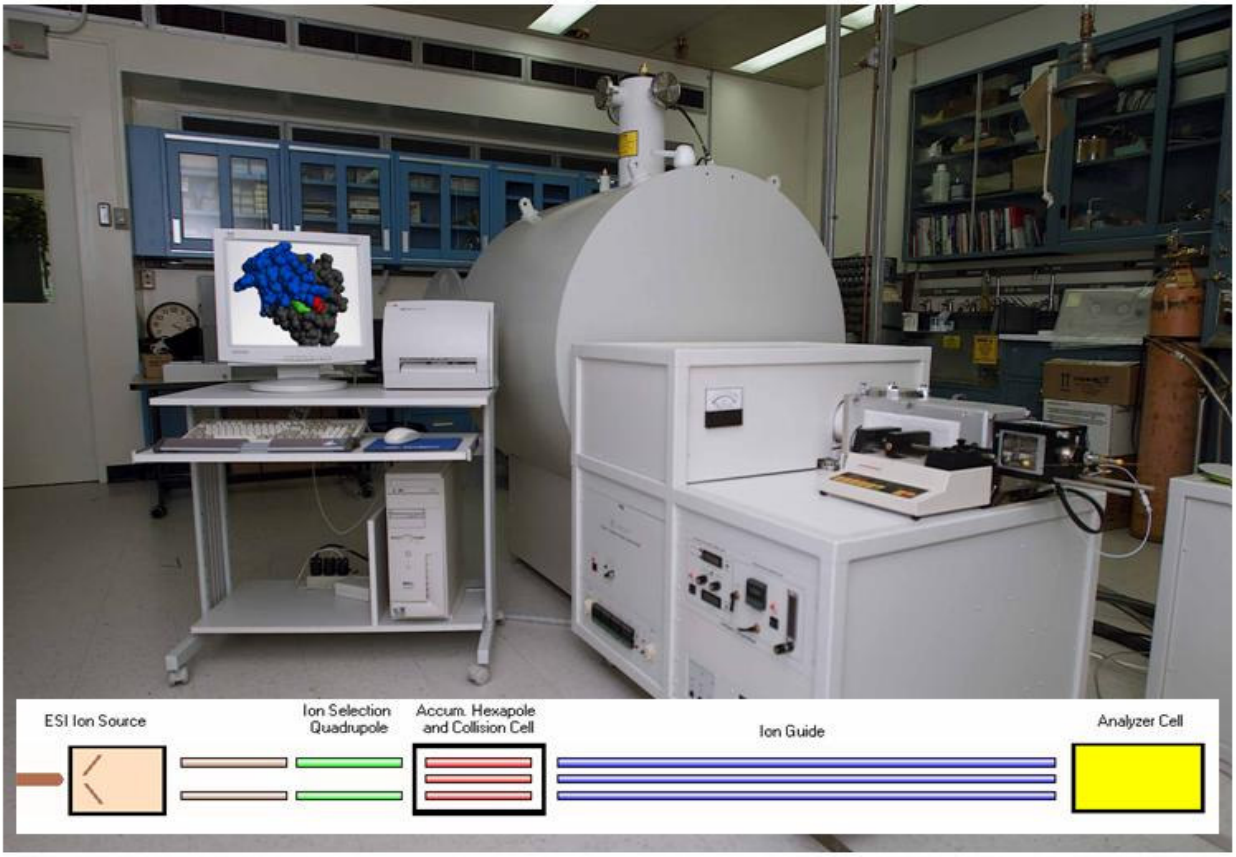


Figure 2-2: Fourier transform ion cyclotron resonance (FTICR) mass spectrometer outfitted with a 9.4-Tesla actively shielded magnet and equipped with Z-spray electrospray ionization (ESI) source.
The inset schematic depicts the transmission path for ions, which are preformed in acidic electrospray solvent and injected for spray desolvation at the ESI ion source. Ions are accumulated in the hexapole for a user-defined duration, then shuttled to the analyzer cell inside the magnet which, along with an activation pulse, cause the ion packets to coherently cycle at a frequencies dependent on their mass-to-charge values.

2.3.1 Intact Protein Mass Spectrometry

Intact proteins were measured by MS to evaluate the extent of oxidation resulting from electrochemical oxidation. After extraction from the non-MS-friendly electrolyte solutions with a C_4 standard bed ZipTip (Millipore, Billerica, MA), eluted protein solutions were then prepared for direct infusion in ACN/H_2O/FA (50/50/0.1, v/v/v). Electrospray-ready samples were infused at 2.5 μL/min by syringe pump (Harvard Apparatus, Holliston, MA) through PEEK tubing directly into the ESI source of a FTICR mass spectrometer (Varian Inc., Palo Alto, CA) outfitted with a 9.4 Tesla actively-shielded superconducting magnet (Cryomagnetics Inc., Oak Ridge, TN). Time domain ion signals were acquired using 1024K data points, averaged over 5-50 scans, and converted by fast Fourier transform (FFT) to the mass-to-charge (m/z) domain with application of a Hann apodization window using Omega software (Varian, Palo Alto, CA). On-line electrochemical experiments were performed in chromatogram mode by acquiring single full-scan 1024K transients without signal averaging.

Oxidation has been shown to occur at the tip of corroded ESI sources resulting in corona discharge (I_{ESI} > 1 μA).[86] However, this process was not observed (I_{ESI} < 50 nA). Control experiments were conducted by passing the protein through the inactive electrochemical cell and directly into the ESI source of the FTICR-MS. The resulting mass spectra did not reveal oxidation of the intact native protein mass, but some of the commercially available proteins did show evidence of low-abundance satellite peaks, which is discussed in Section 4.7.

2.3.2 Proteolytic Digestion

In order to analyze electrochemically oxidized proteins at the level of individual amino acids, it was necessary to cleave these large biomolecules into more tractable units measurable by MS. An aliquot of extracted protein was taken for measurement of oxidation level by direct infusion MS, while the remaining eluant was subjected to proteolysis in a total volume of 250 μL. An aqueous-organic digestion protocol [87] was used to digest small quantities of protein rapidly (one hour, with additional time for disulfide reduction) with the benefit of loading peptides for LC-MS immediately after proteolysis. This transition did not require a peptide extraction step after the proteolysis because sticky chaotropic salts (guanidine, urea) that are used to denature protein in other protocols were lacking from the reaction solution. However, the low concentration of trypsin (~250 ng/100 μL) was still present, which may have contributed to peptide bias (discussed in Chapters 4 and 5) due to occupation of binding sites in the chromatographic resins used. The extracted protein was calculated roughly based on maximum extraction values (~3-5 μg bound each time) reported by the manufacturer, since absorbance at 280 nm was not useful for ubiquitin (no Trp) or oxidized proteins in general since Trp oxidation lowers its UV absorbance.

For the proteolytic digestions, 50 μL of pure acetonitrile (ACN) containing eluted protein from 3-4 ZipTip extractions was added to 20 μL of 50 mM Tris-HCl/10 mM $CaCl_2$. Next, modified trypsin reconstituted in the Tris-$CaCl_2$ solution was added in roughly 1:20 weight ratio to extracted protein, followed by addition of ACN to 60-80% of the total 250 μL volume. Digestions proceeded at 37 ºC with gentle mixing at 300 rpm in a Thermomixer® R (Eppendorf

AG, Hamburg, Germany), followed by a one hour reduction of disulfide-containing proteins in 10 mM dithiothreitol (DTT) at 37ºC. Trypsin was inactivated by addition of 2 µL formic acid before removing the organic solvent by lyophilization (Savant SpeedVac, Thermo Scientific, Waltham, MA). Samples were dried from 250 µL to 20 µL, and then brought to 100 µL with HPLC-grade water to prepare them for binding C_{18} reversed-phase material. Aliquots were allotted as follows: 10 µL for C_{18} peptide extraction followed by FTICR-MS analysis, and 90 µL available for up to three LC-MS/MS runs loading 25-µL for each experiment.

2.3.3 Peptide LC-MS/MS Methods

Peptide level MS analysis was necessary to identify oxidized peptides via mass-shifted parent ions, along with sequence information gained by the acquisition of MS/MS spectra. However, any MS measurement is limited by the dynamic range of abundances for components of the solution being measured, so sensitive detection of unique and/or less abundant constituents is enhanced by reducing the complexity at any given point in the analytical timeframe. Peptides were separated prior to entering the mass spectrometer by either one-dimensional (1D) C_{18} reversed-phase (RP) or two-dimensional (2D) "mini-MudPIT" strategy. MudPIT [75] stands for **Mu**lti**d**imensional **P**rotein **I**dentification **T**echnology, which incorporates a split-phase back column with strong cation exchange (SCX) resin flanked by RP resin in both the back column and the analytical front column. Proteolytic peptides were loaded onto a 100 µm inner diameter (ID) fused silica back column packed with appropriate resins. The back column was then joined by a union filter assembly (UpChurch Scientific, Oak Harbor, WA) to the front column, which

consisted of a 100 μm ID fused silica PicoTip emitter (New Objective, Woburn, MA) packed with 15 cm Aqua C_{18} reversed-phase resin (Phenomenex, Torrance, CA).

For 1D analysis, high pressure liquid chromatography (HPLC) was employed over a 60- or 120-min gradient from 100% solvent A (ACN/H_2O/FA (5/95/0.1, v/v/v) to 100% solvent B (ACN/H_2O/FA (70/30/0.1, v/v/v) with later experiments run over 120-min. Eluting peptides were injected on-line by nano-ESI into the heated capillary of either a LTQ XL or LTQ-Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA). Peptide parent masses were measured in the Orbitrap at 30K resolution, while sequence information was obtained by MS/MS in data-dependent mode in the LTQ with dynamic exclusion enabled to add parent ion masses to an exclusion list (for 30 sec duration) if they were sampled twice within a 30 sec span. For 2D analysis, four gradients were run for a total of 5.5 hours instrument time. The last three gradients began with a short pulse of ammonium acetate (25%, 50%, 100% of 500 mM NH_4OAc), followed by a wash with 100% solvent A before starting the reversed-phase gradient.

2.4 Informatics and Data Mining

Interpretation of mass spectral fragmentation patterns by various search software packages tends to fall somewhere on a continuum (Fig. 2-3) from database protein sequence oriented to *de novo* extrapolation of peptide sequence directly from spectra without reliance on a database. Hybrid approaches incorporate features of each. Both InsPecT [77] and ByOnic [78] use hybrid strategies to accomplish database filtration, or whittling a large list of possible peptide matches down to those candidate sequences that are most likely to match the spectrum.

2.4.1 DBDigger Software

DBDigger [88] is similar to SEQUEST [89] in its strict reliance on a database for spectral interpretation. Peptide tandem mass spectra (MS/MS) were searched using DBDigger [88] software equipped with the MASPIC [90] scoring algorithm to evaluate sequence coverage of the normal, unoxidized protein against a forward-reversed concatenated database. This database contained the aforementioned model proteins plus commonly sampled contaminant proteins such as keratins from human skin, as well as bogus proteins that should not be present, along with the reversal of every database entry for a total of 212 sequences. This software was routinely used for non-PTM proteome searches in our lab, but was not designed to handle the large number of possible mass shifts that can occur due to oxidative modification. For this reason, other software packages created for tackling this type of challenge were employed.

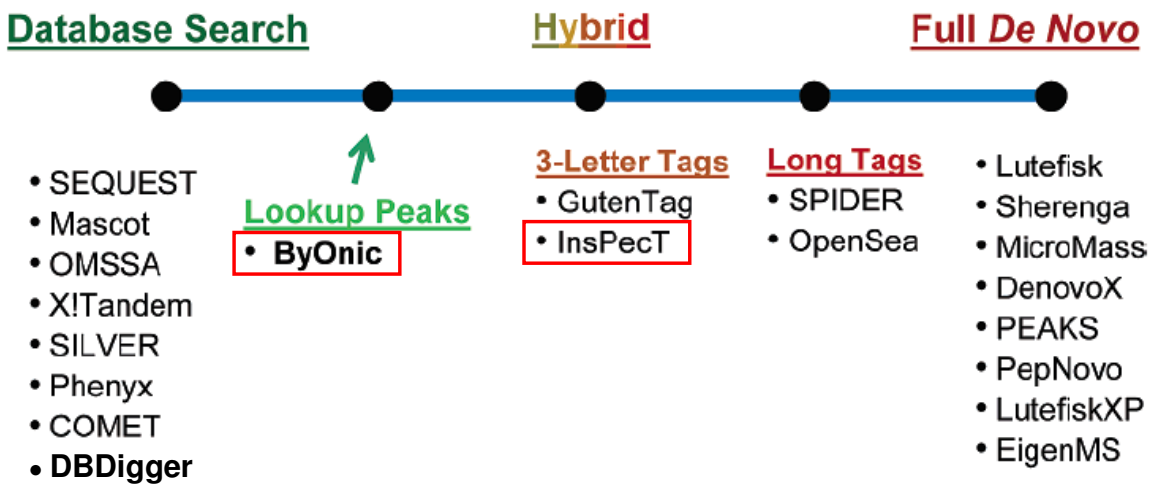


Figure 2-3: Software currently available for searching MS/MS data is displayed on a continuum from total reliance on a database for spectral interpretation to pure *de novo* sequencing that does not access a database. Taken from reference [78] and modified to include DBDigger [88] software.

2.4.2 InsPecT Software

Mass spectral data were initially searched using InsPecT [77] software for more than forty oxidative modifications (Table 2-1). InsPecT does not explicitly limit any type or number of post-translational modification (PTM) mass shifts to be considered, which provides advantages for detecting oxidation-induced modifications. Search input parameters for reported identifications were limited to a single modification per peptide to prohibit spurious matches of multiple mass shifts on a single residue and reduce overall false positives. Searches of MS/MS data were initially conducted against a subdatabase containing six proteins (horse myoglobin, chicken lysozyme, porcine trypsin, bovine ubiquitin, bovine β-lactoglobulin A, and bovine α-hemoglobin), along with their shuffled sequences. Oxidized peptide matches were filtered by p-value ≤ 0.01 and MQScore (liberal = 1.0; conservative = 3.5).

InsPecT relies on local *de novo* sequencing of fragment ion peak patterns to produce three-letter tags used that are used to filter candidate oxidized peptides from the database. While the first study reported in Chapter 3 utilized a relatively small database of 12 sequences total (6 forward plus 6 shuffled) for InsPecT searches, the software curators recommended using a database at least one megabyte in size to reduce false positives among candidate peptides. This newer database contained more than 500 forward sequences, which comprised the model and contaminant proteins along with bogus microbial proteins that should not be present, and was concatenated with the same number of bogus reverse/shuffled sequences. The scores associated with matches to bogus peptide sequences provide a mechanism for setting a filter level that should reduce false positives in forward matches. However, when all oxidative modifications (Table 2-1) are considered on this larger database, the number of possible oxidized

Table 2-1: Mass Shifts of Oxidative Modifications Searched with Two Hybrid Software Packages

Residues	Modifications (Da)
C	+31.99, +47.98, -15.98
M	+15.99, +31.99, +33.97, -32.01, -48.00
W	+3.99, +15.99, +19.99, +31.99, +47.98
F, Y	+15.99, +31.99, +47.98
H	+4.98, +15.99, -10.03, -22.03, -23.02
R	+13.98, +15.99, -43.05
I, L, V, K, Q	+13.98, +15.99
S, T	+15.99, -2.02
P	+13.98, +15.99, +31.99, -30.01
E	+13.98, +15.99, -30.01
D	+15.99, -30.01
N, A	+15.99

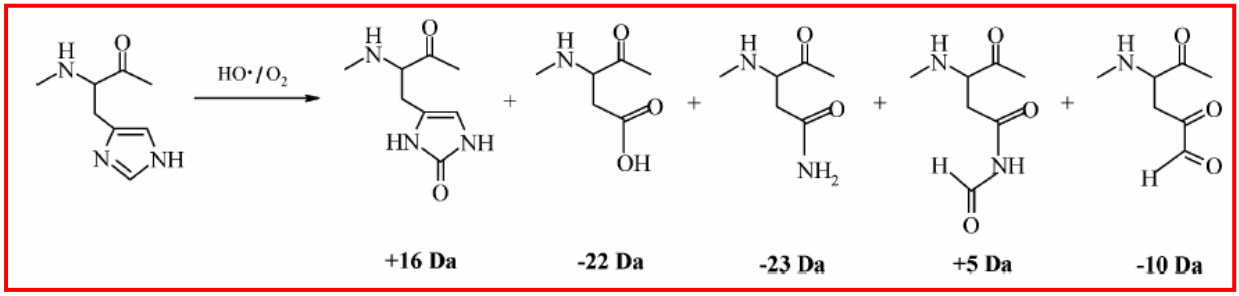


Figure 2-4: Chemical structures for the five possible histidine oxidation mass shifts highlighted in Table 2-1. Structures taken from reference [72]; additional oxidized chemical structures are available in the Appendix materials; see review article for more detailed information.

peptide variants is exceedingly vast and should be investigated by a combinatorial expert to better understand the impact on algorithmic processing during software searches. Consequently, some peptides from forward protein sequences that are clearly not present in the peptide mixture generate sufficiently high scores to pass filter levels, which provides a sense of the challenge in accurately determining false positives.

2.4.3 ByOnic Software

ByOnic [78, 91] emerged on the oxidative mapping scene in 2008 with a oxidation search module [92] that will eventually be available online at http://bio.parc.xerox.com. In contrast to InsPecT, ByOnic computes spectrally-derived "lookup peak" pairs and compares these masses against a much smaller forward-only database to similarly form a candidate list from available sequences. Both software packages can consider the liberal set of oxidized mass shifts (Table 2-1) that represent oxidation events such as hydroxylation (+15.99 or +16 Da), but only InsPecT currently allows the user flexibility in defining input parameters such as number of modifications allowed per peptide (limit of one was recommended by the software curator) or reducing the breadth of mass shifts considered. Furthermore, a ByOnic oxidative search is more liberal by virtue of placing minimal limitations on the number of modifications per peptide and considering additional mass shifts such as deamidation and sodiation. While it was expected that the two searches would reveal different results based on these aspects alone, the method used by each software to filter peptide matches placed unavoidable discrimination on the outcome.

2.5 Metrics of MS Search Software Packages

InsPecT first assigns match quality scores (MQScores) that take into account a variety of spectral characteristics (percentage of fragment ions present, peak intensities, signal-to-noise ratio, etc.) to evaluate the correctness of fit to a particular amino acid sequence. This peptide can belong either to a real forward protein sequence or a bogus reversed/shuffled sequence. Next, the resulting matches are assigned p-values based on how their composite scores compare against the distribution of scores for bogus matches. The p-value is reportedly akin to a false discovery rate (FDR),[93] such that filtering on a p-value ≤ 0.01 would correlate to roughly a 1% FDR. While this is the standard operating procedure for limited PTM searches on large proteomic datasets, it is unclear whether the same FDR would stand when considering many more PTMs on smaller datasets. A simple thought experiment suggests that consideration of many possible mass shifts on a large number of protein sequences could produce a sizable set of spectra matching to bogus modified sequences by chance. If so, this would cause sufficiently low p-values to be assigned only to the very best spectra while lesser quality yet informative spectra are eliminated by the conservative p-value threshold. Empirically, an MQScore of 3.5 or better tended to indicate a good fit with a majority of fragment ions supporting the mass shift(s), and these scores were typically associated with a p-value far below the 0.01 cutoff. However, some matches that passed the p-value filter were associated with MQScores below 3.5, so the confidence level of these matches may be questionable. Similarly, ByOnic search results were not provided the same rigorous p-value treatment, but were filtered by an empirically derived score cutoff of 380 that has been observed by the software curator to give a FDR of roughly 5%.

In contrast with the ~1% FDR for InsPecT results (with the aforementioned caveats), the ByOnic results were filtered more liberally and thus showed more plentiful oxidation matches for both the control and experimental data, a subset of which may likely be attributable to false positives.

2.6 Protein Structure Evaluation

Proteins oxidized by electrochemistry were analyzed for deviation from native structure primarily by the intact protein FTICR-MS, for the reasons described below. Additionally, site-specific oxidation events identified by both InsPecT and ByOnic were compared to SASA values calculated for either static or dynamic protein structures.

2.6.1 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra were acquired on a Jasco 810 spectrophotometer using 1-mm path length quartz cuvette. Anisotropy of a 250 µL solution ensemble was measured over the range of 190-250 nm. Upon conversion of files to ASCII format, spectral data was submitted to the online tool DichroWeb [94] for analysis of secondary structure content. Overlaid CD spectra were generated by Jasco software. Due to evidence of adsorption posed by electrochemical experiments that likely altered protein concentration, this technique was not heavily relied upon due to its concentration dependence.[13] Instead, the intact protein MS was utilized based on findings that structurally sound oxidation could be achieved and evaluated by this technique alone.[30]

2.6.2 Calculation of SASA from Crystal Structures and MD Simulation Results

Oxidized peptide matches were filtered using empirical cutoffs for appropriate software-specific parameters as discussed above (Section 2.4) and initially compared with SASA values calculated using GETAREA [95] using the 1.4 Å probe radius of water for X-ray and/or NMR structures of model proteins ubiquitin (1UBQ, 1D3Z), lysozyme (1AKI, 193L, 1E8L), and holo-myoglobin (1WLA, 1AZI) from in the Protein Data Bank.[96] Later studies employed an in-house script executed by Visual Molecular Dynamics (VMD) software to calculate SASA values for individual amino acids in either static structures or across many structures generated by molecular dynamics (MD) simulations. Three MD simulations were performed over a duration of 20 nanoseconds each in explicitly solvated conditions for the PDB structures 1UBQ (ubiquitin), 193L (lysozyme), and 1WLA (myoglobin). Values for individual residue SASA (average, standard deviation, and maximum) at each 5 picosecond timestep over the duration were extracted using another in-house script. The "SASA factor" was calculated by first taking the average SASA for individual amino acid residues across the MD timestep structures and dividing this by the SASA value calculated for free amino acids that were each subjected to MD simulations for 500 ns in isolation. This relative measurement of the surface area exposed by each residue in the protein structure relative to its unobstructed free form provides a basis for estimating the change in reactivity instigated by involvement in a protein structure. Next, the relative SASA value was multiplied by the by the intrinsic free amino acid reactivity rate that was normalized against the most reactive residue present in the protein sequence.

2.6.3 Evaluation of Oxidation Site Solvent Accessibility by Molecular Dynamics

Determination of solvent accessibility by any covalent probe approach is actually a measurement of the reactivity of the target amino acids. Therefore, interpretation of experimental results can be convoluted by factors impacting the reactivity of susceptible sites. The intrinsic rate of amino acid oxidation by the hydroxyl radical reflects the reactivity of the free amino acid,[40] whereas only a fraction of the amino acid side chain is exposed to solvent from within the protein structure. Additionally, site-specific oxidation could be impacted by interference from highly reactive neighboring residues that could act as a kinetic sink for radical interception or otherwise inhibit reaction with an exposed reactive site. In fact, leucine hydroxylation was shown to be inhibited as anion proximity increases in the primary sequence,[52] and it follows that spatial proximity of influential residue types in the secondary or tertiary structure could affect the modification of an otherwise exposed reactive site. This may help explain why many solvent exposed, highly reactive residues are found in their native unoxidized form only. This caveat is not accounted for within the SASA factor, as nearest neighbor spatial evaluation would require prior knowledge of the protein structure, which is contrary to the application goals for this methodology.

The converse is also true, that seemingly buried residues occasionally turn up oxidized. This could be due to oxidation events occurring subsequent to oxidative-induced unfolding and may pose a more significant problem for less structurally resilient proteins. If indeed the native structure is rapidly oxidized, residues which appear buried in the crystal structure may become oxidized if they experience transient solvent exposure during solution-phase dynamics within the reaction timeframe. This possibility can be explored by running MD simulations and calculating

the SASA for individual amino acids at snapshot intervals over the duration to see whether sufficient exposure might occur. The MD simulations performed for this study spanned 20 nanoseconds, which is long enough to reveal dynamic features, although not sufficient for ascertaining conformational changes that occur on the microseconds to milliseconds timeframe. Statistics calculated from these simulations were used to formulate the aforementioned SASA factor that represents the relative exposure of each amino acid weighted by the intrinsic reactivity rate for each residue. Both average (±standard deviation) and maximum SASA values were calculated for individual residues over thousands of transient structures sampled *in silico* during the MD duration. These values were used to evaluate the site-specific oxidation results from searches of peptide tandem mass spectra discussed in Section 5.4. One exception to this analysis is that methionine residues can be oxidized directly through electron transfer events catalyzed by electrode proximity, peptidyl hydroperoxides, or the presence of hydrogen peroxide formed by hydroxyl radical coalescence. Therefore, no correlation of methionine oxidation with SASA values is expected since radical scavengers were not utilized for the aforementioned reasons.

3. Development of an Experimental Technique for Probing Higher Order Protein Structure with Electrochemical Oxidation and Mass Spectrometry

This chapter is adapted from the following publication:

McClintock, C., V. Kertesz, and R.L. Hettich. *Analytical Chemistry*, 2008. **80**(9): p. 3304-3317.

This chapter presents the first demonstration of electrochemistry to generate covalent oxidative labels on intact proteins in both non-native and physiologically relevant solutions as a surface mapping probe of higher order protein structure. Two different working electrode types were tested across a range of experimental parameters including voltage, flow rate, and solution electrolyte composition to affect the extent of oxidation on intact proteins, as measured both on-line and off-line with mass spectrometry. Oxidized proteins were collected off-line for proteolytic digestion followed by LC-MS/MS analysis. Peptide MS/MS data were searched with the InsPecT scoring algorithm for forty-six oxidative mass shifts previously reported in the literature. Preliminary data showed reasonable agreement between amino acid solvent accessibility and the resulting oxidation status of these residues in aqueous buffer, while more buried residues were found to be oxidized in non-native solution. These results indicate that electrochemical oxidation using a boron-doped diamond electrode has the potential to become a useful and easily accessible tool for conducting oxidative surface mapping experiments.

3.1 Introduction

In this study, two electrodes were selected for their previously characterized ability to generate oxidation products from biologically relevant analytes. Each electrode operates with different mass transport kinetics to facilitate redox reactions at the electrode interface. Porous flow-through (PFT) graphitic carbon electrodes have already been demonstrated to oxidize peptides and intact proteins.[97, 98] Permentier *et al* developed a method for inducing on-line electrochemical protein proteolysis using PFT electrodes, wherein large peptides were generated by C-terminal cleavage at electroactive tyrosine and tryptophan residues.[98] However, MS analysis showed a greater abundance of oxidized intact protein relative to peptides, sparking our interest in its potential ability to controllably oxidize solvent-exposed amino acids for structural studies. Within the PFT electrode, solvent passes through microscopic pores to facilitate analyte mass transport to the electrode surface. However, this route could exacerbate protein adsorption, which has been reported with other graphite electrodes,[99] possibly due to the adsorptive qualities of a few amino acids using various electrodes.[100-102] The boron-doped diamond (BDD) electrode has been previously reported to suffer less adsorption of free L-cysteine relative to other carbon electrodes.[103] The BDD electrode in our configuration is a smooth disc seated in an electrochemical flow-by cell, wherein a very thin layer of solvent flows over the working electrode surface. BDD electrodes have been shown to oxidize organic material in wastewater, including benzene, phenol, and coumarin rings,[104-108] by attack of hydroxyl radicals generated at voltages sufficient for solution electrolysis.[109] Evidence has been provided that both indirect and direct pathways contribute to oxidation using a BDD electrode, with the direct electron transfer mechanism decreasing at higher potentials.[69, 103] While these studies

demonstrated the use of the BDD electrode for oxidizing small molecules, this electrode has not been evaluated with larger biomolecules, in particular proteins.

Electrochemistry is harnessed herein to place a covalent mass tag on accessible and reactive amino acids within protein structures. This is the first study that explores the capacity of electrochemistry to controllably oxidize intact model proteins for the purpose of mapping their surface residues. This approach does not require the addition of chemical reagents or access to sophisticated equipment, and can be conducted under physiologically relevant conditions. In addition to demonstrating the conditions for achieving electrochemical oxidation, some challenges of detecting oxidation products are discussed.

3.2 Results

The essential parameters to consider for electrochemical protein oxidation are applied voltage and duration, the current associated with the supplied voltage, and the solution conductivity. The last factor is governed by the supporting electrolyte (type and concentration) and the dielectric constant of solvent. Flow rate is also critical for controlling analyte residence time during which protein is exposed to the oxidizing environment. Our initial experiments were conducted on-line so that intact protein mass spectra could be monitored for evidence of oxidation. The advantage of this approach was real-time observation of the effects induced by modulating various parameters such as voltage or flow rate, so that oxidation benchmarks could be established. However, efficient oxidation required flow rates unfavorable to maintain steady spray. Also, oxidized protein could not be collected for subsequent peptide analysis. Finally, the buffers that are generally accepted to promote native structure were necessarily excluded because

they suppress MS signal. To circumvent these limitations, subsequent experiments were conducted off-line to enable a greater range of experimental variables to be tested. In this case, fractions were collected upon exit from the electrochemical cell. These fractions were desalted by solid phase extraction, followed by elution into electrospray solvent for both intact protein and peptide analysis. For the purpose of succinct discussion, the solvents used herein fall into one of two categories: native and non-native. While it is recognized that the truest "native" solution is cellular cytoplasm, a variety of physiologically relevant buffers are widely used for *in vitro* protein studies because they retain attributes necessary for proper protein folding. In contrast, the "non-native" solutions presumably do not foster native structure due to low pH values and inclusion of organic solvent.

Three model proteins with compact α-helical structures were chosen for experimentation: ubiquitin (8.6 kDa protein containing a single β-strand and lacking both cysteine and tryptophan residues), lysozyme (14.3 kDa protein tethered by four disulfide bonds), and apomyoglobin (16.9 kDa protein lacking cysteine residues). Ubiquitin is notoriously resistant to proteolysis in aqueous conditions, so a one-hour aqueous-organic tryptic digestion protocol was optimized for our studies to set the stage for high-throughput analysis.[87, 110] Electrochemically oxidized proteins were generated by both electrodes, and tandem mass spectra (MS/MS) of proteolytic peptides were searched for fragment ions matching oxidation mass shifts by InsPecT software. Results for both intact protein oxidation and amino acid oxidation events detected from peptide fragmentation data are discussed below for each of the two working electrodes studied.

3.2.1 Results of Porous Flow-Through (PFT) Graphitic Carbon Electrode Experiments

The large surface area afforded by the PFT electrode, in addition to its reported ability to oxidize a range of intact proteins,[97, 98] made it an attractive candidate for testing its surface mapping applicability. The PFT electrode generated variable oxidation products for each of the three intact model proteins tested, although adsorption issues hindered optimal recovery of apomyoglobin and lysozyme.

3.2.1.1 Oxidation of Intact Proteins

Oxidation of intact ubiquitin could be achieved with a potential of +1.2V applied to the PFT electrode under "non-native" conditions (ACN/H_2O/FA (50/50/1, v/v/v)) similar to those developed by Permentier *et al* [98] for on-line electrochemically-induced proteolysis. The extent of oxidation was limited primarily to two +16 Da oxidation events, with the mono-oxidized species forming the base peak (Fig. 3-1). As detailed in Table 3-1, the most common mass shift possible due to oxidation is +16 Da, followed by +14 Da and others limited to only a few residues. Less common mass shifts occurring on only one or a few amino acids might be present in proportionally low abundance, but are difficult to resolve in the intact mass spectra when appearing near or between abundant peaks shifted by +16*n* Da (*n* is the number of oxygen additions). Note the appearance of fragmented versions of ubiquitin with minor abundances in Fig. 3-1, as evidenced by *m/z* $(817)^{8+}$ and $(934)^{7+}$ ions in Fig. 3-1A which deconvolute to a 6,530 Da neutral species, and $(1111)^{6+}$ and $(1333)^{5+}$ ions in Fig. 3-1B which deconvolute to a 6,661 Da neutral species. The fragment ions observed in Fig. 3-1A match the probable "proline effect" y58 ion.[111] The fragment ions appearing in Fig. 3-1B are likely due to either

electrochemically-induced fragmentation,[97, 98] or fragmentation within the mass spectrometer (possibly as an in-source MSAD [112] fragment ion caused by a high sample concentration). All mass spectra of oxidized intact proteins were examined for fragmentation. In all cases, the major product observed after electrochemical oxidation was the oxidized intact protein, with very little evidence of fragmentation observed. Selected ion monitoring (SIM) plots from on-line experiments with lysozyme conducted in the same non-native conditions as ubiquitin revealed that mono-oxidation occurred during the application of +1.2V (Fig. 3-2). Representative SIMs illustrate our findings that the yield of mono-oxidized protein increased when the flow rate was lowered from 2.5 μL/min to 1 μL/min (Fig. 3-2), effectively increasing residence time of the analyte protein inside the electrochemical cell.

3.2.1.2 Testing of Physiologically Relevant Buffers

While the oxidation of intact proteins was initially achieved in non-native conditions, our goal was to evaluate this methodology for probing native protein structures. The ideal solution for achieving electrochemical oxidation of natively folded proteins would preserve a physiological pH (7.0-7.5) with appropriate buffering capacity against electrogenerated H^+ ions. Additionally, an ionic strength of at least 150 mM would be necessary to simulate the salinity of intra- or extra-cellular environments.[113] Furthermore, an ideal ionic species would be electrochemically inert and exhibit low reactivity with hydroxyl radical.[40, 114] Various aqueous buffers were tested to satisfy these experimental requirements for electrochemical oxidation of folded proteins.

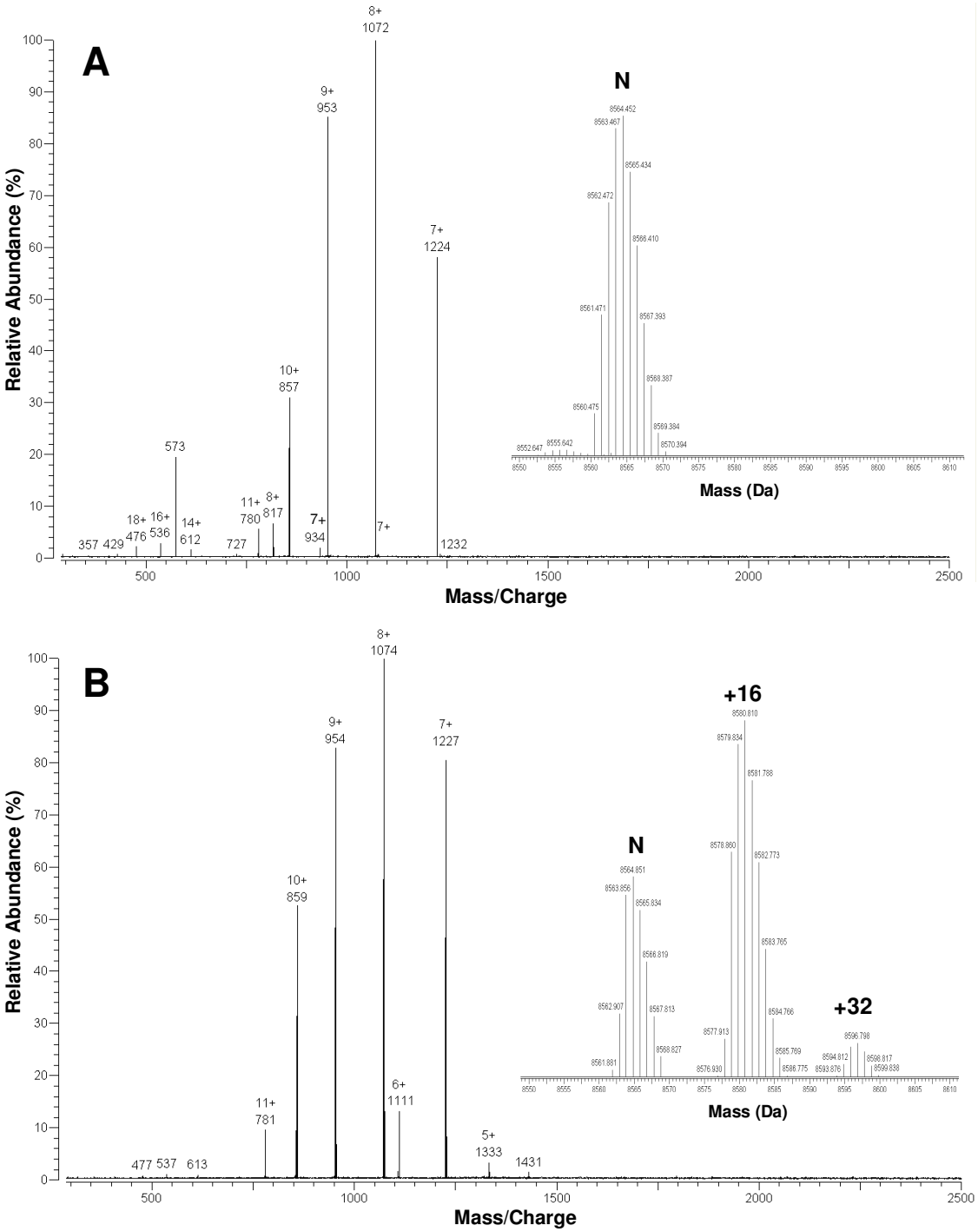


Figure 3-1: Intact protein FTICR-MS of ubiquitin before and after electrochemical oxidation.
(A) Unoxidized form, with inset confirming normal (N) deconvoluted mass of 8565 Da. (B) Oxidized ubiquitin; inset shows three species: N = normal (unoxidized), +16 = mono-oxidized (base peak), +32 = di-oxidized. *Experimental conditions*: 50 µM ubiquitin in non-native solvent at flow rate of 1 µL/min using PFT electrode off-line with +1.2 V potential applied.

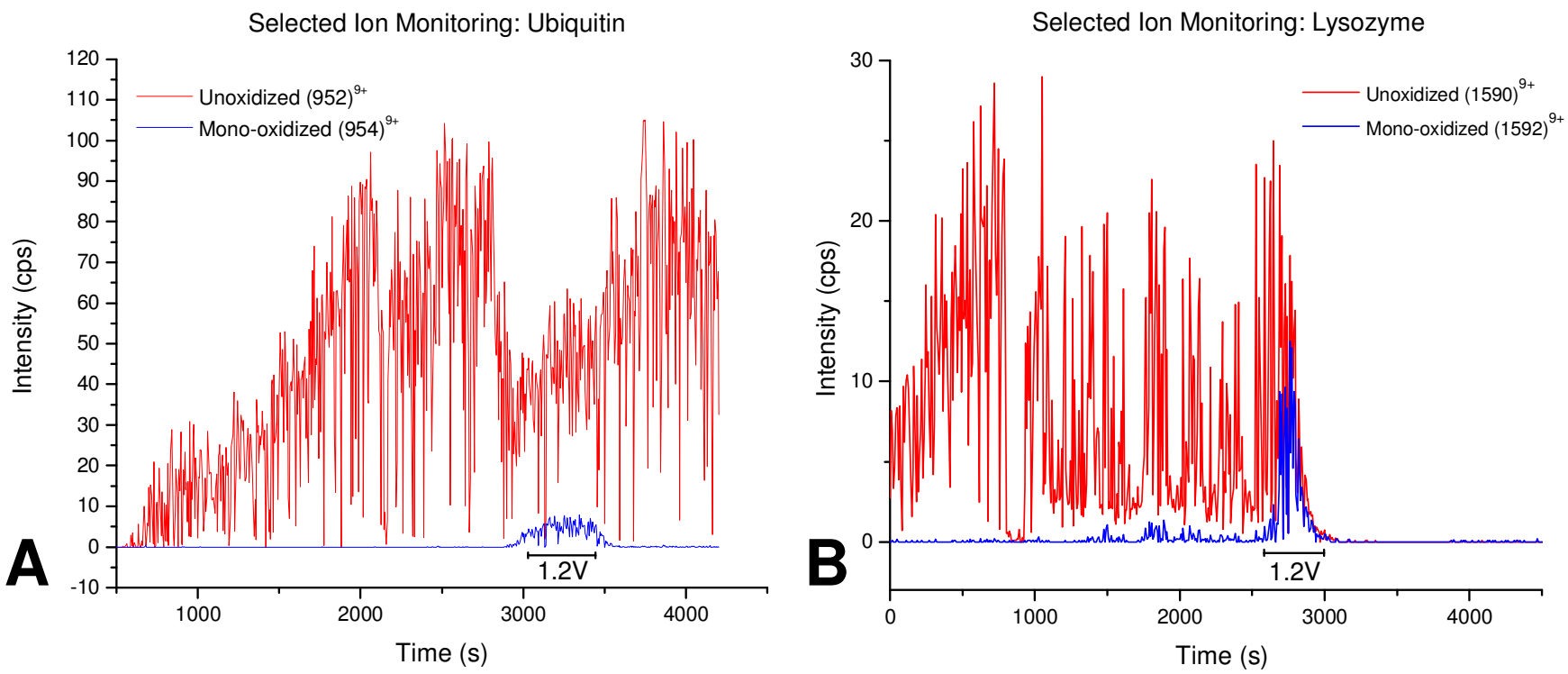


Figure 3-2: SIM of the ubiquitin (A) and lysozyme (B) 9+ charge state, showing unoxidized (red) and mono-oxidized (blue) signals over time. *Experimental conditions*: 1 µM protein in non-native solvent at flow rate of 1 µL/min using PFT electrode (+1.2 V potential applied over 10 min) on-line with FTICR-MS .Note absolute intensity is commonly ~10-fold higher for ubiquitin relative to equimolar lysozyme.

The extent of electrochemical oxidation achieved on intact ubiquitin for each electrolyte tested was monitored by direct infusion ESI-FTICR-MS after reversed-phase extraction followed by elution into ESI solvent. Initially, ammonium bicarbonate and ammonium acetate were tested but discarded as candidates due to low oxidation yields. Excessive bubbling occurred with the bicarbonate solution presumably due to carbon dioxide release, while gentle bubbling observed in off-line non-native studies suggested that solvent electrolysis was induced by the applied voltage. Further experimentation revealed that potassium sulfate is a good electrolyte for achieving electrochemical oxidation of proteins using both electrodes tested. Both ions of this electrochemically inert salt are each only one step more chaotropic than those in the widely used sodium phosphate salts according to the Hofmeister series,[115] suggesting its potential for stabilizing native protein structure. However, in our initial studies, we desired to more closely match the ionic conditions of these buffered solutions to those obtained for the non-native protein electrochemistry measurements. Therefore, we utilized a simple potassium sulfate solution with higher ionic strength (250 mM) and lower pH (6.0) than the ideal physiological buffer described previously. While this "native" solution is likely superior to the non-native solvent for promoting proper protein solvation and folding, it is still not optimum since the higher ionic strength may induce interferences with structurally important electrostatic interactions.[116]

3.2.1.3 Protein Adsorption and Electrode Regeneration

Electrochemical oxidation was tested in different aqueous buffers using the PFT electrode. However, protein adsorption posed a significant recovery issue, and cleaning methods

had to be utilized to achieve reproducibility. This adsorption phenomenon is illustrated in Fig. 3-2 showing the SIM plots of the unoxidized $(1590)^{9+}$ and mono-oxidized $(1592)^{9+}$ lysozyme in non-native solvent. MS signal intensity plummeted after the voltage application, presumably due to protein adsorption onto the porous graphite surface. Lysozyme adsorption seemed to occur across the conditions tested, as evidenced by backpressure increases and/or sample losses. Experiments with lysozyme in the ammonium bicarbonate buffer (500 mM NH_4HCO_3, pH 9, 61.7 mS/cm) often resulted in blockage of the electrode, and eventually protein precipitation in the collection tube revealed the source of electrode clogging. Obstructions encountered while using the PFT electrode may be explainable in that the oxidation of lysozyme lowers its isoelectric point (pI),[62] thus enhancing insolubility at the alkaline pH of this ammonium bicarbonate solution. In contrast, precipitate was not observed in experiments with potassium sulfate buffer (250 mM K_2SO_4, pH 6, 68.1 mS/cm), which facilitated oxidation for both ubiquitin and lysozyme. However, only ubiquitin was recovered in sufficient quantity from this buffer for digestion and LC-MS/MS analysis. Retention was distinctly less pronounced for ubiquitin, as illustrated by SIMs of the unoxidized and mono-oxidized ions in non-native conditions (Fig. 3-2). Importantly, consecutive experiments often showed a decline in the extent of oxidation products generated at the same voltages for which more extensive oxidation had occurred previously. A return to the initially observed level of oxidation could be induced after an overnight wash in the reverse direction with 50%-100% ACN at a flow rate of 2 μL/min (data not shown) to effectively regenerate its oxidative capacity. The appearance of contaminating protein from prior experiments occurred intermittently while using the PFT electrode, confirming the retention of protein within the electrode. Cleaning procedures performed with organic

solvents and/or acids in a reverse flow direction occasionally desorbed protein from the PFT electrode. Protein precipitate was observed during reversed passage of FA/H_2O solution (10/90, v/v) at 2 μL/min, and centrifugation was employed to separate the insoluble fraction prior to tryptic digestion. This observation may explain the 56% sequence coverage for peptides sampled from recovered apomyoglobin. In contrast, all other MS/MS searches resulted in 90% or greater coverage of the functional protein sequence.

Table 3-1: Oxidative Modifications for MS/MS Searches and Protein-Specific Identifications

Ubiq	Lyso	Apo	Residue	EA[a]	Modification type[b] (nominal mass shift in Da)
	●		Ala (A)		Hydroxylation (+16)
●	●		Arg (R)		Deguanidation (-43), Carbonylation (+14), Hydroxylation (+16)
●	●		Asn (N)		Hydroxylation (+16)
●	●		Asp (D)		Decarboxylation (-30), Hydroxylation (+16)
×	●	×	Cys (C)	Y	"Serination" (-16), Hydroxylation (+16), Sulfinic acid (+32), Sulfonic acid (+48)
●	●	●	Glu (E)		Decarboxylation (-30), Carbonylation (+14), Hydroxylation (+16)
●	●	●	Gln (Q)		Carbonylation (+14), Hydroxylation (+16)
●	●		His (H)	Y	Ring open: Asn (-23), Asp (-22), Aspartylurea (-10), Formyl-Asn (+5), Oxo-His (+16)
●	●	●	Ile (I)		Carbonylation (+14), Hydroxylation (+16)
●	●	●	Leu (L)		Carbonylation (+14), Hydroxylation (+16)
●	●	●	Lys (K)		Carbonylation (+14), Hydroxylation (+16)
●	●	●	Met (M)	Y	Aldehyde (-32), Sulfoxide (+16), Sulfone (+32)
●	●	●	Phe (F)		Hydroxylation (+16), Dihydroxylation (+32)
●	●		Pro (P)		Carbonylation (+14), Hydroxylation (+16), Dihydroxylation (+32)
●	●		Ser (S)		Carbonylation (-2), Hydroxylation (+16)
●		●	Thr (T)		Carbonylation (-2), Hydroxylation (+16)
×	●	●	Trp (W)	Y	Hydroxylation (+16), N-Formylkynurenine (+32), Kynurenine (+4)
●		●	Tyr (Y)	Y	Hydroxylation (+16), Dihydroxylation (+32)
●	●	●	Val (V)		Carbonylation (+14), Hydroxylation (+16)

[a] EA = electroactivity status. [b]Taken from compilation published in reference [72]. Structures for these oxidative modifications can be found in the Appendix; see review article for more detailed information. "●" = identified as oxidized; "×" = not present in protein amino acid sequence. Note Gly (G) oxidation was not included in search parameters, as oxidation of the α-carbon typically results in peptide backbone cleavage.

3.2.1.4 Peptide MS/MS Analysis

Oxidized ubiquitin samples were collected off-line from both the non-native and buffered solutions after passage through the PFT electrode, and then were digested for MS/MS characterization. Contained in Table 3-2 are only the ubiquitin residues which exhibited oxidative mass shifts (full survey in Appendix Table A-1); these occurred on a range of both electroactive and non-electroactive amino acids (classifications noted in Table 3-1). Among the most commonly oxidized sites (by peptide spectral count [117, 118]) were solvent-exposed, electroactive amino acids Tyr59 (19) found in both solutions and His68 (9) identified in the buffer only. High-scoring MS/MS matches for two different oxidative mass shifts (-22 Da, +5 Da) on the exposed histidine (His68 ~ 80 $Å^2$) within the same peptide sequence are shown in Fig. 3-3. This example demonstrates that detection of various electrochemically-induced oxidation events on the same amino acid increases confidence in the site-specific identification. The only methionine present in ubiquitin exhibited two different mass shifts, as identified by two MS/MS spectra showing mono-oxidation (+16 Da) in the non-native solvent, and oxidative loss (-32 Da) in the buffer. Many non-electroactive residues (I, L, K, R, P, S, T, Q, E, D, N) were also found to be oxidized, albeit at lower abundances reflected by fewer peptide spectral counts. According to calculated SASA values, 22 residues are buried in the ubiquitin structure. Mass shifts were detected for three of these amino acids from proteins oxidized in *non-native* solution (Leu15, Lys27, Ile61), while only two sites were found among peptides derived from the buffer solution (Ile3, Leu15). Leu15 was identified as both carbonylated (+14 Da) and hydroxylated (+16 Da) in both non-native and buffered solutions.

Table 3-2: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Ubiquitin

Ubiquitin[a]		SASA (Å^2)		PFT Oxidation		BDD Oxidation		Control
Residue	#	X-ray 1UBQ	NMR 1D3Z	Non-Native 1.2V ΔM(#)	Buffer 1.2V ΔM(#)	Non-Native 3.0V ΔM(#)	Buffer 2.1V ΔM(#)	Buffer ΔM(#)
MET	1	19.5	23.6	16(1)	-32(1)		16(3)	16(1)
ILE	3	0.0	0.0		16(1)			
LYS	6	98.3	83.4		14(2)		14(1)	
THR	7	15.1	14.0		-2(1)			
LEU	8	106.7	81.6		16(1)			
THR	12	41.9	39.6		-2(1)		-2(1)	
THR	14	58.2	53.8	16(1)		16(2)	16(1)	
LEU	15	3.1	0.6	14(1); 16(1)	14(1); 16(1)	16(1)	16(3)	
GLU	18	100.5	96.5			14(1)		
PRO	19	52.3	65.7		14(2)	32(2)	14(2)	
VAL	26	0.0	0.0			14(2)		
LYS	27	16.1	8.0	14(1)		14(3)		
LYS	33	81.3	82.6		14(2)	16(2)		
ILE	36	27.3	32.3			14(3)	14(2)	
PRO	37	56.9	67.2			14(3)	14(1)	
PRO	38	48.6	36.5		16(1); 32(1)	14(12); 32(2)	14(2)	
GLN	41	0.0	0.0			14(1)		
ARG	42	81.3	88.7		14(1)	14(10)		
PHE	45	38.3	39.9			32(3)		
LYS	48	92.2	100.6		14(3)	16(1)	14(1)	
GLN	49	80.7	86.9					14(1)
THR	55	34.3	31.9	-2(1)				
SER	57	46.1	43.3	-2(4); 16(1)	-2(1)			
ASP	58	46.3	52.7	16(1)				
TYR	59	24.9	31.5	16(5); 32(1)	16(13)	16(18); 32(3)	16(1)	
ASN	60	102.1	100.9	16(1)		16(1)		
ILE	61	0.0	0.0	16(2)				
LYS	63	122.2	144.2	16(2)				
GLU	64	81.3	108.2	-30(1)				
THR	66	42.6	38.5	16(1)				
HIS	68	75.8	89.8		-22(2); 5(7)		-23(1);-22(3); 5(4)	
LEU	69	0.4	0.6			16(1)		
ARG	72	106.1	93.0			-43(1)		
ARG	74	169.0	171.4	14(1)				

[a]Matches meeting filter criteria reported with mass shifts and peptide spectral counts (in parentheses). SASA calculated by GETAREA;[95] rows shaded gray indicate SASA ≤ 10 Å^2 for at least one of the two structures. Patterned cells denote residues within peptides missing from MS/MS protein sequence coverage.

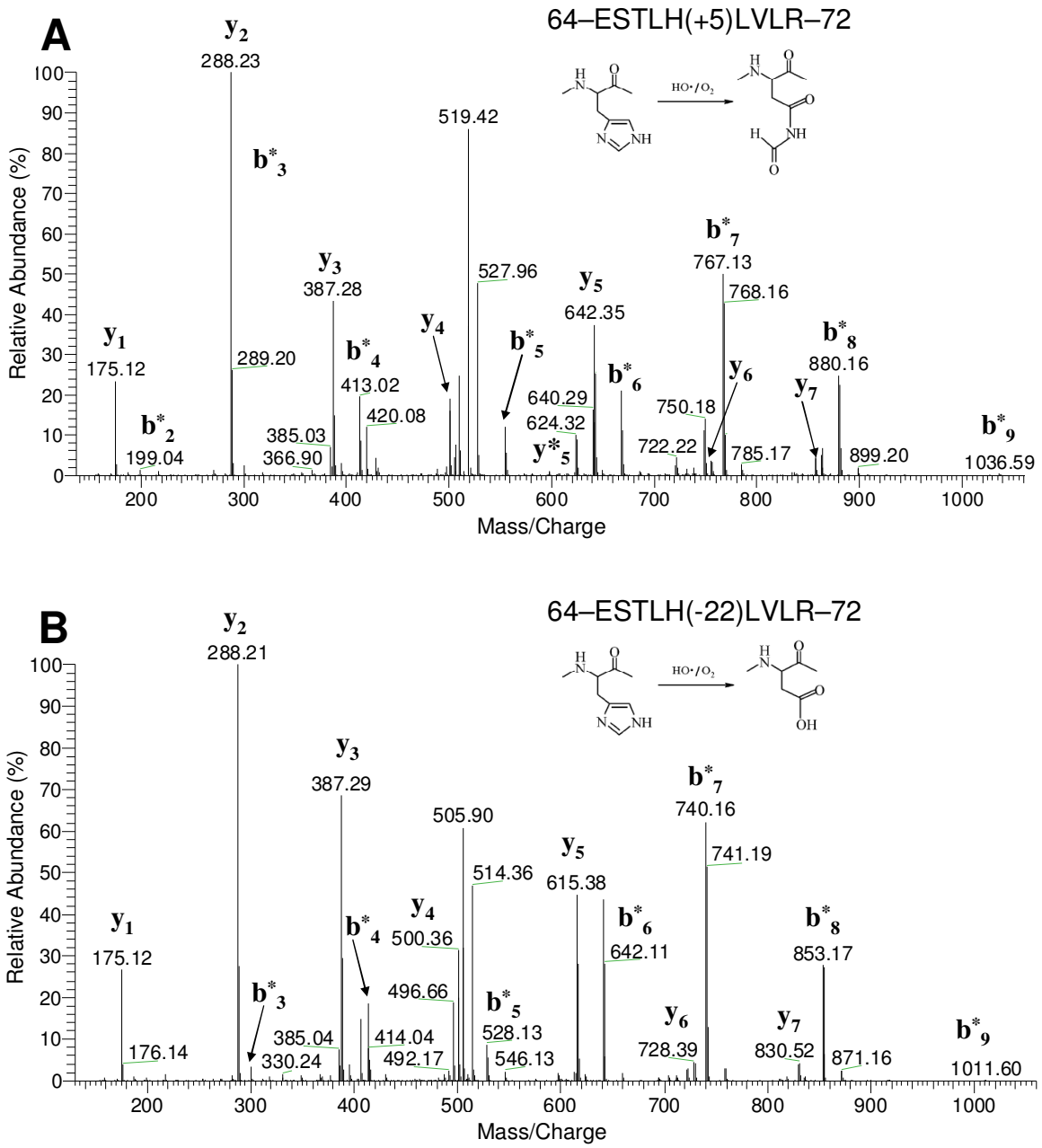


Figure 3-3: Representative MS/MS spectra show two different modifications on the same histidine residue. A) 64–ESTLH(+5)LVLR–72 B) 64–ESTLH(-22)LVLR–72. Induced modifications were achieved by both BDD and PFT electrochemical oxidation of ubiquitin in potassium sulfate buffer. Chemical structures were taken from Xu *et al*, 2007

Apomyoglobin lacks cysteines but contains two methionines, both of which were identified as oxidized (+32 Da) within the protein recovered from the PFT electrode (Tables 3-3 and A-2). Search results revealed oxidation of electroactive tyrosine (Tyr146), as well as three non-electroactive residues (Thr132, Lys133, and Leu149). The latter three modifications suggest contribution by an indirect mechanism of oxidation, although threonine and lysine are C-terminal neighbors to a highly reactive methionine (131-MTK-133). These residues are less reactive than methionine, which may be solvent-exposed in the non-native solvent, so MS/MS spectra were manually evaluated to confirm diagnostic fragment ions showing oxidation at these sites.

Table 3-3: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Apomyoglobin

Apo-Myoglobin[a]		SASA (Å^2)		PFT	BDD
Residue	#	X-ray 1WLA	X-ray 1AZI	Desorbed[b] 0V ΔM(#)	Non-Native 3.0V ΔM(#)
TRP	7	14.8	19.0		32(2)
GLN	8	114.6	120.3		16(1)
TRP	14	4.1	3.3		16(1); 32(1)
LYS	16	38.3	47.8		16(1)
GLU	18	64.4	75.2		14(1)
GLU	27	48.5	54.1		16(1)
LEU	32	8.6	8.0		14(2)
MET	55	6.2	7.3	32(1)	
LYS	56	80.6	84.2		16(2)
VAL	67	86.2	84.2		14(1)
LEU	69	2.4	3.4		16(1)
ILE	107	22.8	22.5		14(1)
MET	131	0.1	0.2	32(2)	16(6)
THR	132	48.6	52.3	16(1)	16(1)
LYS	133	69.6	82.6	16(3)	16(4)
LYS	145	52.6	40.4		16(1)
TYR	146	8.0	6.5	16(1); 32(1)	
LEU	149	47.4	56.6	16(1)	
PHE	151	22.99	19.63	32(1)	

[a]Matches meeting filter criteria reported with mass shifts and peptide spectral counts (in parentheses). SASA calculated by GETAREA;[95] rows shaded gray indicate SASA ≤ 10 Å^2 for at least one of the two structures. Patterned cells denote residues within peptides missing from MS/MS protein sequence coverage. [b]Formic acid wash used for protein desorption not favorable for native structure.

3.2.2 Results of Boron-Doped Diamond (BDD) Electrode Experiments

An attractive feature of the BDD electrode was its smooth surface that seemed promising for efficient recovery of oxidized protein. This was especially important in light of the adsorption problems encountered with the PFT electrode. An aqueous electrolyte solution is all that is needed for a BDD electrode to generate hydroxyl radicals,[109] which are capable of rapid reactions with a wide range of amino acids. This highly reactive species contributed to the oxidation of all three intact model proteins tested in this study, based on matches in the peptide MS/MS data discussed further below.

3.2.2.1 Oxidation of Peptide and Intact Proteins

Our initial BDD experiments conducted on-line with 5 μM methionine enkephalin (1-YGGFM-5) in non-native conditions (ACN/H_2O/FA (50/50/1, v/v/v)) showed that a minimum of +2.0V was necessary to achieve oxidation products at flow rates of 2-5 μL/min. In this case, we utilized flow rates that achieved steady ion current in the ESI source, with the oxidized yield increasing with lower flow rates (data not shown). Spectra scanning the entire mass range were acquired in chromatogram mode to allow selected ion monitoring (SIM) of unmodified and oxidized species. The unmodified parent ion (574.26 *m/z*) remained abundant at these flow rates, with minor (< 10% base peak) oxidation products at 590.25 (+16), 606.25 (+32), and 654.36 (+80). This mixture of products could result (in order of reactivity [51]) from oxidation of methionine to its sulfoxide (+16) or sulfone (+32) forms, and mono- or di-hydroxylation (+16 or +32) of tyrosine or phenylalanine. The presence of a bimodal distribution representing light and heavy oxidation products is not too surprising, as this phenomenon has been observed in our

protein over-oxidation experiments (data not shown). Rather than a smooth, linear increase in oxidation level, it is possible that initial oxidation events may enhance the efficient generation of a heavily oxidized species, thereby even skipping medium oxidation levels. Thus, it will be important to keep the total oxidation level in proteins to a reasonably low level. Given the observed trend in these on-line experiments that lower flow rates produced greater oxidation yields, subsequent experiments on intact proteins were conducted at flow rates of 2 μL/min or less.

Based on these initial peptide oxidation measurements, subsequent experiments were conducted off-line with 50 μM ubiquitin in the standard non-native solution. Potentials were varied in 0.5V steps from +2.0V to +3.0V, and three consecutive 20 minute fractions (0-20, 20-40, 40-60 min) were collected during each hour-long voltage application. Fig. 3-4 illustrates the mono-oxidized protein yield from the fractions collected within the voltage application window over incrementally increasing working electrode potentials. It is not yet clear why the oxidation level increases over time as illustrated in the fractions collected using the BDD electrode at +3.0V in Fig. 3-4. It is possible that mixing within the volume of the electrochemical cell allows sampling of an increasingly oxidized protein population over time. In later experiments (data not shown), the flow rate was lowered from 2 μL/min to 1 μL/min to effectively increase residence time of both solution and analyte inside the electrochemical cell while decreasing turbulence of the flowing solution. Intact mass spectra from these experiments showed that mono-oxidized ubiquitin or di-oxidized apomyoglobin formed the base peak among proteins collected in the first 20-minute fraction. However, this was not the case with lysozyme, as the unoxidized species retained base peak status with two abundant +16 Da mass shifts.

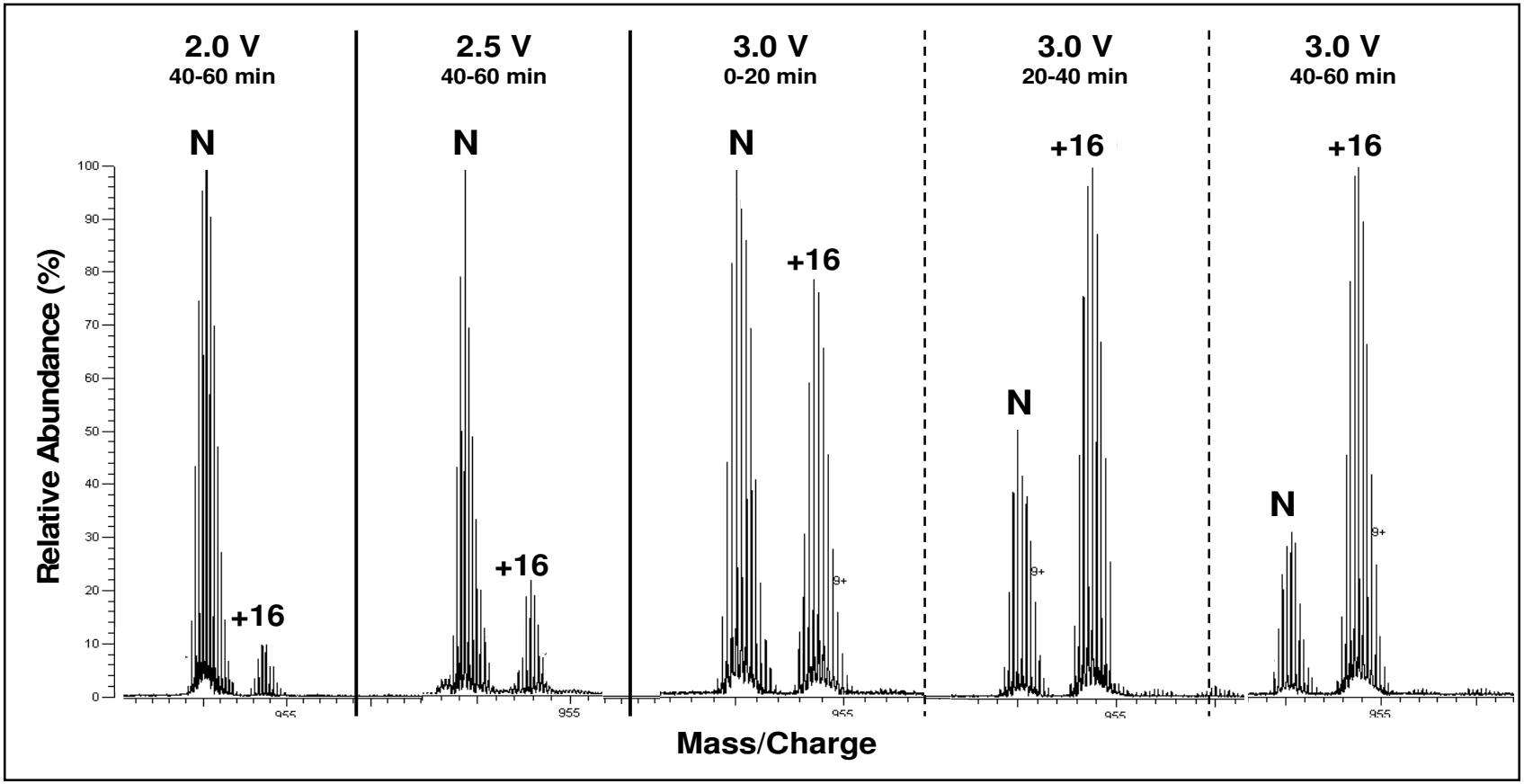


Figure 3-4: Mass spectra from three separate experiments showing increasing abundance of mono-oxidized ubiquitin relative to the unoxidized protein. *Experimental conditions*: 50 μM ubiquitin in non-native solvent at flow rate of 2 μL/min using BDD electrode off-line with range of potential applied (+2.0 to +3.0 V).

3.2.2.2 Oxidation in Physiologically Relevant Buffer

Electrochemical testing of the potassium sulfate buffer using the BDD electrode showed that application of at least +2.0V generated light bubbling in the collection tube indicative of solution electrolysis. Experiments with proteins in this buffer revealed a moderately increased level of oxidation, as shown in Fig. 3-5 for both ubiquitin and lysozyme relative to levels observed in non-native solvent. While unoxidized lysozyme formed the base peak in the intact mass spectra from proteins exposed to the electrochemical conditions generated by +2.05V at the BDD working electrode (Fig. 3-5B), mono-oxidized lysozyme was the most abundant species in the +2.10V collection (data not shown). Both of these samples showed a decreasing abundance (after the base peak) of increasingly oxidized protein; populations representing up to six +16 Da modifications on intact lysozyme were visible for collections from each voltage.

3.2.2.3 Peptide MS/MS Analysis

Overall, the BDD electrode exhibited a slightly lower number of detectable oxidation mass shifts relative to the PFT electrode for oxidized ubiquitin, based on high-scoring matches of singly-modified tryptic peptides contained in Table 3-2. Only the residues which underwent oxidative mass shifts are reported here (Table A-1 is unabridged), in addition to sulfur-containing residues Cys and Met regardless of their oxidation status. While electroactive tyrosine (Tyr59) claimed the highest peptide spectral count (21) among residues oxidized in the non-native solvent using the BDD electrode, its modification was detected only once among peptides derived from ubiquitin oxidation in the buffered solution. Three MS/MS spectra indicated hydroxylation (+16 Da) of the solvent-exposed methionine in the buffered solution,

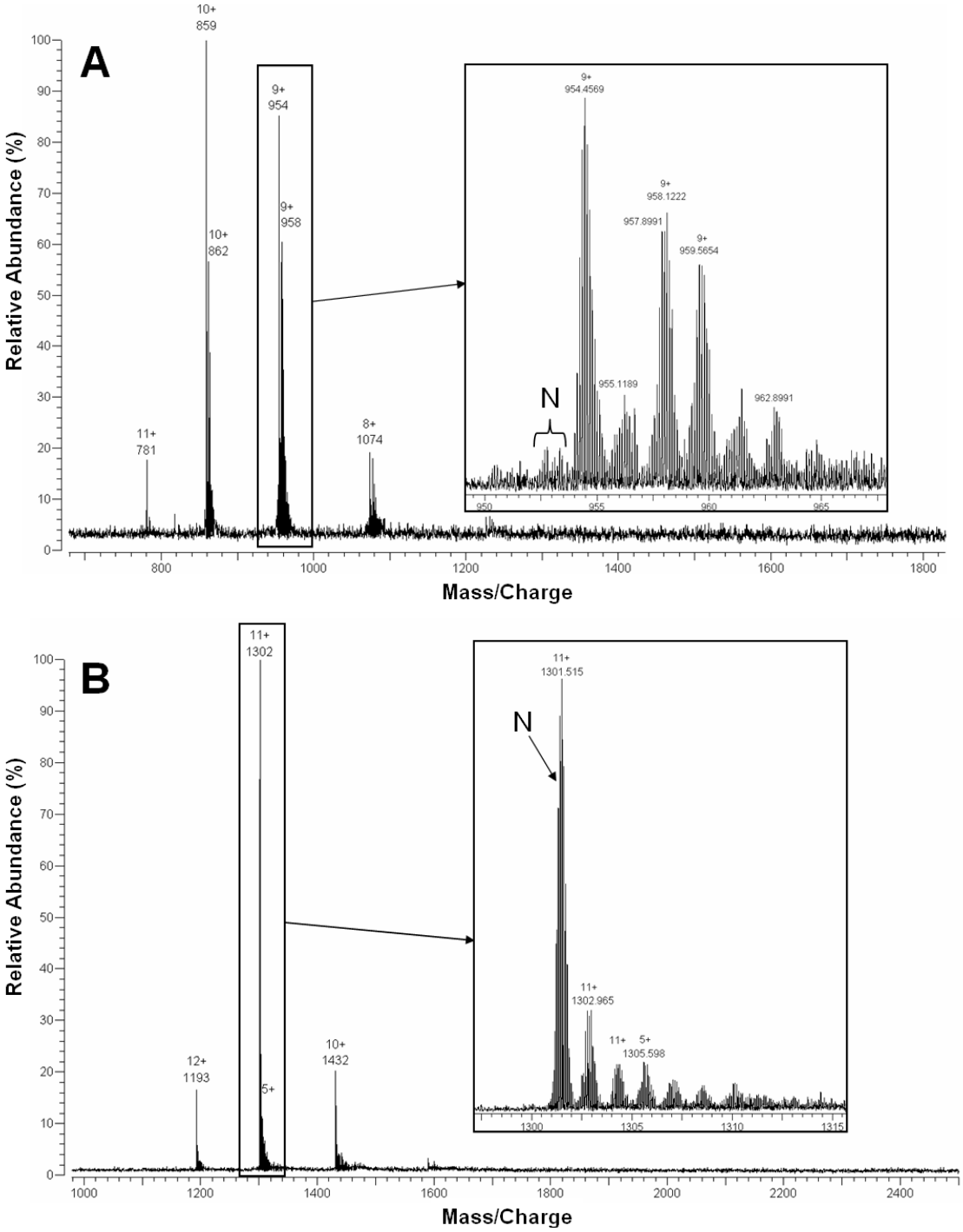


Figure 3-5: Mass spectra of proteins electrochemically oxidized in physiologically relevant solution. A) Ubiquitin inset shows mono-oxidized base peak with multiple mass shifts at higher *m/z* values; normal (N) unoxidized peak is absent. B) Lysozyme *m/z* values for unoxidized charge states are shown above each heterogeneous peak. *Experimental conditions*: 50 µM protein in 250 mM potassium sulfate solution at flow rate of 1 µL/min using BDD electrode off-line with +2.1 V potential applied.

while none reflected its oxidation in the non-native solvent. Three different modifications (+5 Da, -22 Da, -23 Da) occurred on His68 when electrochemically oxidized in the buffered solution, two of which are illustrated in aforementioned Fig. 3-3. Additionally, a +5 Da shift was found on a different tryptic peptide (55-TLSDYNIQKESTLH(+5)LVLR-72), further corroborating His68 oxidation in this sample. A survey of total peptide spectral counts for oxidation events among non-electroactive residues of proteins oxidized in both solutions showed that proline (24) and arginine (11) residues ranked highest. These were followed by fewer MS/MS spectra pinpointing oxidative mass shifts on a range of other residues (F, I, L, K, V, T, Q, E, N). A total of five buried residues belonging to the non-electroactive subset were oxidized in the *non-native* solvent (Leu 15, Val26, Lys27, Gln41, Leu69), compared with only one buried residue (Leu15) hydroxylated (+16 Da) in the buffered solution. This site was found to be oxidized in both non-native and buffered solutions using both the PFT and BDD electrodes. Interestingly, its neighboring residues are calculated to be solvent-exposed.

Apomyoglobin oxidized in *non-native* solvent using the BDD electrode with +3.0V applied resulted in oxidation events summarized in Table 3-3 (see Table A-2 for unabridged version). Only one of two methionines was identified as oxidized, while oxidative mass shifts occurred on both tryptophans. Neither of the two tyrosines nor any of the eleven histidine residues was found to be oxidized. Given that two oxidations could occur on a single peptide based on the intact MS, searches were conducted that allowed two modifications per peptide (data not shown). This resulted in several high-scoring histidine oxidations with coincidental modification of a nearby residue. Additionally, these search results indicated that tyrosine hydroxylation (103-Y(+16)L(+14)EFISDAIIHVLHSK-118) occurred adjacent to a carbonylated

leucine. Singly-oxidized peptide matches passing stringent filters revealed oxidation events on a range of non-electroactive residues (W, L, I, V, K, T, Q, E) with low peptide spectral counts.

Lysozyme was collected after electrochemical oxidation using the BDD electrode in two different solutions: non-native (ACN/H_2O/FA (50/50/1, v/v/v)) and buffered (250 mM K_2SO_4) at two different electrode voltages. The search results indicative of high-scoring singly-oxidized peptide matches are summarized in Table 3-4; only the modified residues are reported here (see Appendix Table A-3 for complete survey). Both methionines were identified as oxidized with +16 Da mass shifts in the non-native solvent, while neither was found to be modified in the control digest or after oxidation in the buffer solution. Oxidative modification was detected for three of the eight disulfide-bonded cysteines in non-native solvent; all three are calculated to be buried according to high resolution structures. The only exposed cysteine (Cys6 ~ 40 $Å^2$) was not detected as oxidized in the non-native solvent, but underwent singly- and doubly-oxidized mass shifts in the samples electrochemically oxidized in buffered solution. All residues found oxidized in the buffered samples exhibit solvent-exposed sidechains based on SASA calculations of high resolution lysozyme structures, with no oxidation of buried sites observed.

3.2.3 Interpreting Oxidation Search Results

Though the analytical details discussed thus far are important for conducting electrochemical oxidation experiments, interpretation of the data gathered is a crucial process and can be rather laborious for a number of reasons. First, the extensive list of modifications (Table 3-1) that were included in the MS/MS search parameters poses a formidable combinatorial challenge even when only one protein is involved, as database search space

Table 3-4: Modified Amino Acid Sites Identified from Electrochemically-Oxidized Lysozyme

Lysozyme[a]		SASA (Å^2)		BDD Oxidation Results			Control
Residue	**#**	**XRC 1AKI**	**NMR 1E8L**	**Non-Native 3.0V ΔM(#)**	**Buffer 2.05V ΔM(#)**	**Buffer 2.10V ΔM(#)**	**Tryptic Digest 60%ACN ΔM(#)**
CYS	6	42.6	37.5		-16(1);32(1)	32(1)	48(1)
ALA	11	6.2	17.4	16(1)			
MET	12	0.1	0.3	16(3)			
LYS	13	63.7	89.2	16(2)	16(1)		
ARG	14	129.4	159.3	16(4)	16(1)	16(2)	
HIS	15	38.8	34.5	16(1)			
ASN	27	15.2	10.2	16(1)			
TRP	28	0.0	0.0	4(1);16(1);32(1)			
CYS	30	1.1	0.1	-16(2)			
LYS	33	59.3	44.2	16(1)			
GLN	41	94.3	84.2	14(1)			
ARG	45	116.3	132.5	14(1)			
TRP	63	30.1	35.9	32(1)			
PRO	79	77.9	67.6			14(1)	
CYS	94	0.5	0.0	-16(1)			
SER	100	22.7	28.6	-2(1)			
ASP	101	70.8	79.3	16(1)			
MET	105	0.0	0.0	16(3)			
ASN	106	33.2	58.7	16(2)			
ALA	107	16.6	18.8	16(1)			
TRP	108	6.4	7.5	16(1); 32(1)			
TRP	111	10.8	28.7	32(2)			
ARG	112	90.5	101.3				16(1)
ARG	114	109.9	128.2			14(1)	
CYS	115	0.0	1.1	-16(3)			
LYS	116	100.8	107.3	14(1)		14(1)	
TRP	123	49.0	75.9	4(1); 32(4)	4(1);32(1)	32(2)	
ARG	125	115.4	168.9	16(1)		16(1)	

[a]Matches meeting filter criteria reported with mass shifts and peptide spectral counts (in parentheses). SASA calculated by GETAREA;[95] rows shaded gray indicate SASA ≤ 10 Å^2 for at least one of the two structures. Patterned cells denote residues within peptides missing from MS/MS protein sequence coverage.

expands significantly with each additional mass shift. The InsPecT software used to search MS/MS data in this study was critical for handling so many modifications. Searches of MS/MS data conducted in our lab typically use the SEQUEST algorithm,[89] which limits searches to a maximum of three PTMs (static or dynamic), and DBDigger,[88] which performs well for non-PTM proteome searches but has been reported to suffer high false discovery rates when multiple modifications are considered.[119] InsPecT allows all PTMs to be considered simultaneously, so allowance of more than one modification per peptide requires extensive manual evaluation to counter increased false positives, evidenced by several instances of multiple mass shifts on a single residue. Therefore, only singly-modified peptide results are confidently reported in this study. Stringent filters were established to maximize confidence in oxidized matches, with the tradeoff that potentially informative spectra were disregarded. For this reason, both analytical and technical replicates should be acquired to increase the likelihood of obtaining high-scoring peptides representing each oxidized residue.

Another challenge lies in the potential for co-eluting oxidized peptide isomers that are isolated and fragmented together due to having isobaric parent ion masses, which can result in the loss of site-specific oxidation data. Generation of shorter peptides by a less specific protease like pepsin could be advantageous in producing a more heterogeneous peptide mixture to decrease the likelihood of peptides with identical or overlapping retention times. Shorter peptides would also minimize the number of reactive residues contained in a single peptide to better match the single modification per peptide limit imposed on the data reported here. The unfiltered search output includes all of the abovementioned characteristics used to rank peptide matches, and alternate matches to the top-scoring oxidized peptides can be accessed to manually

evaluate the possibility that multiple species were fragmented in a single MS/MS scan. While this interpretation process may appear computationally cumbersome, such exercises can provide critical structural information in a relatively swift manner.

3.3 Discussion

The goal of this study was to evaluate the efficacy of using electrochemically-generated hydroxyl radicals to oxidize model proteins maintained in buffered, physiologically relevant conditions. Both electrodes investigated in this study were useful for the generation of oxidized intact protein with characteristic oxidative modifications identified on the peptide level. However, significant sample losses concomitant with pressure spikes that required cleaning procedures suggest the PFT electrode is subject to hysteresis, in which past events collectively affect subsequent performance. Collection of insufficient protein yield from the electrochemical cell prohibited further analysis of PFT-oxidized lysozyme. This observation might be attributable to its potentially adsorptive cysteine residues, which are absent from the other two proteins tested. As expected due to the non-porous electrode surface, significantly higher yields of oxidized proteins were observed using the BDD electrode relative to the PFT electrode. This is almost certainly due to less efficient mass transport of the protein to the BDD electrode surface, facilitating recovery and peptide MS/MS analysis for all three oxidized lysozyme samples. Although significant sample losses indicative of irreversible adsorption are not observed with this electrode, we cannot rule out the possibility that some reversible adsorption may occur. However, the BDD electrode generates copious amounts of hydroxyl radicals, so it seems likely that oxidation reactions with analyte protein would occur primarily in solution.

This study demonstrated that the BDD electrode could potentially be useful in protein structure studies, with less protein adsorption problems than the PFT. It should be noted that some reactive residues calculated to be solvent-exposed were not identified as oxidized in this preliminary study. While this could be attributable to the imposed search constraints of a single modification per peptide, false negatives such as these could be evaluated by extensive electrochemical oxidation of intact proteins to obtain a comprehensive list of vulnerable sites. These would be used for comparison with results from modest oxidation of folded proteins, which is important because even exposed reactive residues could be kinetically occluded by neighboring residues. Note that a balance must be achieved between false positive and false negative levels. On one extreme end, very light protein oxidation will likely ensure that native protein structure is retained, but may not provide extensive enough coverage to ascertain much information about the protein structure. At the other extreme, very extensive protein oxidation will undoubtedly label most of the reactive sites, while providing less useful information with respect to protein structure determinations. The key aspect here is to conduct moderate oxidation that enables fairly detailed labeling representative of true solvent accessibility in the folded protein. Fortunately, a simple change of the working electrode potential can directly alter the extent of oxidation observed on the intact protein. Additionally, use of a less conductive solution that is more favorable for promoting native protein structure may allow further fine-tuning of the electrochemical conditions. Even under optimal conditions, the false positives (i.e. buried sites that are oxidized) may not be completely zero; thus, it is important to empirically determine appropriate experimental oxidation conditions and proper filters on search results to minimize the false positives. These purported false positives, such as the labeling of Leu15 in buffered

solutions of ubiquitin using both the PFT and BDD working electrodes, may serve as an indicator of either extended oxidative damage or dynamic protein flexibility in solution.

Some additional factors will need careful consideration as electrochemistry is further evaluated and optimized for surface mapping applications. As mentioned previously, complications such as low abundances across a range of different oxidized peptides, a wide variety of possible mass shifts, and fragmentation of co-eluting oxidized peptide isoforms pose challenges to data interpretation. The utilization of multidimensional chromatography [120] in addition to longer aqueous-organic gradients could minimize these potential problems while increasing the dynamic range to allow identification of more low-abundance species relative to one-dimensional results. Also, secondary oxidation events occurring after exposure to the controlled electrochemical environment due to lingering ROS could be avoided by the incorporation of peroxidases and/or hydroxyl radical scavengers.[121] We omitted these reagents to avoid adding variables to an already complex scenario. Clearly, additional work is needed to more fully optimize the buffered conditions (ionic strength, pH, etc.) to ensure native protein structure. Furthermore, as electrochemical oxidation is optimized in future studies under more native conditions, use of spectroscopic methods to monitor alteration of secondary structure as a function of oxidation yield will facilitate understanding how oxidation level may distort native conformations. Nonetheless, our focused preliminary experiments demonstrate a clear potential for this method to evaluate a wide range of solvent accessible amino acid residues differentially labeled between native and non-native protein conformations.

3.4 Conclusions

Electrochemical oxidation has been shown to generate useful oxidative labels on intact proteins, in both non-native and buffered solutions. Also, the extent of oxidation can be systematically regulated by adjusting experimental parameters including the solution flow rate, applied working electrode voltage, and solution composition. Digestion of the oxidized proteins collected during off-line experiments allowed identification of oxidation sites using MS/MS analysis and evaluation against calculated SASA values. We observed electrochemical oxidation of a wide range of amino acids far beyond the electroactive subset. Additionally, search results showed agreement between solvent accessibility and the oxidation status of many ubiquitin residues oxidized in buffered solution. This provides encouragement that electrochemistry could be used for oxidative surface mapping experiments to evaluate the solvent accessibility of amino acids in a natively folded protein structure. Further optimization and extensive testing of electrochemical oxidation using the BDD electrode on more model proteins will facilitate wider application of this mass spectrometry-based approach in protein structure research.

4. An Electrochemistry-Based Pipeline to Probe Solution-Phase Protein Dynamics with Rapid Oxidation Exposure and Multidimensional Peptide Separations

This manuscript will be submitted to *Analytical Chemistry* with the following author list:

McClintock, C.; Parks, J.M.; Ghatty, P.K.; Kertesz, V.; Hettich, R.L.

4.1 Introduction

While a range of techniques are currently practiced for generating hydroxyl radicals (•OH) for SASA probe experiments, most involve highly regulated and costly equipment that is not readily accessible to many researchers. The goal of this work was to develop a method that accomplished comparable research objectives using a BDD electrode, whose basic operation was demonstrated in Chapter 3. A large body of literature describes the capacity of the BDD thin film electrode to generate •OH from water, thus allowing indirect oxidation events while negating the use of harsh oxidizing reagents. However, like all other electrodes, the BDD can also achieve oxidation by a direct pathway, which would involve physical interaction with the analyte protein that could potentially distort native structure.

Key adjustments were made to the previously reported electrochemical protein oxidation protocol.[76] The changes involve a different cell activation approach for more constant •OH production, and a reduced flow rate for more rapid exposure to the oxidizing environment inside the electrochemical flow cell chamber. Additionally, concentrations of both electrolyte and analyte protein were altered to better match oxidant flux, allowing improved regulation by the user. These changes facilitate a quick oxidation of proteins in a more physiologically relevant

ionic strength salt solution, with a dramatically shorter residence time during which adsorption could occur. While this may not be the ideal timescale to match the dynamic movements of solution-phase proteins, these findings have laid the groundwork for potential improvements.

The goal of this work is to establish an experimental "pipeline" for electrochemical surface mapping of proteins that could prove to be a burgeoning conduit for high-throughput structural proteomics screening. In this scenario, proteins would be rapidly oxidized, extracted for intact protein MS analysis while most of the oxidized protein undergoes a quick one-hour proteolytic digestion, followed by peptide level MS analysis via two chromatographic approaches. The first separation would allow a fast survey of oxidized sites, while a second, more sensitive approach would permit a deeper probe of oxidized sites. With this arrangement, all data for an experiment could be acquired in under 8 hours, provided the instrumentation is available. The most time-intensive aspect of this pipeline would be searching the tandem mass spectral data for oxidized sites, depending on the size of the database employed, the oxidative mass shifts considered, and the number of spectra to be interpreted. These informatic challenges are discussed in Chapter 5.

This chapter is focused on an optimized approach for achieving electrochemical oxidation of structurally robust model proteins ubiquitin, lysozyme, and apomyoglobin, with the latter presenting the case for an unresolved protein structure. The purpose for conducting extensive studies using model proteins is to foster a clearer understanding of residue-specific labeling patterns in relatively simple and well-characterized systems, with two of the three devoid of the highly reactive and often disulfide-bonded cysteines. The dynamic range challenge in detecting uniquely modified peptides within an oxidized protein sample was addressed by

separating the mixture according to electrostatic charge (strong cation exchange) in addition to the standard characteristic of hydrophobicity (reversed-phase). Oxidation sites mined by the InsPecT software package with updated input parameters are reported for both previously described and newly acquired MS data, with accurate detection of oxidized peptides continuing to present a complicated issue.

4.2 Experimental Refinement of Electrochemical Parameters

Previously we reported a method for the surface mapping of proteins utilizing electrochemistry to achieve oxidation in a physiologically relevant salt solution.[76] While sufficient to demonstrate the application potential, the experimental conditions were not optimal for a rapid oxidation reaction favoring the indirect •OH-mediated route instead of a direct (heterogeneous) electron transfer event which could involve interaction with the electrode surface. Both mechanisms are possible, with contribution from the indirect route increasing with voltage.[69] A dramatic reduction in the time that protein resides inside the electrochemical cell was achieved by a 100-fold increase in flow rate. The electrochemical compartment contains a volume of approximately 1 μL between the working and counter electrodes when separated by a 50 um gasket, so the residence time was lowered from 1 min at 1 μL/min to 1.2 seconds at a flow rate of 100 μL/min. The rapid flow rate virtually eliminated the voltage oscillation attributable to the nucleation of electrolytic bubbles on the BDD surface by hastening their exit from the cell – a change that could extend the useful lifetime of the BDD electrode. Bubble formation can damage the BDD film by gradually decreasing the electrode surface area contacted by solvent, which threatens the dwindling interface area with a rapidly increasing

current density. The performance of the BDD electrode used in the previous report declined over the course of one year, likely due to a combination of damage caused by bubble accumulation at low flow rates and fouling that occurred during testing of different electrolytes in addition to large cysteine-containing proteins. Observations discussed below suggest that a moderate hysteresis effect may contribute to increasing voltages with the application of same current level over subsequent runs, whether due to adsorption of electrolyte, protein, or both.

The use of a new BDD electrode allowed us to achieve oxidation of lower micromolar concentrations (1-10 μM) of protein that appeared to better match oxidant flux. This was accomplished in half the previously used concentration of inorganic salt, 125 mM potassium sulfate (pH 6.0), which remains a physiologically relevant concentration of a weakly buffered electrolyte that absorbs minimal ultraviolet (UV) light.[122] An interesting consideration about this electrolyte is whether the sulfate anion could become radicalized ($SO_4^{2-} \rightarrow e^- + \bullet OSO_3^-$) by direct oxidation and either attack proteins directly or coalesce into peroxydisulfate ($S_2O_8^{2-}$) whose reactivity is similar to hydrogen peroxide.[123] If so, this reactivity reportedly could confer advantages to protein studies,[71] with the same cautions that generally apply to peroxides. The oxidation-reduction potential of peroxydisulfate is +2.1 V, which directly corroborates the only CV peak observed with potassium sulfate solution alone inside the electrochemical flow cell, as shown in Fig. 4-1. Additionally, this has been empirically identified as the potential that most often results in moderate protein oxidation. Interestingly, the final reaction product would effectively mirror the most common oxidative mass shift for hydroxyl radical labeling (typically +16 Da), because the protein-attached potassium sulfate (KSO_4, +135 Da) would only be an intermediate state that is then converted to an alcohol (Fig. 4-

1 inset).[124] While the reaction components shown on the inset were not explicitly added to the solution, the pathway evokes contemplation for what could be taking place within the activated electrochemical cell.

Changes in electrochemical parameters included a decrease in applied current from 500 μA (+2.1V with older BDD) to 100-300 μA, which produced a potential ranging from +1.35 V to +2.10 V under these experimental conditions with the new BDD electrode. All three model proteins were reproducibly oxidized at variable levels depending on the controlled flow rate or electrode current employed, while occasionally producing some evidence of interaction with the electrode surface.

4.3 Protein-Electrode Interaction Studies

Factors which may influence the propensity of a protein to interact with an electrode surface include net charge, cysteine content, and hydrophobicity. A number of electrochemical studies have documented the adsorption of free cysteine to various electrode surfaces;[101, 102] however, BDD electrodes were shown to perform much better than glassy carbon electrodes because they did not suffer the same measurable adsorption of this sulfur-containing residue.[103] If either cysteine or the disulfide-bonded cystine indeed contributes any tendency bind the BDD surface, then proteins containing more of these residues would adsorb more readily. Other features that could affect adsorption include both the isoelectric point (pI) and the hydropathicity index (GRAVY),[125] which are unique to each protein based on the amino acid composition. Proteins retain a net charge when the solution pH differs from the pI value, so the likelihood of electrostatic interaction between protein and electrode increases with stronger net

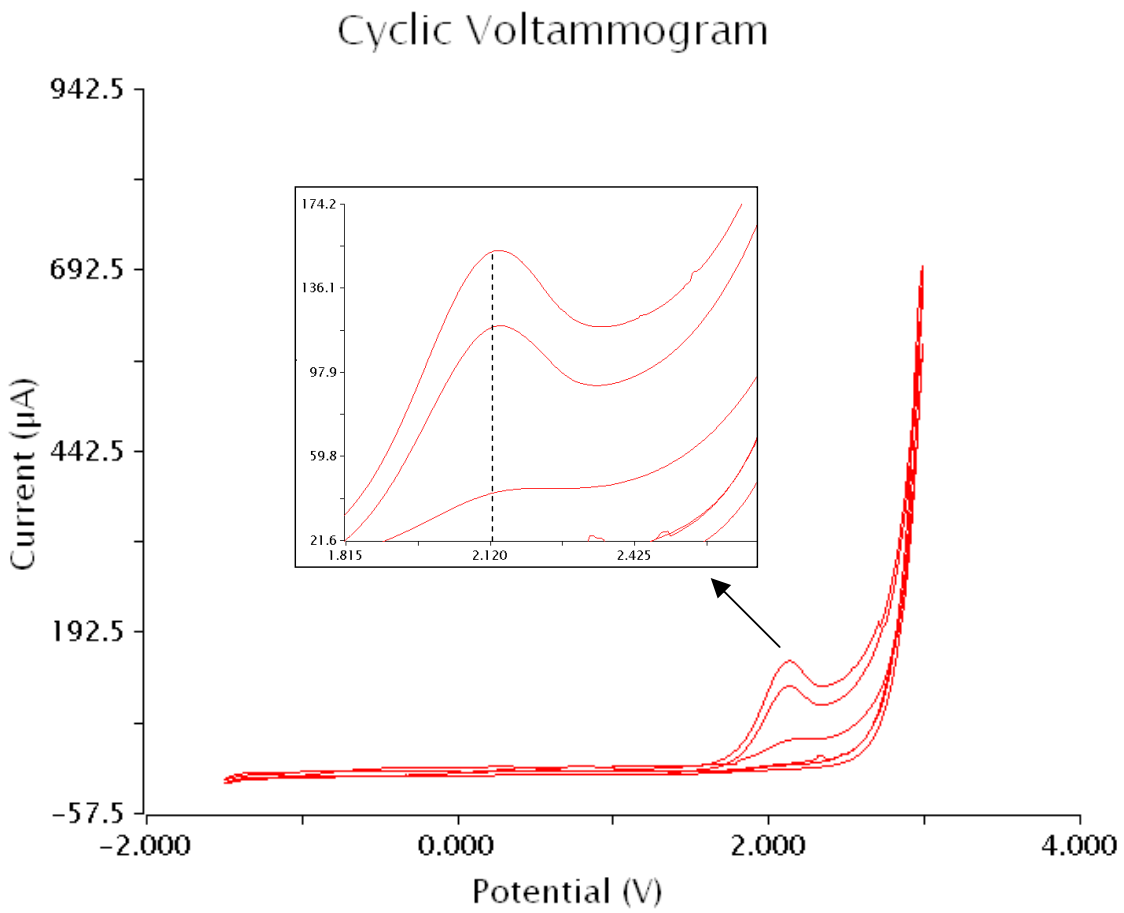


Figure 4-1: Electrochemical detection of oxidation-reduction potential for potassium sulfate solution. The cyclic voltammogram shows the current required to scan the voltage range of -1.5 V to +3.0 V over three cycles. During the rising phase of each cycle, an anodic potential pulls electrons onto the BDD electrode, resulting in oxidation of nearby solution components. The inflection suggests oxidation of the sulfate anion (E_O = +2.12 V).

charge in pH 6.0 potassium sulfate solution (Table 4-1). Oxidation can alter the intrinsic protein pI by causing loss of charged functional groups, in addition to increasing surface electron density through the addition of electronegative oxygen atoms via hydroxyl or carbonyl groups.[72] The GRAVY value indicates greater hydrophilicity with increasingly negative values, so apomyoglobin would be most likely to experience hydrophobic interaction of the three model proteins tested.

Proteins exiting the electrochemical cell were measured by a dual-wavelength (205 / 280 nm) ultraviolet absorbance monitor in an effort to detect changes in protein concentration that would indicate adsorption of protein during electrolysis (Fig. 4-2). Absorbance at 205 nm indicates presence of the peptide bond, while 280 nm signifies primarily tryptophan, with minor contributions from tyrosine and disulfide bonds.[122] Controls with potassium sulfate solution showed an increase of 15-30 mAU at 205 nm (< 2 mAU at 280 nm) due to activation of the electrochemical cell alone. Comparable increases are seen during electrolysis of protein solution at a 10-30% increase from baseline protein UV absorbance. However, a drop below the protein signal can be seen upon electrode activation and/or deactivation with some proteins. Ubiquitin showed an initial drop in 205 nm absorbance of roughly 3% relative to the total protein signal (2.5 μM ~ 250 mAU) just as the BDD was activated (data not shown). Lysozyme did not appear

Table 4-1: Physical features of model proteins subjected to electrochemical oxidation

Protein	MW	pI	GRAVY	Met	Cys	Trp	Ext-Coeff
Ubiquitin	8565	6.6	-0.489	1	0	0	1490
Apomyoglobin	16951	7.4	-0.396	2	0	2	13980
Lysozyme	14306	9.3	-0.472	2	8	6	37470

to suffer the initial drop, but showed a dip of about 5% relative to the protein signal (5 μM ~ 690 mAU) upon BDD deactivation before returning to the baseline protein UV absorbance (Fig. 4-2). This discrepancy in the timing of absorbance changes occurring early versus late is speculated to provide a clue about the physical property that most strongly confers adsorption propensity. These drops are not observed in control experiments of BDD activation alone, suggesting there may be a change in protein concentration during the electrolysis timeframe. These potential losses do not hinder recovery for further analysis even at the low micromolar concentrations tested, but any change in protein concentration could impact oxidation kinetics due to altered match with flux of hydroxyl radicals over the experimental timeframe.[126]

Interrogation of the BDD electrode conductivity as a measure of surface integrity was gauged by open circuit voltage (OCV) readings and cyclic voltammetry (CV) cleaning procedures conducted between protein oxidation experiments. Details about each of these methods are outlined in Chapter 2. Immediately following activation of the electrode with a positive voltage, the OCV reading would fall rapidly from a positive reading back to a negative potential, then decay slowly to resting potential. With the new BDD electrode, resting OCV readings in potassium sulfate solution were in the range of -150 mV to -250 mV and did not deviate significantly over the course of one year. It is not well-defined what level of insulation would cause a gradual drop of 100 mV during this time. Conversely, decreases in current magnitude at voltage endpoints in CV plots were occasionally noted after a protein experiment relative to the one acquired just prior to it, which suggested some level of adsorption had occurred to insulate the electrode surface. Evidence of this phenomenon is visible in Fig. 4-3, wherein the CV current endpoints changed after the experiment from near ±5.0 closer to ±2.5.

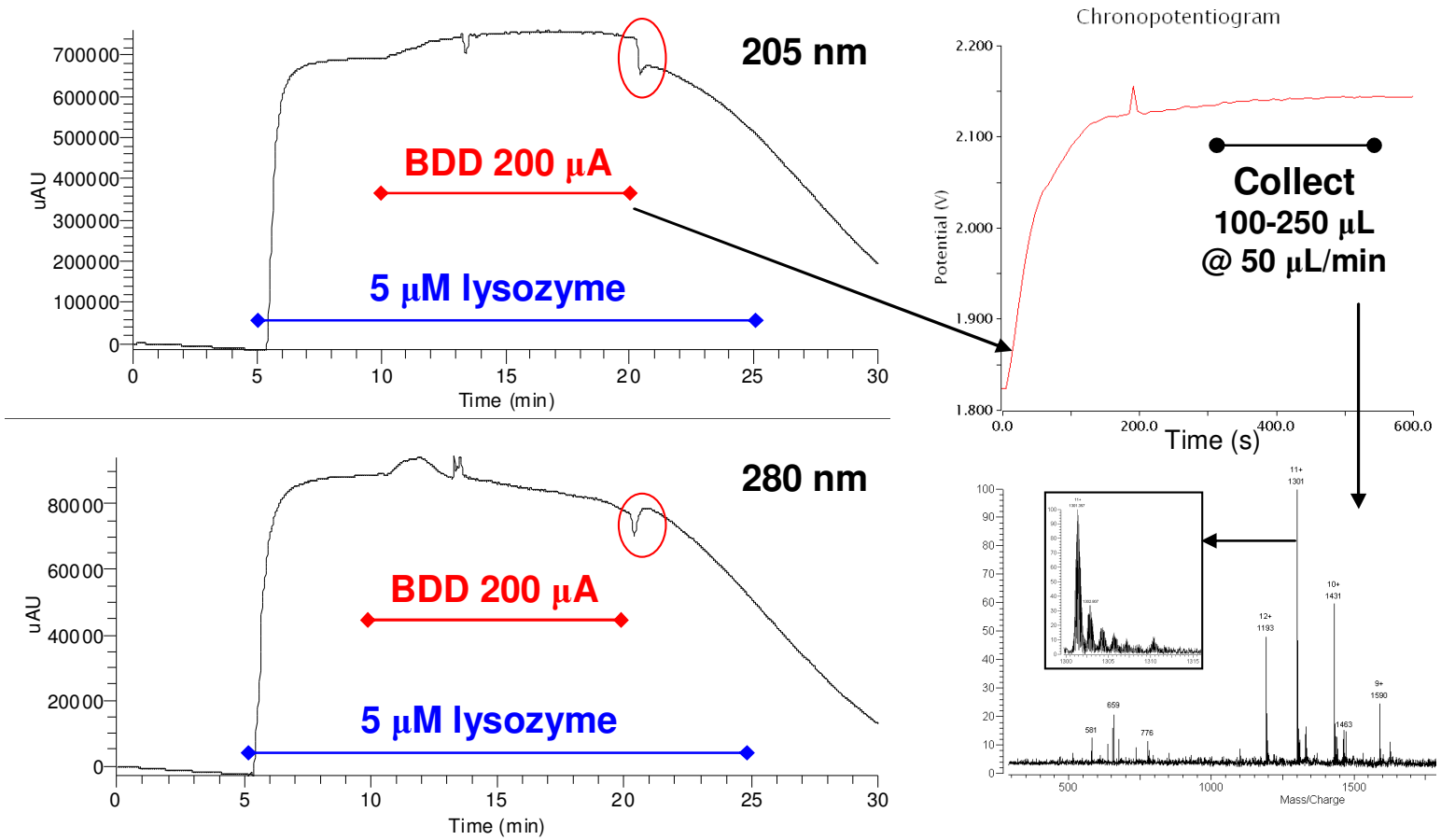


Figure 4-2: Workflow for monitoring protein concentration with the UV detector to detect any change during electrode activation.
Baseline absorption of potassium sulfate solution is obtained prior to injecting protein solution into the electrochemical cell flow chamber over 20 min duration. During the protein injection, the BDD electrode is activated (noted in red), and oxidized protein samples are collected once a steady state is reached, as evidenced by a voltage plateau in the chronopotentiogram at top right. Samples are analyzed by FTICR-MS for extent of oxidation revealed by expansion of mass/charge region near native protein peaks, as shown in MS inset at bottom right.

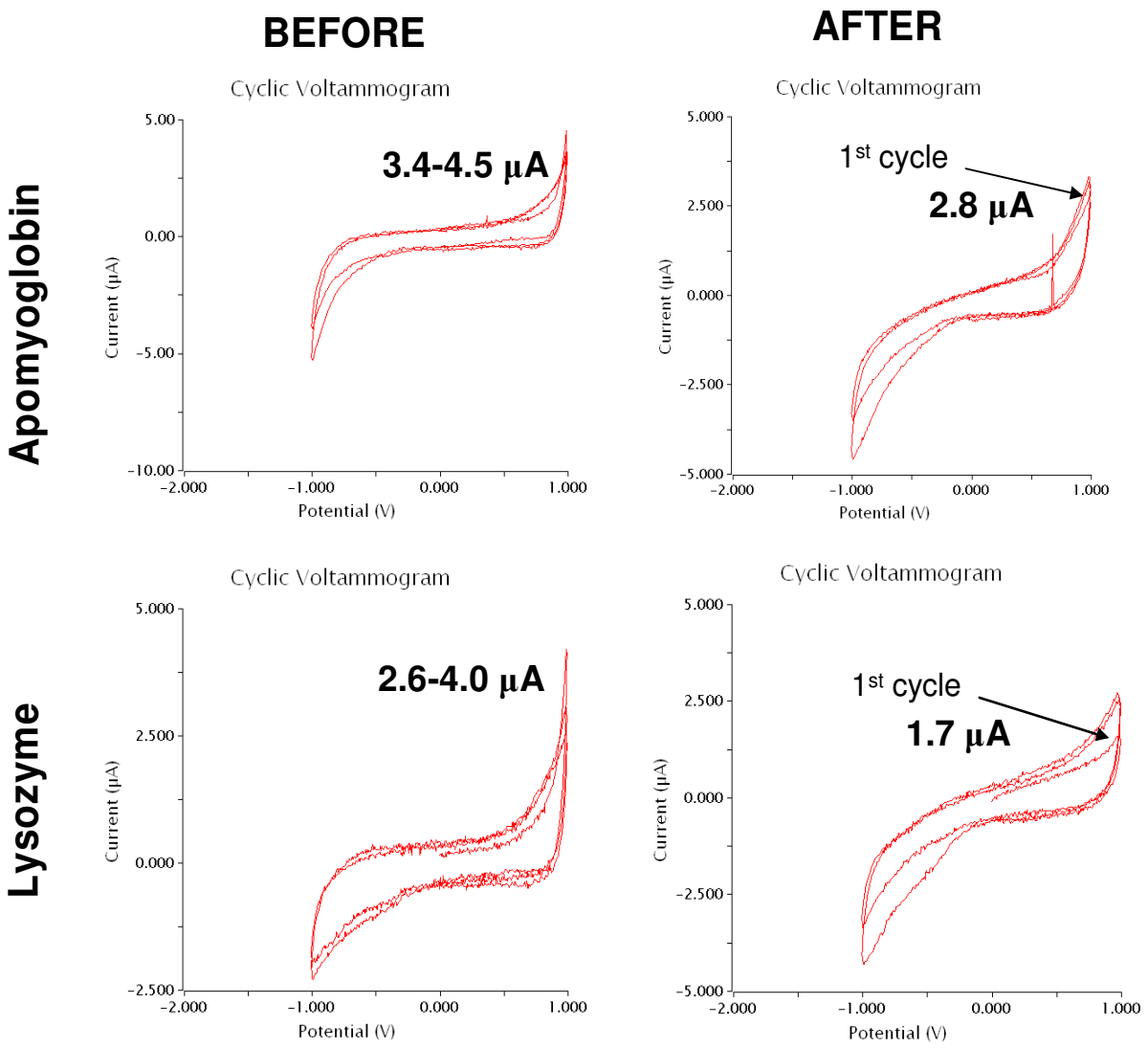


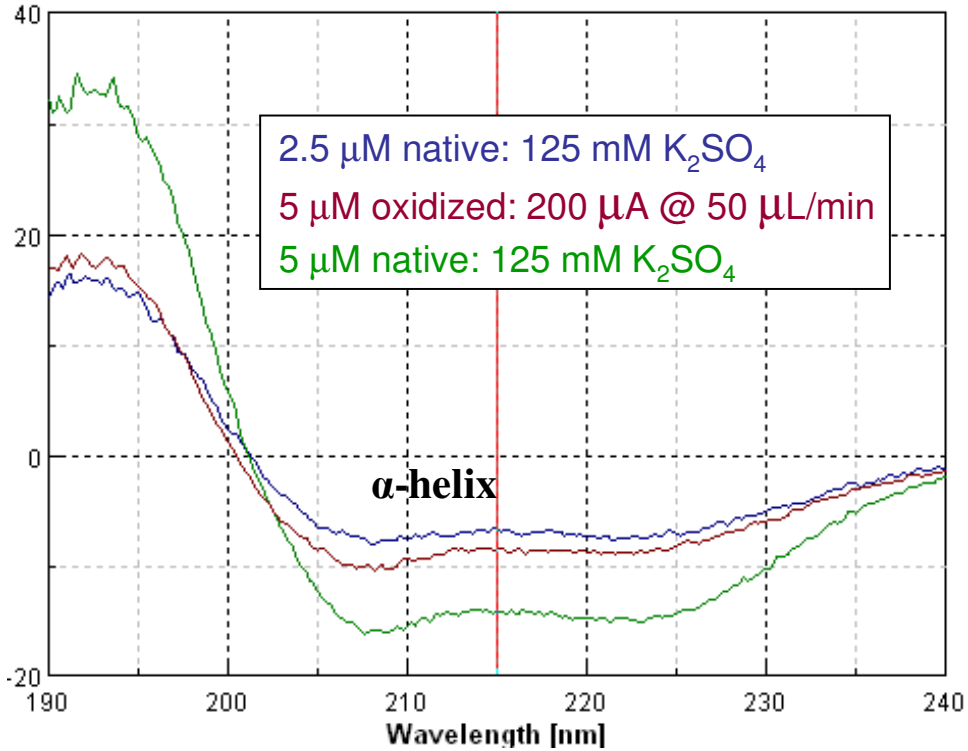
Figure 4-3: Monitoring electrode changes with CVs performed before and after protein oxidation. Relative to CV scans performed prior to extended activation of the BDD electrode in the presence of protein, the decreases observed in current magnitude at the voltage endpoints (-1.0 V and +1.0 V) as noted in black text above suggests that some level of adsorption has occurred for both apomyoglobin and lysozyme. Relative to the first cycle, apomyoglobin CV shows little change across successive CV cycles while lysozyme appears to show some resolution with subsequent CV voltage cycling.

This suggested that the electrode became insulated after an experiment with apomyoglobin, and was relatively unchanged by three cycles of voltage oscillation from -1.0 V to +1.0 V. In contrast, only the first cycle after a lysozyme experiment reveals evidence of any change to the electrode surface and subsequent cycles appear to restore the electrode to its prior condition. These suggested that a fraction of the apomyoglobin was irreversibly adsorbed onto an electrode, and that perhaps lysozyme is reversibly adsorbed to some extent. Because these proteins are both positively charged at pH 6.0, it stands to reason that they interact with the negatively charged cathode instead of the BDD anode. However, the pI of apomyoglobin is much closer to the solution pH, so its postulated interaction may be more hydrophobic than electrostatic. Over time it was observed that the voltage produced by a specified current level was consistently increasing, albeit in small increments of roughly 100 mV per experiment. This is thought to be due to some insulation of the electrode surface after its activation, despite whether the experiment involved protein or electrolyte only. For this reason, radical scavengers that are commonly used by other oxidative mapping groups to prevent spurious methionine oxidation events were not used in these experiments to circumvent possible electrochemical complications due to small molecule adsorption. The same rationale was applied in the decision not to use glutamate to reduce the likelihood of proteins encountering hydroxyl radicals after potential oxidation-induced unfolding.[44] Moreover, two of these model proteins were already shown to be structurally resilient to oxidative damage by circular dichroism (CD) measurements and intact protein mass spectrometry.[30] This latter metric was used for gauging structural integrity since CD is sensitive to any change in protein concentration which would interfere with structural interpretation of results.[13, 14] For example, apomyoglobin (5 μm ~ 0.08 mg/mL) stock was

measured by CD in 125 mM potassium sulfate, which was calculated by DichroWeb [94] to contain 45% α-helix and 0% β-strand secondary structures. A sample volume collected during the application of 200 μA current at a flow rate of 50 μL/min produced a spectrum that was calculated to contain only 23% helical structure along with a surprising 8% strand signal. The raw CD spectra are shown overlaid in Fig. 4-4. The table embedded in this figure reveals that by setting the input protein concentration incrementally lower, the helical percentage interpreted from the oxidized spectrum increased while the strand content decreased. In fact, half the original concentration (0.04 mg/mL) gave very similar values to the stock solution. The protein concentration was not reliably measurable by 280 nm absorbance since tryptophan is a highly reactive target and its oxidation products have shifted UV absorbance profiles.[127] Examination of the intact mass spectrum revealed a primarily unimodal distribution of predominantly native, singly, and doubly oxidized species that was shifted toward the unoxidized form. It seems more likely that this sample suffered some decrease in concentration rather than a striking deviation from its native helical structure. However, there was very slight evidence of bimodality along with oxidation of satellite peaks, so the conclusion remains ambiguous as to the degree of CD signal change caused by either altered concentration or structure.

4.4 Mass Spectrometry Analysis of Intact Protein Oxidation

After passage through the activated electrochemical cell, proteins were measured by ESI-FTICR-MS for oxidation level as function of voltage-dependent •OH flux and exposure time. Intact proteins consistently showed an increase in overall number of oxidation events when either flow rate or protein concentration was decreased, or when current magnitude was



Concentration	Helix (%)	Strand (%)	Disorder (%)
80 µg/mL	**44.5**	**0.0**	**20.3**
80 µg/mL	23.1	7.7	26.4
70 µg/mL	26.5	6.2	26.7
60 µg/mL	32.6	2.7	23.7
50 µg/mL	39.7	2.1	21.0
40 µg/mL	48.6	0.9	20.4
30 µg/mL	57.9	-1.2	15.0

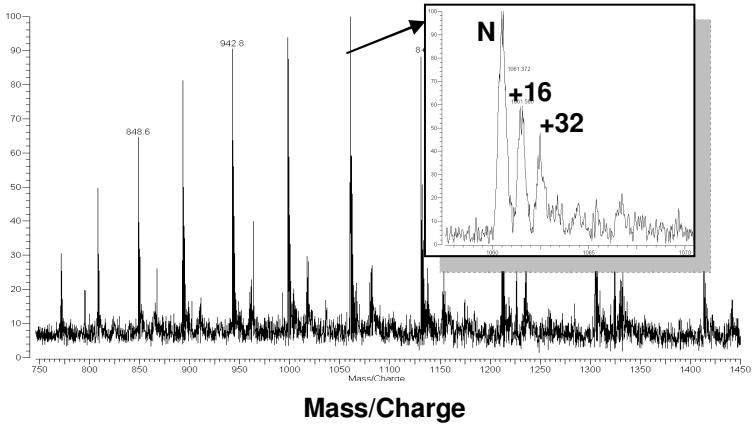


Figure 4-4: Monitoring change in protein structure and/or concentration with CD data for unoxidized and electrochemically oxidized apomyoglobin. The color-coded raw CD spectra show a strong α-helix signal for all three cases, while intensity for the oxidized sample most resembles the unoxidized or "native" sample at half the original concentration. Spectral data was submitted to DichroWeb for analysis of secondary structure content, and as shown in the top right table, the results for apomyoglobin oxidized at 5 µM (80 µg/mL) concentration are most consistent with the results of the 2.5 µM (40 µg/mL) unoxidized form. The mass spectrum on bottom right depicts oxidized apomyoglobin, with the inset 17+ charge state revealing at least two +16 Da oxidation products.

increased. Exposure to moderate oxidizing conditions usually resulted in a unimodal distribution of intact protein oxidation products, while more extreme oxidizing conditions typically produced a multimodal distribution (Figs. 4-5, 4-6, and 4-7). Such multiple distributions in oxidized samples are thought to be representative of different folded conformations.[30] A singly oxidized protein suggests the oxidation occurred within one snapshot of its dynamic solution-phase behavior, without additional events that may have occurred after some oxidation-induced structural change.

Ubiquitin revealed a bias toward the mono-oxidized form in a variety of experimental conditions, which remained the base peak upon over-oxidation (Fig. 4-5). However, even a sample that appears primarily mono-oxidized may contain an ensemble of over-oxidized proteins, as the most relatively abundant mono-oxidized form may be masking all other species present. The N-terminal methionine is the solitary sulfur-containing residue present in ubiquitin, and its sulfoxide form may account for a large proportion of the mono-oxidized protein signal intensity. Likewise, apomyoglobin shows a unimodal distribution in the moderately oxidized sample (300 μA at 50 μL/min) centered around +16 Da with flanking peaks representing the native and di-oxidized forms (Fig. 4-6). With heavier oxidation, additional distributions appear starting at greater masses than the di-oxidized species, suggesting the presence of partially unfolded conformers that expose more reactive sites. Lysozyme has ten sulfur-containing residues, including eight cysteines which can each be triply oxidized (+48 Da), so this mono- or di-oxidized phenomenon is not readily observable due to the large number of mass shifts possible (28 events of +16 Da each) on these residues alone (Fig. 4-7). It is possible that use of scavenger small molecules could result in oxidation product distributions for ubiquitin and

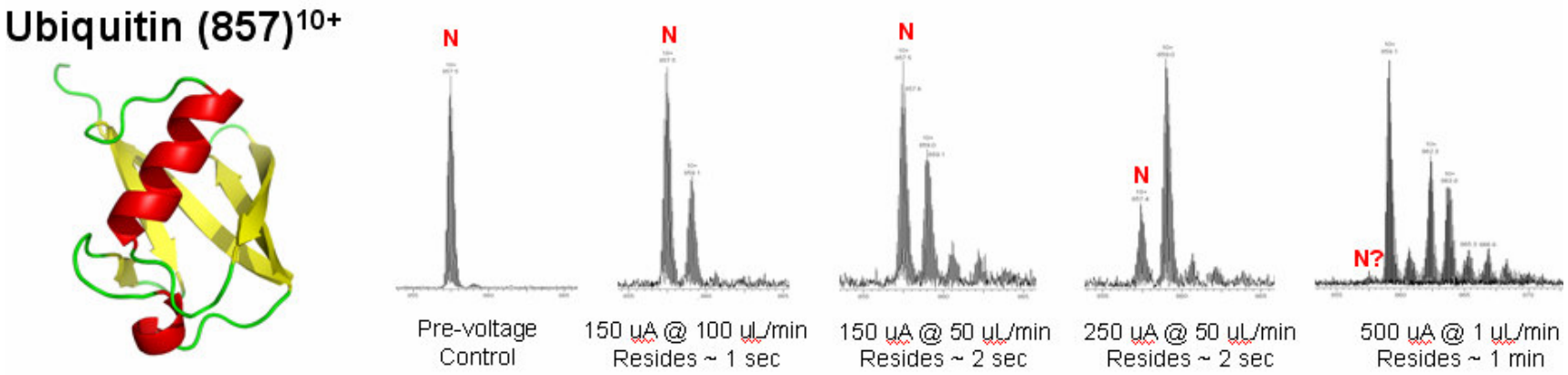


Figure 4-5: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. Each spectrum zooms into the 10+ charge state of ubiquitin around mass-to-charge 857. Adjustable parameters include current (•OH concentration) and flow rate (residence time for protein inside electrochemical flow chamber). Peaks occurring at higher mass-to-charge are +16 Da heavier than the native molecular mass (8565 Da). Note the bimodal distribution apparent in the over-oxidized sample at far right, suggesting that the portion of the ubiquitin ensemble with more than one oxidation event has undergone structural changes relative to the mono-oxidized form.

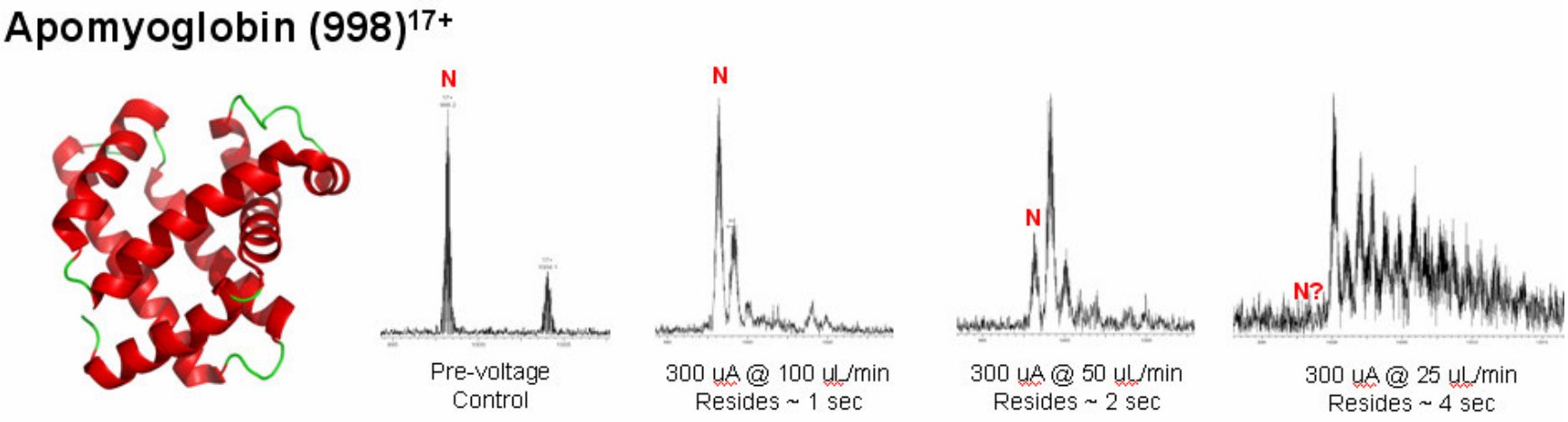


Figure 4-6: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. Each spectrum zooms into the 17+ charge state of apomyoglobin around mass-to-charge 998. Adjustable parameters include current (•OH concentration) and flow rate (residence time for protein inside electrochemical flow chamber). Resolvable peaks occurring at higher mass-to-charge are +16 Da heavier than the native molecular mass (16951 Da). Note satellite peak present in control spectrum that accumulates oxidation with increased exposure.

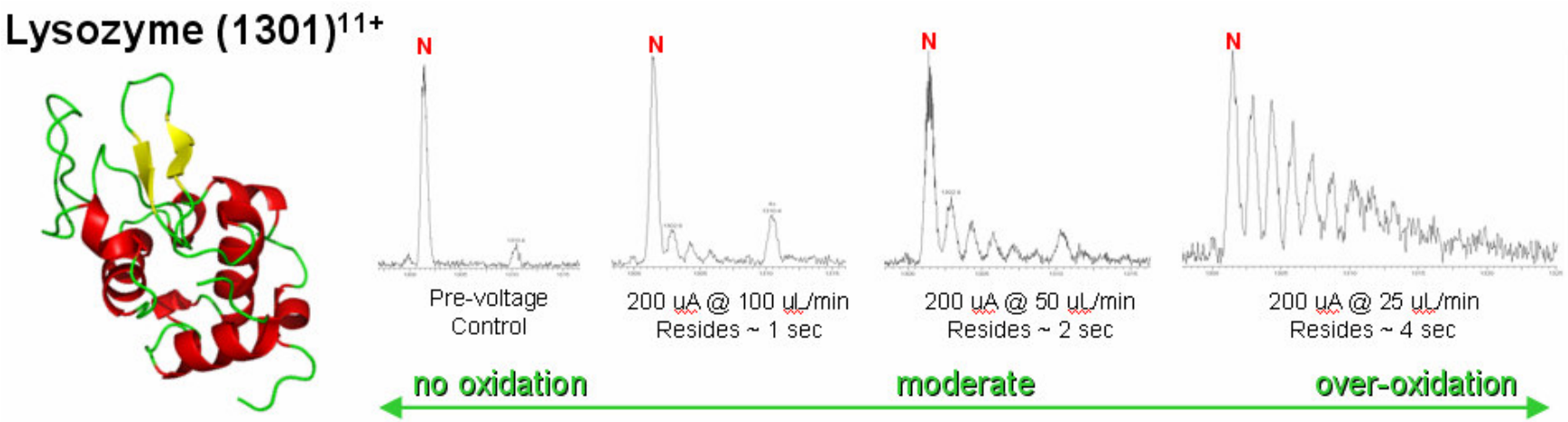


Figure 4-7: Native (N) and oxidized protein FTICR-MS showing oxidation level induced by the experimental conditions reported below each spectrum. Each spectrum zooms into the 11+ charge state of lysozyme around mass-to-charge 1301. Adjustable parameters include current (•OH concentration) and flow rate (residence time for protein inside electrochemical flow chamber). Peaks occurring at higher mass-to-charge are +16 Da heavier than the native molecular mass (14306 Da), which comprises four disulfide bonds. Note satellite peak present in control spectrum that accumulates oxidation with increased exposure.

apomyoglobin that show less agreement with the number of sulfur-containing residues and more compliance with first-order kinetics dependent on exposure to oxidizing conditions. However, these electroactive residues may be targeted by both indirect (peptidyl hydroperoxide or radical mediated) and direct (electron transfer) oxidation mechanisms.

4.5 Peptide Separation Strategies

Most amino acids are susceptible to oxidation by the hydroxyl radical, and often multiple mass shifts are possible for any particular residue representing different oxidation products.[72] Tryptic peptides typically contain more than one reactive amino acid, so this means that an oxidation experiment on a pure single protein suddenly becomes a complex mixture when all the reactive sites are considered in many different combinations. Peptides are typically sampled proportional to their relative abundance,[118] creating a dynamic range problem for detection of peptides modified at less reactive sites and in unique mass shift combinations. Of particular interest to this experimental consideration is the fact that reactivity rates measured for free amino acids span four orders of magnitude, from the most reactive (Cys ~ 3.5×10^{10} reactions per mole per second) to the least reactive (Gly ~ 1.7×10^{7} reactions per mole per second).[72] The significance of this observation is that while less reactive residues should be oxidized less frequently than more reactive sites, the information is no less important to mapping solvent accessibility. Therefore, a two-dimensional (2D) miniature MudPIT [74] chromatographic strategy was employed over a longer experimental timescale with the goal of increasing the breadth of peptide sampling relative to a simple reversed-phase (1D) separation. Oxidation events typically increase peptide polarity, resulting in a reduced affinity for reversed-phase resin

and thus a lower chromatographic retention time. Oxidation can also change peptide electrostatic nature by causing the loss of both acidic (decarboxylation, D/E -30 Da) and basic (deguanidation, R -43 Da) functional groups,[72] which may alter peptide charge and/or length due to missed proteolytic sites. It has been found that hydrophobic interactions between peptides and strong cation exchange (SCX) resin are remedied by the addition of acetonitrile to the mobile phase,[127] so this phase is separating primarily on affinity for an ion exchange resin. Derivatization of oxidation sites with a cationic adduct could dramatically increase the effectiveness of this additional chromatographic dimension, in addition to facilitating the enrichment of oxidized peptides.[128, 129]

Both 1D- and 2D-LC approaches have strengths and weaknesses, with potentially biased tendencies in peptide sampling. The 1D approach separates peptides based on a single characteristic, requires a reasonable amount (1-3 hours) of instrument time per sample, and produces a single chromatogram that makes relative quantification of peptide peak elution areas a rather straightforward task. However, the smaller chromatographic timeframe and limited separative capacity hinder the dynamic range of 1D analysis, which impacts finding uniquely modified peptides within the complex oxidized mixture. The 2D approach offers separation based on two orthogonal physical properties, hydrophobicity and electrostatic charge, to theoretically reduce the overall complexity of peptides eluting at any given retention time and increase the chances for detecting uniquely modified peptides that are likely present in very low abundance. The longer 2D experimental timeframe (5-10 hours) offers more measurements thus higher peptide spectral counts to gauge statistically relevant estimation of abundance relative to unoxidized counterparts. One caveat to the 2D approach is that with a deeper probe of the

sample contents, more oxidations are identified in tryptic digestion controls that are not ordinarily measured from the same sample in 1D analysis. Potential causes of and remedies for these background oxidation events are discussed below and warrant consideration relative to oxidation events identified from experimental samples.

The widely accepted method of relative quantitation for oxidative mapping involves integration of native and oxidized peptide peak intensities with selected ion chromatograms (SIC) to achieve a relative measurement.[33] There are three major drawbacks to this approach: the first is that SICs provide information at the peptide level rather than for individual amino acids, which is the purported strength of oxidative mapping over HDX in addition to its critical covalent mass tag which does not suffer from back-exchange. This means that parent ion SIC cannot discriminate between differentially oxidized versions of the same peptide whether they coelute or vary in their retention times, so the tandem mass spectra must be relied upon for site identification if this information is desired. Secondly, an overwhelming number of possible mass shifts make an exhaustive analysis cumbersome, so just the most common mass shifts (+16*n* Da) are typically considered. One solution used to counter these issues involves monitoring depletion of the unoxidized form in the experimental sample relative to the control digest against the relative intensity of an internal standard.[36] It is not clear whether this approach took into account the possibility of a decrease in unmodified form within the oxidized peptide mixture simply due to cleavage sites missed during proteolysis. Interestingly, these researchers used the mono-oxidized value as the "unoxidized" form when the myoglobin-derived peptides contained methionines (< 10 $Å^2$ exposed) due to minuscule levels of the native isoform, despite using glutamine scavenger during the experiment. These chromatographic extraction

methods were not performed in our study as the goal was a high-throughput pipeline that relied on spectral counts for estimating relative abundance. However, the most effective strategy may involve optimizing a shorter 2D approach based on the elution profile of the peptides in the control digest and building this into a single instrument method for ease of quantifying disappearance of the native peptide peak relative to a spiked internal control. This would allow a metric for comparing oxidized amino acid statistics derived from tandem mass spectra within specific peptides to the parent ion peak intensities measured from full scan mass spectra, which are not subject to the sampling directives introduced by dynamic exclusion as discussed in the following section.

4.6 Entry Points for Bias in Data Acquisition

Experimental variables can impact the actual measurement of peptides containing oxidized sites, which would affect the resulting tandem mass spectra available for oxidative data mining discussed in the next chapter. These parameters include the experimental methods used to handle protein or peptide solutions, as well as instrumentation options utilized to acquire spectra in a data-dependent fashion. Examination of these entry points for sampling bias could lead to protocol improvements that allow for impartial peptide sampling frequencies.

4.6.1 Liquid Chromatography Strategy

The increased overall sampling inherent in 2D analysis relative to 1D could foster a semi-quantitative approach for calculating the occurrence of oxidation events at specific sites as a ratio of modified spectral counts to the total observed for any given peptide. However, search results showed a significant discrepancy in spectral counts among tryptic peptides in 2D experiments.

While it is common to see small hydrophilic peptides missing from the sequence coverage, some peptides present in 1D were either overwhelmingly represented or noticeably lacking in the 2D results. For example, one abundant ubiquitin peptide (12-TITLEVEPSDTIENVK-27) had a relatively tight peak elution profile (~2 min) in 1D that was found with similar retention characteristics (~32-37% solvent B) in all three salt pulses. The first two salt pulses were followed by a slower gradient of 5 to 50% solvent B, in contrast to the 5 to 100% gradient used in both 1D experiments and after the last 2D salt pulse. Consequently, this peptide experienced a ten-fold increase in sampling (from ~20 to over 200) with the 2D strategy and novel sites of modification were detected in control samples, as discussed below (Section 5.7). In contrast, the N-terminal neighboring tryptic peptides (1-MQIFVKTLTGK-11) from the same protein were typically sampled less than 15 times in control 2D samples, and InsPecT reported them absent altogether from the experimental 2D samples. Since no extraction step was performed before proteolytic peptides were loaded onto the analytical column, it stands to reason that this apparent sampling bias could be due to differential desorption from the chromatographic resin within the split-phase column. In other words, irreversible adsorption of some peptides to either of the solid phase resins employed may have occurred. Another possibility is that some peptides adhered to walls of the reaction tube, resulting in variable abundances present in the final mixture. This might be attributed to the acidification step to quench the enzymatic activity of trypsin prior to lyophilization, wherein some peptides may become insoluble upon hitting their isoelectric points during the rapid pH drop. Due to the small volumes and low concentrations used, such a phenomenon might not visible to the naked eye. Peptide FTICR-MS scans of a control ubiquitin tryptic digest reveals that the monoisotopic peak at $(894.5)^{2+}$ is the most

abundant ion and thus the base peak of the spectrum, which corresponds to the neutral mass of the aforementioned "TITLE" peptide at 1787 Da. However, the proteolytic mixture was extracted by C_{18} ZipTip prior to mass spectrometric analysis, so this measurement was also subject to potential retention bias posed by reversed-phase material. Nevertheless, increased sampling of this abundant peptide along with its oxidized products has revealed interesting information about oxidation events detected in both control and electrochemically oxidized samples, as described below (Sections 4.8 and 5.4).

4.6.2 Mass Spectrometry Approach

Dynamic exclusion (DE) is a powerful feature of data-dependent acquisition of mass spectra, wherein scans of the full mass range (MS1) are referenced to direct the subsequent isolation and fragmentation (MS2) of the most abundant parent ions. After measurement of a particular parent ion mass a certain number of times (repeat count) within a particular timeframe (repeat duration), this value is added to an exclusion list to avoid excessive sampling of abundant ions. Enabling DE has been shown to positively impact detection of low abundance species from proteomic peptide mixtures, with an optimal exclusion duration of roughly 90 seconds.[130] Optimization of this data acquisition feature is critical because oxidized peptides can occur in many different varieties, so each version is typically present in very low abundance. Additionally, the most common oxidation event involves incorporation of an oxygen atom (+16 Da), effectively increasing the peptide polarity and thus hydrophilicity. While this predictably reduces affinity for reversed–phase resin, it also results in localization of oxidized species to the interior of electrosprayed droplets which suppresses their ionization, thereby resulting in lower

intensity parent ions.[73] With all of this considered, oxidative experiments were run with DE set to add a given parent ion mass to the exclusion list for 30 sec if it was sampled twice within a 30-sec repeat duration. A shorter exclusion duration was chosen (relative to the abovementioned proteomic study) to avoid temporarily excluding all peptide signals resulting in isolation of noise peaks, since the oxidized mixture is complex but not to the level contained in a proteome analysis. The reason for setting the repeat count to two was to facilitate redundant measurements of low abundance oxidized species for improved confidence in oxidation site identifications relative to single spectral count matches. However, it could be argued that the repeat count should be set to one with a shorter exclusion duration in order to speed the subtraction of more abundant species and rapidly dive down to sampling lower intensity parent ions.

4.7 Importance of Control Measurements

All experiments benefit from control measurements to delineate features intrinsic to the subject of interest versus those introduced by steps in the experimental protocol. In the case of oxidative mapping, it is important to identify any endogenous oxidation events to avoid false positives, and would be ideal to flag all possible oxidation sites to minimize false negatives. Controls for the proteolytic digestion can help the experimentalist identify problem areas when peptides are found missing or large sampling discrepancies exist, since oxidation of basic sites can remove the charged functional groups[72] resulting in missed cleavage sites. Additionally, control data on the LC-MS/MS strategy allows gradient optimization for measuring chromatographically unique peptides from the protein of interest. Finally, tandem mass spectra of unoxidized peptides allows manual evaluation of peptide-specific (length, hydrophobicity, etc)

scoring in determining appropriate spectral quality filter levels, with the caveat that oxidized peptide dissociation may result in variable intensities of resulting fragment ions relative to their unoxidized counterparts.[131] The most reliable filter is manual evaluation for all peptide matches, but this labor intensive task for 1D data would be even more burdensome for routine 2D data analysis, especially if replicate measurements are obtained. Therefore it may be advantageous to combine the unmodified SIC depletion metrics with statistically relevant results of tandem mass spectral searches filtered based on manually evaluated peptide-specific score thresholds from the native digest for a relative comparison of site-specific oxidation extent in experimental samples. While in retrospect the proposed strategy may be very effective, this study focused only on comparing peptide spectral counts for the two chromatographic approaches.

4.7.1 Negative Control to Identify Background Oxidation

It is critical that search results of tandem mass spectra from oxidative mapping experiments are compared against those from a control unoxidized proteolytic digestion, as oxidation may unintentionally occur at some point in the process. The first potential source of this so called "background" oxidation is during preparation of commercially available purified proteins. Commercial hemoglobin was found to contain multiple oxidation events (+16 and +32 Da) on primarily methionine-containing peptides, but the authors conceded that additional modification types must exist since α-globin (single Met) showed a +48 Da signal, albeit weakly.[132] Lower abundance non-Met oxidized peptides were not measured by tandem mass spectrometry of the course of a 60-min reversed-phase gradient, though this detail might have

been resolved with a longer gradient and/or multidimensional separations. It is not surprising that hemoglobin proteins would show some level of endogenous oxidation since binding and transporting O_2 is their primary function. However, multiple oligomeric and oxidized species were observed for commercial hemoglobin that were not found in the much cleaner spectrum from freshly purified hemoglobin. After lyophilization followed by reconstitution and freezing, the thawed sample of freshly purified hemoglobin mirrored the commercial stock. Similar to hemoglobin, the presence of the iron-containing heme prosthetic group in apomyoglobin may have contributed to the background oxidation prior to its ejection.

In the case of oxidation appearing in commercial proteins, the intact protein mass spectrum should reveal measurable satellite peaks aside from the expected native protein signal near average molecular mass. However, if very concentrated samples are tested, or long collection times are allowed in the ion trap, then the native species could be overly abundant. In either case, the resulting measurement could suffer from dynamic range limitations such that lower abundance modified protein satellite peaks are not visible relative to the overwhelming native signal. Additionally, oxidized species are more polar and thus may not ionize as well as their unoxidized counterparts.[133] Satellite peaks were present for the proteins tested in this study, but the peaks observed are typically ~100 Da heavier, as in the case of lysozyme (Fig. 4-8) and apomyoglobin. Low abundance (< 5% intensity) peaks near the primary satellite peak have occasionally been observed that appear to be in increments of +16 Da. One exception to this is the appearance of a low level of mono-oxidized ubiquitin can be seen in Fig. 4-5, just to the right

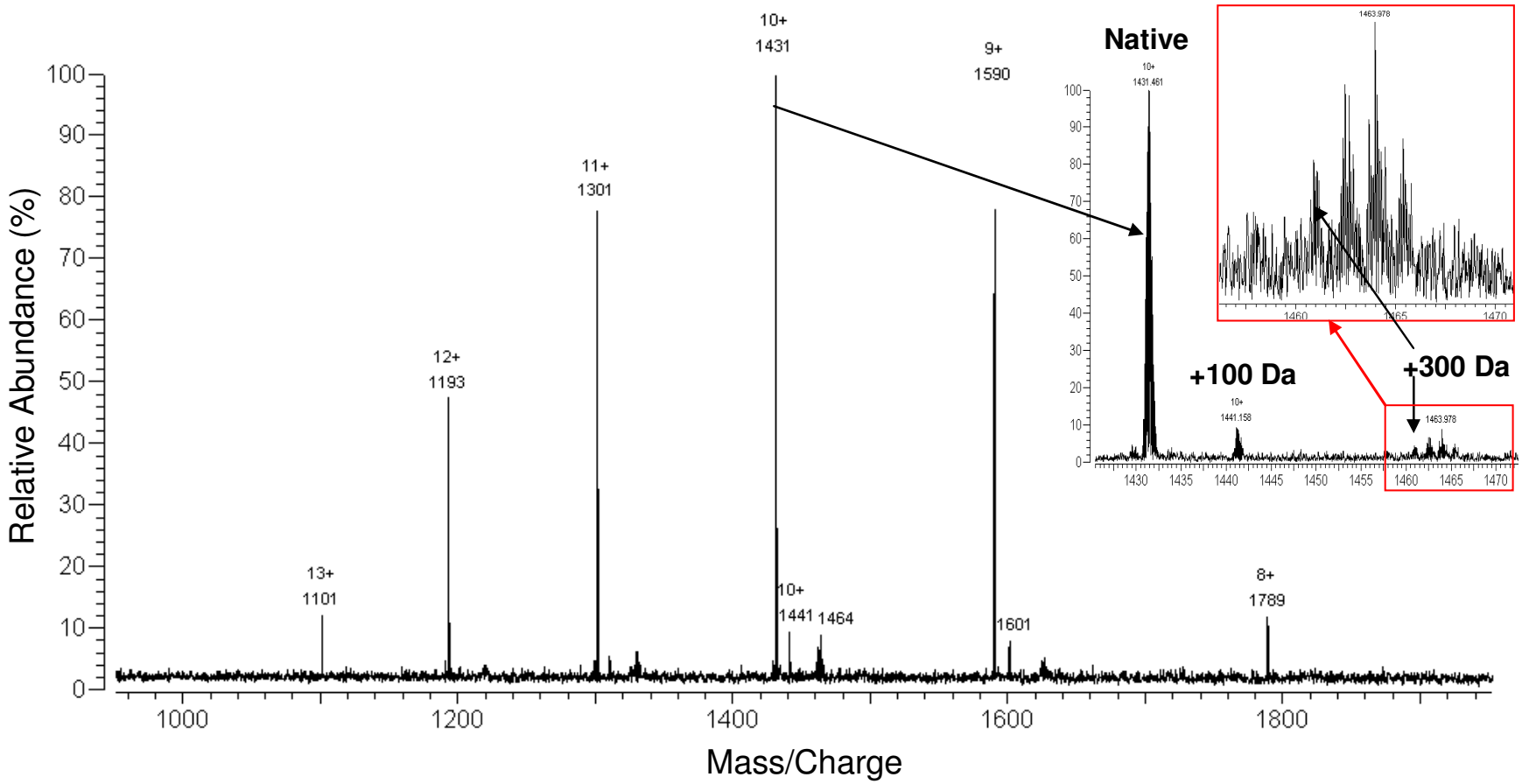


Figure 4-8: Lysozyme FTICR-MS from freshly prepared, commercially-purified stock protein with satellite peaks showing oxidation signature. Inset zooms into a closer view of peaks that appear to be oxidation events present on the +300 Da species, which is an unknown modification but presumably an isoform of lysozyme since the satellite peak is present for the most abundant unoxidized charge states.

of the native species. However, this could be an artifact of the charge balancing redox that occurs during electrospray ionization from a metal capillary,[86] likely resulting in the direct oxidation of methionine on a small fraction of proteins.

Another potential source of oxidation is exposure to oxidants is during proteolysis. One method to reduce the content of molecular oxygen (O_2) is by sparging the reaction solution with an inert gas such as nitrogen or argon. While it is not known whether radicals (hydroxyl, superoxide, etc.) are present in solution during the digestion process, the use of a radical-scavenging antioxidant such as dithiothreitol, glutathione, or free methionine could counter this minor threat. Hydrogen peroxide formed by collision of two hydroxyl radicals can be decomposed to water and molecular oxygen by the catalase enzyme. Peroxides reportedly cause two-electron oxidation of methionine independent of solvent exposure, and can be cleaved into hydroxyl radicals either by trace iron present in solution or ultraviolet light emitted by indoor fluorescent bulbs.[134] Furthermore, the tryptophan oxidation product kynurenine can act as an oxidative photosensitizer when bound to protein and exposed to far-UV wavelengths, catalyzing the oxidation of tyrosine to crosslinked dityrosine and 3,4-dihydroxyphenylalanine.[135] A third possible source of background oxidation mentioned previously is the charge-balancing electrochemical redox reactions that occur during electrospray ionization. However, this seems less plausible in nanospray LC-MS/MS analysis since peptides do not pass over the gold working electrode on their way to the mass spectrometer through fused silica tubing, not a metal capillary.

The appearance of oxidation events in the control digestions could be taken as false positives since they were not induced and thus were not expected to be found at significant levels. One seemingly reasonable option for subtracting background oxidation from

experimental results is to filter peptide matches based on the scores for oxidized peptides appearing in the control sample. However, this is effective only if the assumption is true that background oxidation in the control sample is false. In the case that background oxidation is real, then the assumption is false, and good match quality spectra would be discarded with such conservative filters. This perspective is supported by the finding of oxidized methionine and tyrosine at significant levels, as well as a few spectra showing hydroxyproline,[78] in the well-known ISB (Institute for Systems Biology) data set containing tryptic peptides of commercially available reference proteins.[136] For the abovementioned reasons, spectra matching peptides for both control and experimental samples were filtered by the same criteria, as discussed in detail within Chapter 5. By subjecting control digestions to 2D measurements and oxidative searches with InsPecT software, a survey of background oxidation either already present in commercially available proteins or induced by sample preparation steps is reported below.

4.7.2 Positive Control to Obtain List of All Oxidation-Sensitive Sites

In theory, generous exposure of protein to oxidizing conditions would promote oxidized-induced unfolding, resulting in protein denaturation and all susceptible amino acids would become oxidized. Consequently, sites that are resistant to hydroxyl radical attack would emerge as potential false negatives in moderately oxidized samples in the absence of the positive control. However, there are difficulties with this approach that are discussed in greater detail within Chapter 5. Briefly, these challenges include increased false positives concomitant with an extreme expansion of database search space, and confident identification of oxidized species

when modification occurs near peptide endpoints or more than once on a single peptide. Finally, over-oxidation could primarily target sulfur-containing and aromatic residues once exposed due to oxidation-induced unfolding, leaving other reactive residues unmodified.

4.8 Survey of Modified Sites in Control and Oxidized Proteins

A survey of oxidized sites was obtained with the use of InsPecT [77] software to search for more than forty oxidation mass shifts (Table 3-1). The overall assessment was similar to the one performed in Chapter 3, except that an even more extensive mass shift list was employed (Table 2-1) along with a larger database based on advice from the software curators. This broad interrogation mirrored the oxidative modification list developed for ByOnic software, which was recently used in another oxidative mapping study.[92] A few caveats to this comparison of InsPecT results include the number of modifications allowed per peptide and the size of the databases employed for each search. The earlier study allowed only one mass shift based on cautious advice from InsPecT curators, while the later search allowed three to emulate the more liberal ByOnic settings. Additionally, there is a significant difference in the size of the databases referenced by each InsPecT search, from twelve initially to over one thousand sequences in order to facilitate improved p-value assignments. Admittedly, a fairer comparison would involve changing only one parameter at a time, for instance, using the same small database for both searches while expanding the mass list for only one of the two searches. Post-processing and preparation of search results is a fairly work intensive process, involving a variety of script-based commands for each experiment, followed by extraction of spectral counts into a table format.

These searches were conducted based on advice received from InsPecT curators for future publication purposes with the comparison of results for the previously published data set as an afterthought, but a discussion of these results is nevertheless relevant to the reasons for exploring ByOnic software. This comparative discussion is followed by a general analysis of oxidation sites identified by InsPecT for control versus electrochemically oxidized proteins followed by 2D separations of tryptic peptides.

These newer search results showed a markedly changed oxidation report for the same 1D data on searched in Chapter 3. In contrast to the ubiquitin oxidation sites reported in Table 3-2, recent search results (Table A-4) show additional modification sites at Phe4, Ile13, Ile30, Gln31, while no oxidation was reported on either Thr12 or Pro19. Residues identified in both searches include Met1, Thr14, Leu15, Ile36, Pro37, Pro38, Lys48, Tyr59, and His68, although spectral counts tended to be lower in the new search. The new search conditions turned up more modified sites in the 1D lysozyme data (Table A-5), for both the +2.05 V and +2.10 V experiments, while only two residues were consistently found modified in both experimental conditions reported: Cys6 and Trp123. In the +2.05 V results, new sites included Arg5, Glu7, Ser24, Trp28, Cys30, Ser91, Cys94, and Trp108. In the +2.10 V results, new modification sites were revealed with single spectral counts for Phe3, Ala11, Phe34, Glu35, and Cys115, along with 3-5 spectral counts for Asn106, Trp108, and Trp111. Some of these sites are considered buried in the static structures, but as discussed in the next chapter, these may be explained either by transient solvent exposure during solution-phase dynamics, or perhaps by virtue of challenges inherent in mining oxidation sites from mass spectrometry data.

There are also distinct differences in the reports obtained for control and oxidized proteins, which show that electrochemical oxidation under the new experimental conditions still generated modifications on a breadth of non-electroactive residues within the proteins of interest. The 2D approach produced significantly higher spectral counts (contained below in parentheses) for some peptides, which often fostered detection of more oxidized sites in those more abundant peptides. For example, the longest tryptic peptide in ubiquitin registered some background oxidation at Ile13 (1), Thr14 (1), Leu15 (2), Glu16 (1), Glu18 (1), and Lys27 (4). Theoretically, the amino acid sites identified as oxidized in the control should also be found in the deliberately oxidized sample. Oxidation events detected in the experimental (300 μA at 25 μL/min) counterpart to the aforementioned control do in fact include the same sites in addition to new ones: Thr12 (3), Ile13 (5), Thr14 (3), Leu15 (7), Glu16 (1), Glu18 (4), Ser20 (1), Thr22 (2), Val26 (3), and Lys27 (2). In contrast, the same peptide from a lighter oxidation exposure (150 μA at 50 μL/min) and run three months later on the same instrument revealed oxidation at a unique blend of *some* of those residues in addition to two new sites: Ile13 (1), Glu18 (2), Ser20 (2), Glu24 (1), Asn25 (1), and Val26 (1). Different protein stocks were used between the earlier and later experiments, so it is not surprising that this oxidized subset does not fully comprise the abovementioned control subset. This type of differential reporting illustrates the need for controls to be run with every experiment. Additionally, the variable results realized by these different searches run on the same data sets corroborated empirical findings that tyrosine-only versus liberal oxidative searches performed on bovine serum albumin (BSA), which was labeled by either tyrosine-selective or oxidative experiments, yielded different results about the frequency of tyrosine oxidation. However, this latter observation could be due to the assignment

of oxidized sites to tyrosine when in fact a neighboring site was modified, such that peptide fragment ions matched this "next best" oxidized version well enough. Nevertheless, it was admitted by the previous InsPecT curator that their software was designed to accommodate significantly fewer modifications than the complete oxidation set, so another software was utilized for comparison. ByOnic was a natural choice due to its presence in the field, so a more comprehensive view of the search results from both software packages are reported along with a discussion of informatic challenges in Chapter 5. Those results comprise experiments performed using the more recently developed electrochemical conditions discussed above.

4.9 Conclusions

The protocol first published in 2008 for conducting protein electrochemical oxidation was refined for a more rapid exposure to oxidizing conditions and better controllability of the resulting oxidation level. While some evidence of adsorption was reported above, this technology still holds some promise for mapping the dynamic surface of solution-phase proteins. Examination of the oxidized sites by multiple search softwares and evaluation of these results by molecular dynamics, as discussed in the next chapter, may shed light on whether a sufficiently deep probe of surface residues was accomplished. Regardless of this outcome, valuable data on control protein digestions revealed the need to use freshly purified protein to minimize background oxidation and increase likelihood of labeling native structure. However, any *in vitro* sample processing, including purification, lyophilization, and freezing, could impact structural studies since they may further the tendency for background oxidation to occur regardless of attempts to counteract such events. The benefit of increasing dynamic range with 2D

chromatography was demonstrated, albeit with room for protein-specific gradient optimization, by the identification of considerably more low-abundance oxidized peptide spectra relative to 1D measurement. Improvements in the overall methodology for electrochemical protein oxidation have furthered the potential for its use in oxidative surface mapping experiments, and generally expanded knowledge of how the BDD electrode may become valuable to future protein research.

5. Mining Oxidative Modifications from Multidimensional LC-MS Data for Increased Dynamic Range in Electrochemical Mapping of Protein Dynamics

This manuscript will be submitted to *Journal of Proteome Research* with the following authors:

McClintock, C.; Parks, J.M.; Ghatty, P.K.; Bern, M.; Kertesz, V.; Hettich, R.L.

5.1 Introduction

Mass spectrometry has become the lynchpin technology for the characterization of complex protein mixtures, such as proteomes identified from genomic sequence data. Proteomic studies seek to assess the host of proteins present in a particular sample by measurement of constituent peptides found within a database of protein sequences. Depending on the experimental design, protein abundance can arguably be estimated by a relative isotopic labeling approach [137] or by utilizing an absolute peptide spectral counting strategy.[138] Either of these quantification tactics would suffer from a pure database search due to the presence of PTMs, because a covalent modification of any given peptide produces a mass shift in many of the resulting fragment ions. Therefore, PTMs can impact the number of peptide tandem mass spectra that match the protein database, although their abundance may be relatively low. Because this biologically relevant information impacts overall protein identifications as well as quantification efforts, the mining of PTMs from proteomic data has taken center stage in recent years.[119]

A number of computational algorithms have been developed to tackle the challenge of detecting both expected and unanticipated PTMs, while also handling the intrinsic variability among peptide mass spectra. Oxidative mapping presents a unique challenge to the computational strategies developed for mining PTMs, not only due to the large number of possible mass shifts, but because uniquely oxidized peptide variants are generated at low abundances. Therefore, peptide mixtures were separated by either reversed-phase only (1D) or a split-phase MudPIT column (2D) for increased dynamic range and thus measurement sensitivity.[139] Mass spectral data were searched by two different software packages, InsPecT and ByOnic, for a suite of possible oxidative modifications. By using resolved crystal structures as templates for molecular dynamics (MD) simulations, we gained insight about transient solvent exposure for some buried sites in the static crystal structures. Oxidative modifications mined by both the InsPecT and ByOnic software packages are reported and compared with both average and maximum SASA values obtained from MD trajectories, as accurate detection of oxidized peptides continues to present a formidable combinatorial challenge.

5.2 Interpreting Tandem Mass Spectra of Native and Oxidized Peptides

Collision-induced dissociation (CID) of a parent ion typically produces a tandem mass spectrum (MS/MS or MS2) containing positively charged fragment ions from both the positively charged N-terminus (b-ions) and the C-terminus (y-ions).[140] Fragmentation by CID occurs at the weakest link (O=**C—N**) adjacent to the peptide bond, producing b-ions that contain the amino group and y-ions that often include a basic site (K/R) due to the cleavage specificity of trypsin.[141] The peak intensity pattern of CID fragments tends to favor a bell-shaped Gaussian

distribution, wherein mid-range fragment ions are most intense while outermost ions typically contribute much lower signals that may become obstructed by spectral noise. Consequently, both low and high indices (*i*) for both b_i- and y_i-ions may be difficult to interpret. For this reason, oxidized peptides with modification sites located on or near the peptide endpoints may produce spectra with residual uncertainty about the specific amino acid site affected.

5.3 Potential Sources of Information Loss

During the course of searching mass spectra for oxidative modifications, there are some pitfalls that can impact peptide identifications. These include the challenges with interpretation of spectra that involve multiple oxidation sites on the same peptide, unexpected modifications, and mixtures of more than one version of an oxidized peptide. Recognition of these difficulties does not imply they have been resolved, only that the results of any oxidative search should be viewed with these caveats in mind.

5.3.1 Multiply Oxidized Peptides

Identifying a singly modified peptide is a reasonable challenge, because the modified position is reflected by mass-shifted fragment ions along with their unmodified complementary ions from either terminus. For instance, the spectrum from an octopeptide (1-PEPTIDEX-8) modified at the 4th position (T+16) would theoretically show unmodified native masses for b_1-b_3 and y_1-y_4 ions, with masses shifted by +16 Da for b_4-b_8 and y_5-y_8 ions. In this case, the diagnostic mass-shifted fragment ions and their unmodified complements (b_3/y_5 [+16], b_4 [+16]/y_4) should show the transition to a shifted mass occurring at the 4th position. However, the situation becomes much more complicated when more than one mass shift occurs in a given

peptide. For example, when both peptide ends are modified, then oxidized fragment ions no longer correspond to an unmodified complement. Since the protein sequence is known, the primary challenge lies in the many possible combinations of mass shifts that can coincide among various reactive sites on the same peptide. One example of a triply oxidized ubiquitin peptide shown in Fig. 5-1 illustrates the need for allowing more than one modification per peptide to fully interrogate the oxidized peptide mixture. Interestingly, this spectrum appears to be correctly identified as the first two ubiquitin peptides M(+16)QIFVK(+14)TL(+14)TGK by InsPecT, but ByOnic matched this spectrum in a regular oxidative search to a low-scoring (below filter level) bogus myoglobin peptide: LFTGHP [+16]ET [+22]LEK. ByOnic uncovered the correct mass shifts as M(+16)QIFVK(+28)TLTGK with a high score in a blind search, but assigned both +14 Da modifications to Lys6. This presumably is an example of the differences among search algorithms, as "lookup peaks" were admitted by the authors to be less proficient in identifying peptides modified at or near both endpoints.[78]

5.3.2 Unanticipated Oxidation Events

Any modification that results in a stable mass shift for which provisions are not explicitly included in the search parameters will result in a fraction of missed peptide identifications. For example, ByOnic peptide match results commonly showed evidence of sodiation (+22 Da), which is a mass shift that was not considered in InsPecT searches. The use of a more conservative list generated by cutting less common mass shifts or less reactive residues may result in faster searches and fewer false positives. However, if these are excluded from search parameters when the modifications actually exist, the result will be lower peptide spectral counts

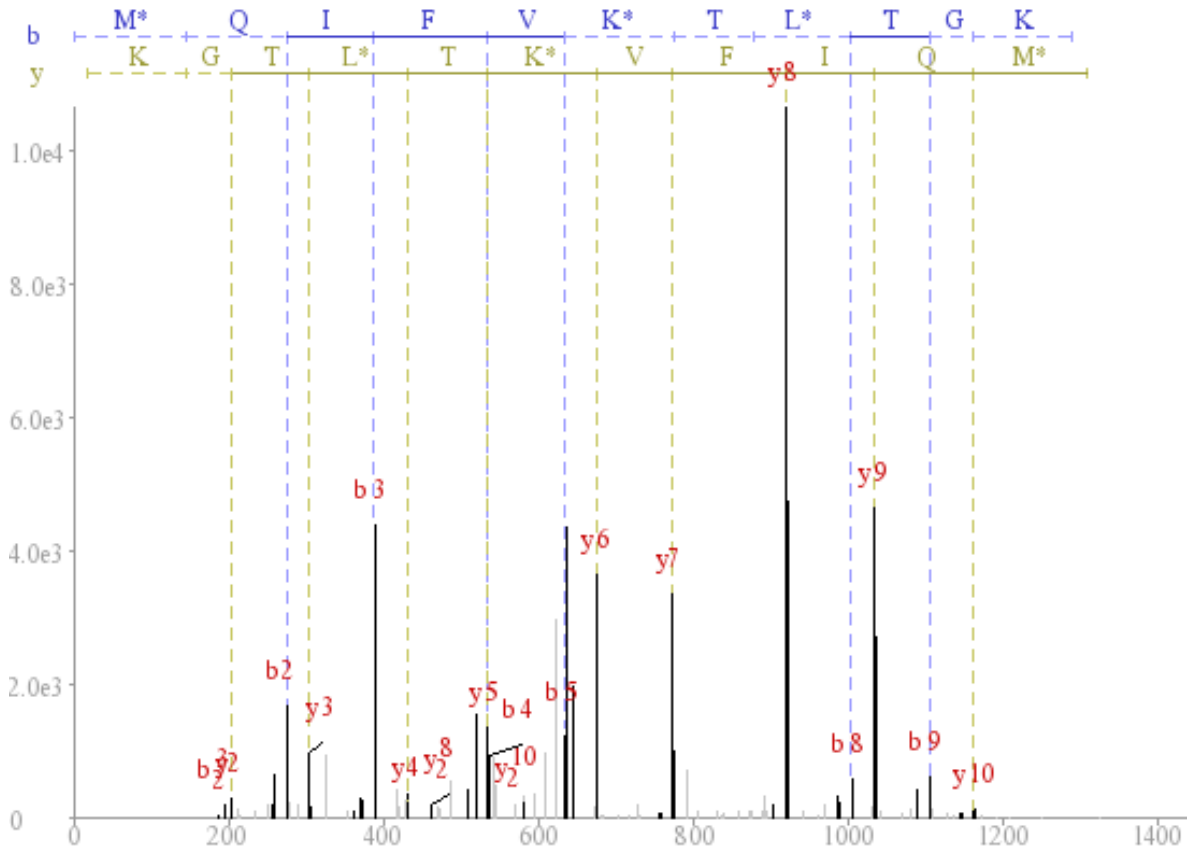


Figure 5-1: Triply oxidized ubiquitin peptide 1-M(+16)QIFVK(+14)TL(+14)TGK-11 identified by InsPecT. This tandem mass spectrum was correctly matched by InsPecT software with MQScore of 3.46 and p-value < 0.01, highlighting the need for relatively liberal search constraints that allow more than one modification per peptide. *Experimental conditions*: 150 µA current at solution flow rate of 50 µL/min.

and false negatives that would skew information about reactivity and thus solvent accessibility. The flip side of this issue is that imposition of more modifications than the software can effectively handle may prove detrimental to reporting of true matches, in addition to requiring extensive processing time to run searches. With some oxidative studies it may be advantageous to conduct unrestrictive searches, as in the case of glycosylated proteins where polysaccharide mass shifts are often unpredictable.[142, 143] So called "blind" searches of eye lens proteomic data have yielded interesting discoveries about PTM relevance in crystallin proteins and founded speculation about the role of oxidation in cataract formation.[144, 145]

5.3.3 Coeluting Oxidized Peptide Isoforms

One potential complication inherent in oxidative experiments is the coexistence of the same peptide sequence with differentially oxidized sites. These oxidized isomers or "oxidisomers" have the same parent ion mass (isobaric) and similar affinities for reversed-phase resin. In this case, it may not be possible to chromatographically resolve oxidisomers such that the coeluting parent ions are fragmented simultaneously, producing a mixture spectrum like the one shown in Fig. 5-2. It is not known how often these isomers occur, as software search results report only one match per spectrum and second-best matches must be manually investigated on an individual basis. However, analysis of one electrochemically oxidized ubiquitin data set (5 μM ubiquitin exposed to 300 μA at 25 μL/min) by the ByOnic curator revealed that over 30% of matched peptide spectra may be some mixture of cofragmented species (unpublished work by Bern *et al*), while nearly 50% of identifications may register some residual level of uncertainty. Resolving oxidisomers chromatographically was one goal of the 2D approach, but due to the

difficulty in reliably detecting their frequency of occurrence, it remains unclear whether a significant impact was made relative to 1D measurements.

5.4 Peptides Identified from Oxidized Proteins by Two Software Packages

With growing awareness of the complexity in mining MS data for oxidative modifications, an evolution occurred in the choice of software used to undertake this challenge. Early studies utilized DBDigger to search for +16 Da mass shifts on a modest set including the most reactive amino acids. Later the InsPecT software package was utilized for its ability to dynamically accommodate an unlimited set of user-defined modifications. The publication of the results reported in Chapter 3 highlighted the need for more broad-based search capabilities to adequately match the number of possible oxidative mass shifts. Another oxidative study followed which showcased ByOnic software for this purpose and reported a wide range of modifications in addition to the +16 Da shift most commonly observed.

Peptide tandem mass spectra were searched for the liberal list of oxidative modifications designated by the ByOnic oxidation module and listed in Table 2-1. For InsPecT searches, peptides were allowed up to three oxidation events and a conservative threshold (p-value ≤ 0.01) was used to filter the resulting matches. These results are contrasted with those from ByOnic, which does not strictly demarcate the number of modifications that can occur on the same peptide, and has a p-value mechanism built into the quality calculation (Score ≥ 380) upon which matches were filtered. The following sections include a discussion of amino acid level oxidation events identified on peptides from control unoxidized and electrochemically oxidized model

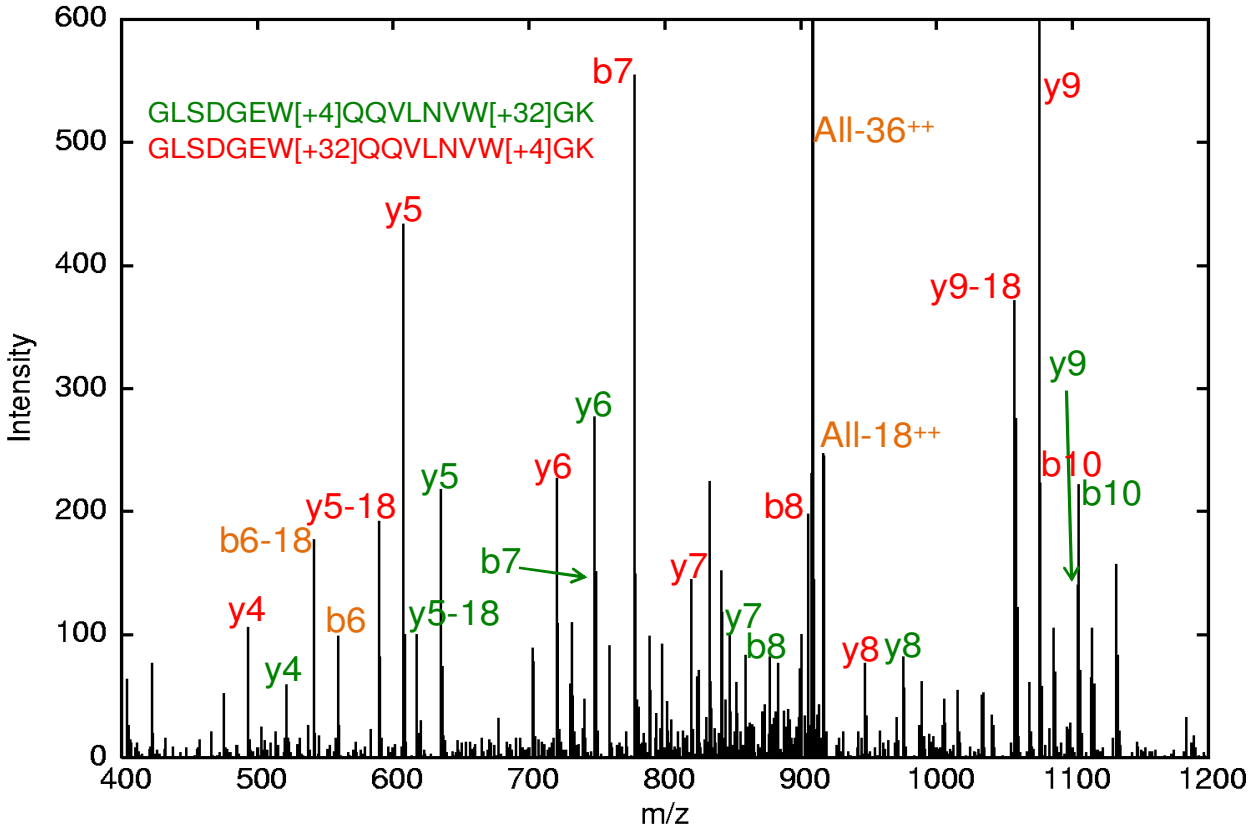


Figure 5-2: Mixture "oxidisomer" tandem mass spectrum from electrochemically oxidized apomyoglobin. Color-coded fragment ions belong to the matching oxidized parent ions shown at top left. This tandem mass spectrum was identified as containing two differentially oxidized peptides during investigation performed by curator of ByOnic software as discussed in Section 5.3. Search results reveal a list of interpreted spectra with each matching a single peptide identity, so this example illustrates the existence of a richer set of information beneath the best scoring oxidized peptide matches typically reported. *Experimental conditions*: 300 µA current at solution flow rate of 25 µL/min.

proteins ubiquitin, lysozyme, and apomyoglobin. Results from both search software packages are discussed in light of known secondary structural features, as well as SASA factors, whose calculation is detailed in Section 2.6.2.

Overall, ByOnic tended to register higher peptide spectral counts for both total (oxidized + unoxidized) and oxidized peptide variants, which could be attributed to filtering at a 5% false discovery rate (FDR) relative to the 1% used by InsPecT. However, confidence about such FDR distinctions should be guarded for the previously discussed (Section 2.4) differences in algorithmic strategy and input parameters, along with the inherent challenge in defining false positives within oxidative search results. Results from both software searches turned up a surprising number of background oxidation events in control experiments, with a troublesome level detected in lysozyme and especially apomyoglobin. As mentioned earlier, both InsPecT and ByOnic identified unexpected oxidation events on tyrosine and proline within the ISB reference protein digest data set, while oxidized methionine was not unforeseen. Oxidation was often found on these three residues in 1D controls, but a significantly wider breadth of residues appeared to be affected with the increased sensitivity in 2D sampling.

One solution to coping with background oxidation events is to compare the ratio of oxidized to total peptide matches for both the control and experimental samples to obtain a relative change in oxidation percentage. However, excessive background oxidation may suggest that some structural perturbation was present in a fraction of the intact protein ensemble prior to electrochemical oxidation. For this reason, protein FTICR-MS scans should be carefully conducted (low ion trap accumulation times and multiple scans for adequate signal averaging) to maximize accurate detection of satellite peaks. While ubiquitin appeared to exhibit less

contamination by satellite peaks, the intact protein mass spectra for both lysozyme and apomyoglobin suggest the presence of a modified form that may be more vulnerable to oxidation. This speculation arises from the observation that the oxidative +16 Da signature appears on the satellite peak for each protein, with the satellite distribution of products leaning toward more the heavily "decorated" variety. Therefore, a detailed interpretation of ubiquitin oxidation events is highlighted with the expectation that native structure should have been labeled, with less emphasis on the other two proteins. Based on satellite peaks observed from intact protein mass spectra, it seems likely that some uncharacterized species within lysozyme and apomyoglobin solutions were electrochemically oxidized and thus contribute oxidized peptides to the final tally of site-specific labeling.

5.4.1 Ubiquitin

Ubiquitin is a small (8.5 kDa) protein conserved throughout evolution, as its covalent linkage at Lys48 to other proteins targets them for destruction by the proteasome.[146] Its secondary structure consists primarily of β-strands, along with one long α-helix and two smaller helices. While informally known as a "rock" due to its proteolytic resistance in aqueous conditions, peptides were readily generated in the organic-aqueous digestion solution employed by these studies. Ubiquitin was chosen for these studies for a variety of reasons, including its commercial availability at high purity (>98%) and the limitation of sulfur-containing residues to its single methionine. Because of its small size and rigid structure that has been shown to resist oxidation-induced unfolding, it was thought that this protein should show good correlation with SASA values under optimal oxidative labeling conditions.

Proteolytic sequence coverage was typically 72-100%, and interestingly the lower coverages were affiliated with 2D experiments. Missing tryptic peptides typically included a rather polar peptide from the middle (28-AKIQDK-33), and four residues at the C-terminal end of the protein (72-LRGG-76). Semi-tryptic peptides measured by 2D and identified by InsPecT revealed backbone cleavage which can be induced by oxidative damage.[72] This empirical finding suggests that MS search parameters in oxidative studies should routinely include provision for this non-specific cleavage event, albeit at a cost in terms of processing time due to a spike in the number of different peptides. The N-terminal peptide with a missed cleavage site (1-MQIFVKTLTGK-11) is lacking primarily from the 2D results for InsPecT, but ByOnic tended to find spectra matching one or both of these two short tryptic peptides. Interestingly, neither of the two short peptides comprising this longer fragment is sampled in isolation, but the latter tryptic sequence (TLTGK) is occasionally found attached to the longest peptide (12-TITLEVEPSDTIENVK-27) which is also the most commonly identified. This "TITLE" peptide was sampled most heavily in 2D experiments, with a large discrepancy of usually more than two hundred peptide spectral counts relative to less than forty for most other tryptic peptides. The second ranked peptide in terms of 2D sampling is 55-TLSDYNIQK-63, with nearly a hundred spectral counts in some measurements.

While the reason for the large discrepancy in peptide sampling frequencies remains unclear, the advantage of increased overall spectral counts in 2D experiments allowed closer examination of oxidative labeling patterns. Oxidation sites found by both softwares in two different control 1D samples include Met1, Phe4, and Tyr59; while additional oxidation events appeared to be centered around Pro37/38 and scattered throughout the TITLE peptide most

heavily on the N-terminal end from Ile13 to Pro19. The 2D results also showed a background oxidation "hotspot" around Pro19 that was shifted more toward the middle and C-terminal end of the 16-residue peptide. This hotspot includes an indication of a likely modification at Glu18, with at least one spectral count matched by both softwares in both 2D controls. The 2D control results corroborated the 1D finding of significant oxidation at Tyr59 and likely oxidation of Met1, which along with multiple Pro sites comprise those residue types identified as oxidized in the ISB data set. Finally, ByOnic results suggested that both Phe45 and His68 suffered some level of background oxidation.

In the previous publication, it was noted that the solvent-excluded or "buried" Leu15 residue was found oxidized by 1D analysis in both non-native and native solution conditions, as well as the control experiment. Here we found that Leu15 was indeed oxidized in a 2D control sample (2.2%, 2 of 92 peptide spectral counts) and oxidized to a slightly larger extent (+0.8%) in the experimental 2D counterpart (3.0%, 7 of 235), according to InsPecT (results refer to Table A-6 columns "2D Control B" and "2D Expt B" with EC conditions of 300 μA at 25 μL/min). ByOnic showed lower level of control oxidation (0.7%, 1 of 135) for a greater change (+1.6%) relative to the experimental (2.3%, 12 of 422). This subtle change reported by both softwares suggested that Leu15 was sufficiently exposed during the experimental timeframe for a small fraction of ubiquitin proteins to become oxidized at this site. This residue is one of those considered buried in the crystal and NMR structures, which also has a low average SASA value of almost 4 $Å^2$ in MD simulations (SASA factor = 0.3). However, this residue became transiently exposed during MD simulations (Fig. 5-3) to nearly 30 $Å^2$, which corresponds to roughly 17% of the area maximally exposed in the free amino acid (max SASA factor = 2.4).

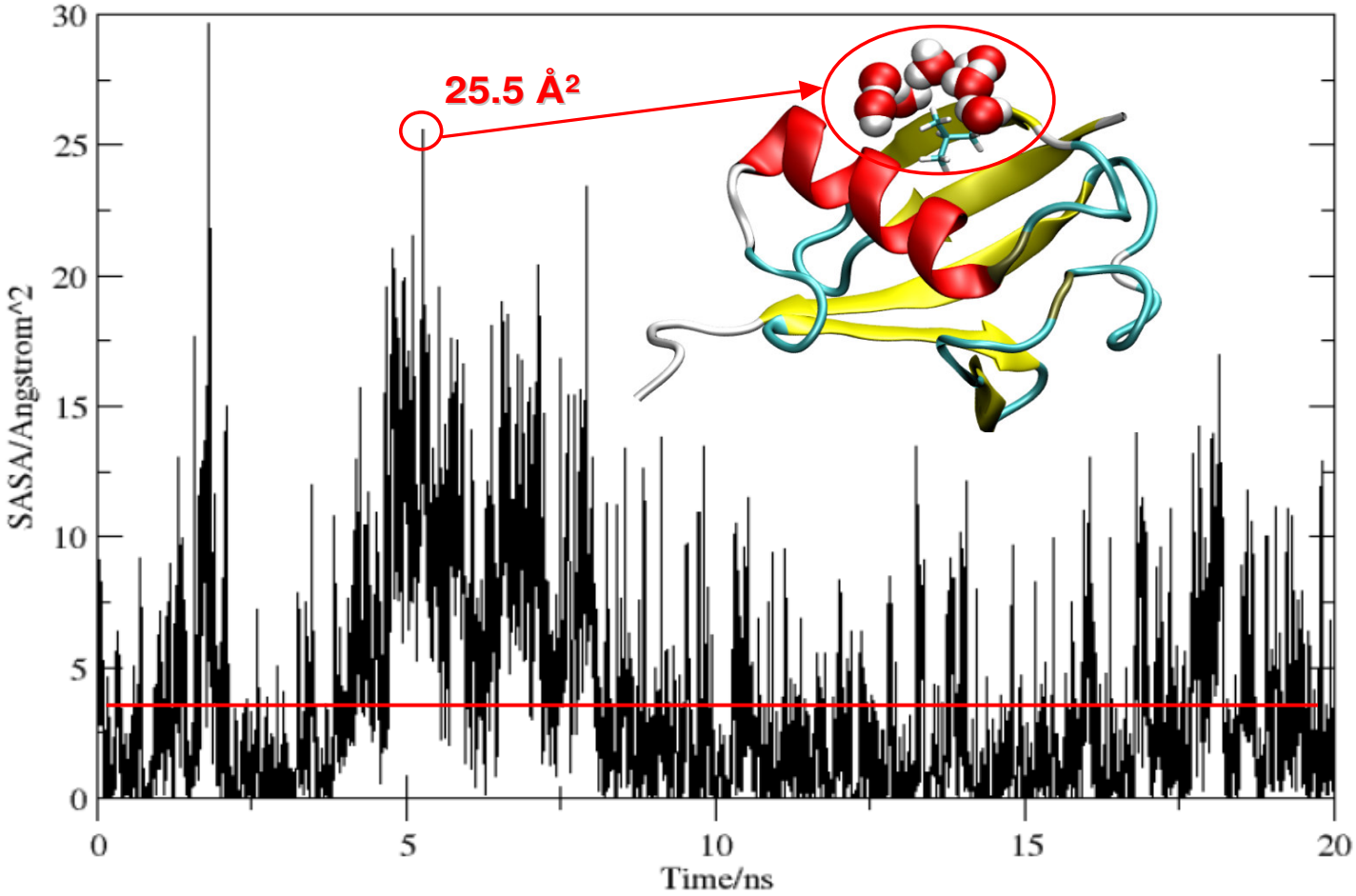


Figure 5-3: Plot of SASA values for ubiquitin Leu15 over MD simulation timeframe. Inset ubiquitin structure shows the Leu15 sidechain in the presence of water molecules during the second-highest exposure at 25.5 Å^2. Line roughly shows average SASA value of 2.9 Å^2. SASA factors: Average = 0.3, Max = 2.4.

Another example of this phenomenon is found in the same control and experimental samples with nearby Ile13, which appears buried with only 1.5 $Å^2$ exposed in the crystal structure and 13 $Å^2$ dynamically exposed in one NMR model. The fact that this residue shows a distinct increase in labeling from the control (ByOnic – 0%; InsPecT – 1.1%) to the oxidized form (ByOnic – 5.2%; InsPecT – 2.1%) is supported by the insight provided by the SASA factor of 0.9, three times the factor of Leu15. This aliphatic sidechain frequently samples more than 20 $Å^2$ with spikes up to 52.5 $Å^2$, which is more than 30% of the SASA maximally exposed in the free amino acid (max SASA factor = 4.3). For perspective, the highest SASA factor of 21.2 (max SASA factor = 36.4) is assigned to the most reactive ubiquitin residue Tyr59.

ByOnic reported a significantly higher level of induced oxidation in the TITLE peptide center loop that connects the N-terminal β-strand to the C-terminal α-helix, which is not unexpected considering the greater flexibility of loop regions. Lower levels in the InsPecT results are primarily clustered on the N-terminal end of this peptide, which is calculated to be relatively more exposed and reactive. Both software results showed that ubiquitin generally undergoes a distinct increase in oxidation events upon exposure to 300 μA current at a flow rate of 25 μL/min. Intact scans showed primarily mono-oxidized (+16 Da) protein signal that appeared to severely overwhelm less abundant peaks around +32 and +48 Da. These experimental conditions may have over-oxidized the protein as evidenced by complete disappearance of the native unoxidized peak, leaving the number of modified forms a mystery shrouded by the dynamic range limitation. In this case, it is not surprising that the oxidized spectral counts do not seem to correlate strongly with solvent accessibility, but more so with reactivity alone as experimental counts increased at or near control hotspots (Pro19, Tyr59,

His68). The SASA factors for highly reactive residues overshadow the reactivity of neighboring residues regardless of their solvent accessibility. Somewhat better agreement is realized when search results are compared with SASA factors in peptides that lack these highly reactive residues (Fig. 5-4). The TITLE peptide illustrates the more sensitive evaluation that can be accomplished without the interference by reactive hotspots. Additionally, the likelihood of a lower reactivity residue becoming oxidized while a nearby hotspot remains unoxidized is doubtful, so accurate detection of oxidation at these sites will depend on the ability of the search software to interpret spectra for which multiple modifications occur on a single peptide.

Moderate oxidation induced by 150 μA current at a flow rate of 50 μL/min (labeled "2D Expt C" in Table A-6) registered a similarly modest level of site-specific oxidation events. Comparison of the ByOnic 2D ratios for sites corroborated by InsPecT showed that the moderate (150-50) and more heavily (300-25) oxidized samples revealed less modification in the lighter exposure for Ile13 (0.6% vs 5.2%), Glu18 (3.5% vs. 6.9%), Ser20 (10.4% vs. 12.8%), and Tyr59 (10.3% vs. 19.0%). In contrast, both software results reported increased labeling of His68 (92.9% vs. 28.7%) in the less oxidized sample, which might be an artifact of the challenges posed by multiply modified peptides present in the more oxidized sample. This experimental 2D sample was also run twice by 1D to gauge reproducibility in peptide spectral counts and detection of oxidized sites. Some peptides showed significant variability in total spectral counts, and there appeared to be more overlap in ByOnic oxidation sites relative to the spottier and less consistent InsPecT identifications. Overall, the results for this relatively simple protein are not consistent enough with solvent accessibility to make any solid interpretations without prior knowledge of the protein structure.

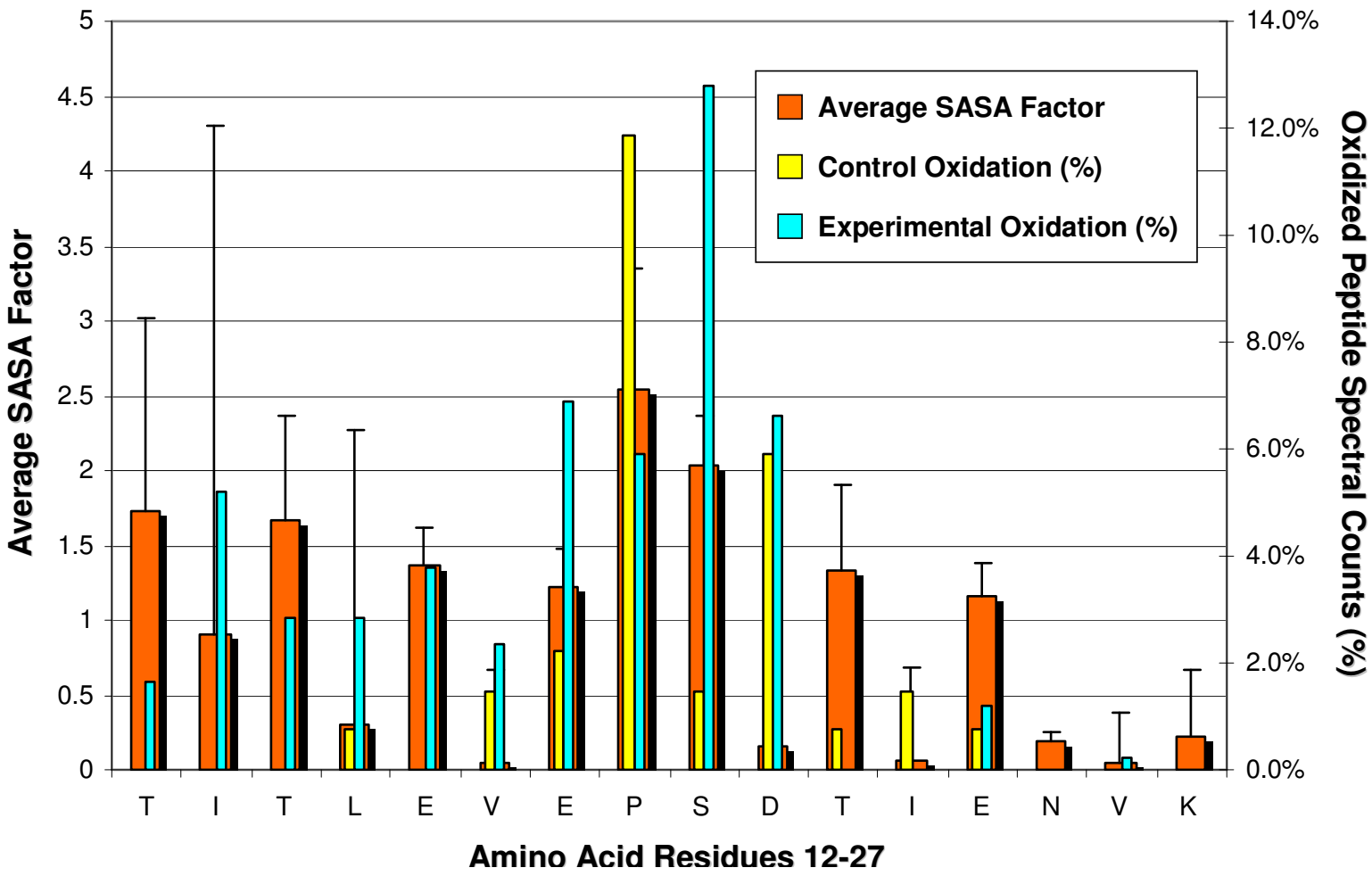


Figure 5-4: SASA factors for the ubiquitin TITLE peptide, which illustrates a relatively level playing field due to its absence of high-reactivity residues. Orange bars show average SASA factors, with error bars indicating the maximum SASA factors. Yellow bars represent the background oxidation in a control sample and blue bars reflect the oxidation in an electrochemically oxidized sample; percentages are the ratio of oxidized to total (unoxidized + oxidized) peptide spectral counts identified by ByOnic software.

5.4.2 Lysozyme

Lysozyme is a 14.3 kDa enzymatic protein that catalyzes the degradation of bacterial cell wall peptidoglycans, making it an effective antimicrobial agent. This important protein is abundant in egg white and can be found in many immunological defense fluids, including tears, mucus, saliva, and milk for the developing infant.[147] For this reason, it has been the focus of much research, and a multitude of its structures have been deposited in the Protein Data Bank. Its secondary structure is mostly α-helical with requisite loops, turns, and intermittent β-strands.

Proteolytic sequence coverages for lysozyme ranged from 67-100%, with the lower coverages due to the almost complete disappearance of the longest tryptic peptide (74-97) from an entire batch of samples. Interestingly, a few spectral counts measuring this peptide were obtained in 1D analysis. Regardless, most other tryptic peptides are measured in significant quantities, and often more frequently in oxidized samples. One hypothesis for this phenomenon is that oxidation typically diminishes peptide affinity for reversed-phase resin, thus reducing the incidence of irreversible adsorption. Background oxidation was rampant in control 2D samples, most commonly affecting the usual suspects: Cys, Met, Trp, Tyr, and Pro residues. However, background events were also identified by both softwares for a wide breadth of less reactive residues albeit with lower spectral counts. Interestingly, the active site residues Glu35 and Asp52 were modified in at least one of two 2D controls according to InsPecT, though not to the extent of the Asp101 substrate binding site (4/210 ~1.9% and 10/93 ~10.8%) as ByOnic reported for both controls. Such modifications would likely impact the results of functional enzymatic assays that were considered as a test of structural integrity. A reduction in the percent oxidation level was achieved at most sites with the early addition of antioxidant DTT along with bubbling

nitrogen gas in the digestion solution prior to capping for the incubation period. However, background oxidation levels were still clearly non-zero and these steps were not performed for most samples that were generated prior to these exploratory tests.

Lysozyme (5 μM) was subjected to 200 μA current at three different flow rates (25, 50, and 100 μL/min) with each doubling correlated to a halving of the residence time (exposure duration of 4, 2, and 1 seconds) inside the electrochemical cell. In general, oxidized peptide spectral counts tended to dwindle as residence time increased, which is counterintuitive to logic as well as the intact mass spectra showing increased oxidized variants. Therefore, two theories are presented to cope with this observation. The first is that longer residence times lead to oxidative unfolding, allowing oxidation events to pool in the most reactive residues regardless of solvent accessibility. This theory is supported by a large increase in Met oxidation according to ByOnic (from 176/453 ~ 38% at 100 μL/min to 426/528 ~81% at 25 μL/min). The second theory involves the likelihood that multiply oxidized peptides were causing difficulties in achieving proper oxidative annotations and sufficiently high scores, which resulted in diminished oxidation reporting. The truth probably employs both factors to some degree, along with the reality that a number of buried sites with less than compelling maximum SASA values were significantly labeled according to these search results.

5.4.3 Apomyoglobin

Much like hemoglobin, myoglobin is a critical carrier for molecular oxygen in muscle tissues. Its structure consists of eight helices connected by loops to form a hydrophobic pocket surrounding the doubly His-linked heme prosthetic group. This physiologically important

protein is widely studied as a model protein in its "apo" form, wherein the heme group has been removed. Some deviation from the holo-myoglobin structure is expected with heme removal, but currently no high resolution structures exist for this isoform. In addition to insights gleaned from its folding kinetics, the study of this protein structure may provide clues about other proteins that bind prosthetic groups or more transitory ligands.

Apomyoglobin peptide sequence coverage ranged from 64-100%, with most oxidized samples lacking peptides sporadically from two residue spans: 43-87 and 97-118. A relatively high frequency of basic sites, especially in the first region, can result in shorter tryptic peptides that were apparently lost at some point for most of the oxidized samples and two of the four 2D controls tested. Interestingly, these were all analyzed by the same instrument (LTQ2), but the two control 2D samples that were analyzed by the high-performance Orbitrap-LTQ, which has a different type of LC system, only suffered loss of the 103-118 peptide. A 1D control run on the Orbitrap-LTQ from the same digest revealed this missing peptide present in the mixture with almost 30 peptide spectral counts according to ByOnic and around 6 spectral counts identified by InsPecT. This peptide appears in only one other sample from apomyoglobin oxidized by 250 μA at 50 μL/min. While this region (residues 97-118) only registered two single-count oxidation events in the InsPecT search, it was interesting to note that this peptide suddenly appeared with over 30 counts by both software accounts in the *second* run from the same sample. In other words, two consecutive 2D analyses revealed the appearance of a missing peptide after the sample had already been exposed to the analytical reversed-phase front column, which implies this peptide was irreversibly adsorbed on the first pass. This also happened with residues 63-87, which were altogether absent from the first 2D run and then present with over 200 spectral

counts (tryptic 64-77 peptide) in the second run. These observations spark concerns about reproducibility among replicate measurements, and whether steps like column pre-conditioning can be systematically taken to minimize adsorption.

Replicate 2D controls that varied only in loading volumes allowed examination of the Orbitrap-LTQ measurement reproducibility with only a change in peptide concentration. Parent ions are measured in the Orbitrap with high resolution and mass accuracy; parent ion mass error based on the latter metric is utilized along with spectral features to evaluate the correctness of peptide matches. These controls provide a compelling account of the reason why many oxidative mapping studies, which typically employ only 1D peptide separation on lower performance instruments, do not report significant background oxidation events. The consistently more conservative InsPecT results detected oxidation at 15 sites in the control protein by 1D analysis, but only 3 of these (Trp7, Trp14, and Met131) received more than a single InsPecT spectral count along with ByOnic confirmation. The increased sampling provided by 2D analysis revealed a plethora of background oxidation, with validation of the 1D sites in addition to more than 50 new ones. Of these, 40 were corroborated by InsPecT results from both controls, with 34 of the doubly reported sites registering more than one spectral count in both reports, and 26 of those sites confirmed with ByOnic results. Both Trp7 and Trp14 registered at least 50 spectral counts in both controls, with ByOnic counts topping out at 118 (13%) for Trp7 and 245 (39%) for Trp 14. While these two amino acids display an extreme level of background oxidation, similarly problematic scenarios appear to be scattered throughout the stock protein on Tyr, Phe, His, Leu, Ile, Val, and Gln residues.

While the change in oxidation ratios can be compared among differentially oxidized samples, it is unclear what advantage that offers when the background levels are so high. This phenomenon could be due to sample processing artifacts, but intact scans show satellite peaks that may be responsible for some fraction of the modifications reported. In this case, native structures might have been labeled for say four-fifths of the solution ensemble, but it is not distinguishable at the peptide-level which identifications can be attributed to either native or non-native structure. Such a distinction could theoretically be made with expensive isotopic labeling experiments as discussed in Chapter 6, but the combinatorial challenge of considering two oxygen isotopes in searches of tandem mass spectra may prove to be computationally intractable.

5.5 Conclusions

The challenge in properly acquiring and interpreting oxidative mapping data remains daunting. It is not surprising that 1D-LC does not provide the sensitivity that 2D-LC offers with its increased dynamic range and longer analytical timeframe. However, 1D analysis appears to suffer less from missing sequence coverage while detecting many of the most commonly modified residues at comparable ratios. Although background protein oxidation is not desirable, it cannot be ignored since this would compromise structural integrity for a fraction of the solution ensemble undergoing experimentation. Differences in software algorithms and input parameters can influence reporting of oxidized sites, and it is likely that the overall strategy employed with InsPecT searches may be too conservative while ByOnic searches could be overly liberal. Their different approaches to searching mass spectra lend increased confidence in peptide matches which are identified as oxidized by both algorithms, while those sites uniquely

reported by each were often fruitful upon spectral evaluation. This is especially important given all the factors that may convolute spectral interpretation, since a reliable calculation of false positive rates does not appear to be well-established for oxidative mapping studies. While it remains unclear whether electrochemical oxidation was performed under optimal conditions for surface mapping applications, it should be recognized that the accurate detection of oxidized sites is essentially a discovery process without a pre-defined set of modifications. Even model proteins contain a variety of factors that can impact amino acid reactivity and thus may preclude a strong correlation between mass spectral search results and solvent accessibility.

6. Conclusions

This dissertation research has demonstrated that proteins can be controllably oxidized with electrochemistry by exploiting the differential reactivity of amino acids to probe their solvent accessibilities within folded proteins. The combination of this covalent labeling approach with high performance mass spectrometry has revealed a new experimental approach for oxidative characterization of protein structures. While the electrochemical technique will require further development and optimization, the work to date has established a solid platform for oxidative labeling, deep MS measurement, and computational extraction of surface mapping information.

Several attributes of the boron-doped diamond (BDD) are particularly attractive for this work, including its smooth surface and low cell volume for a relatively low adsorption profile which might be further reduced with higher flow rates. Its flow-by design is highly compatible with increasingly rapid exposures of analyte protein to the oxidizing environment inside the flow chamber. Importantly, the BDD showcases a documented ability to produce •OH directly from water within its wide working potential window. The additional contribution to the oxidant flux from sulfate radical may offer an interesting advantage for surface mapping with BDD electrochemistry when properly harnessed.

The computational search results of BDD electrochemical oxidation data provide evidence that a wide breadth of amino acid residue types can be labeled. However, more effort will be required to counter the challenges that remain in successfully mining MS data for myriad oxidized sites with reliable filter levels to minimize false positives. Reliance on distinct and

complementary software approaches appears to be a viable strategy for identifying highly probed sites. For example, the two algorithms utilized here for searching MS data each presented a hybrid approach, with *de novo* spectral interpretation factors designed to cope with the possibilities posed by a multitude of oxidative mass shifts that can occur in any combination. The identification of reactive hotspots within the interrogated protein structures by oxidized sites reported by both algorithms shows reasonable agreement with the calculated SASA factor, which encompasses both the intrinsic reactivity of free amino acids weighted by the relative solvent accessibility of their dynamically exposed side chains. It is not surprising that the most reactive residues form the majority of these hotspots that are most often identified with typical reversed-phase LC-MS/MS analysis. However, the orders of magnitude spanned by differential rates of amino acid reactivity seem to be more effectively accommodated with increasingly sensitive measurements that allowed detection of less reactive sites. When considering whether analysis of lower reactivity sites is worth pursuing, it is important to remember that fewer spectral counts is a reflection of the lower reactivity, not the lesser importance of the information revealed.

Detection of the less reactive subset of amino acids, relative to those easily targeted by either direct or indirect oxidation mechanisms, is greatly enhanced by the increased dynamic range of 2D chromatography paired with high performance mass spectrometry. The multidimensional separation of peptides offers the critical advantage of decreasing the peptide complexity over the analytical timeframe to allow sampling of oxidized peptide variants present in low abundance. When this advanced technique is paired with the technological benefit of highly accurate parent ion masses and charge states scanned in the Orbitrap, the overall effect is

a vital enhancement of measurement sensitivity with availability of key metrics to improve the compilation of candidate peptide matches.

Another challenge lies in conquering background oxidation that interferes with data analysis and may jeopardize the structural integrity of the solution ensemble experimented upon. However, these issues do not overshadow the fact that proteins have been successfully oxidized by the BDD electrode, whose utility in protein structure research remains to fully realized. The demonstrated refinement of experimental parameters for more rapid exposure has shown that further improvements could make this methodology compatible with a more ideal reaction timescale. Finally, the proposed pipeline provides foresight in moving toward a high-throughput method for gleaning structural information from protein mixtures.

6.1 Next-Generation Optimization of Electrochemical Surface Mapping

The experimental parameters optimized as part of this dissertation research allowed achievement of predictably controllable oxidation signatures that were unique to the protein of study, which was likely dependent on the number of highly reactive residues. Therefore, knowledge of the number of aromatic and sulfur-containing residues in the protein structure can provide guidelines for gauging moderate versus heavy oxidation conditions. However, watching for a unimodal distribution on this experimental timeframe may not be sufficient to successfully label only native structures, due to the threat posed by oxidation-induced unfolding. While two of these model proteins have been shown to be structurally resilient under a moderate oxidation load,[30] other proteins of interest might not benefit from this comfort. Therefore, a necessary improvement would be an increase in the flow rate to further reduce the protein residence time

inside the electrochemical flow chamber. Ideally, proteins would be oxidized on the microsecond timescale or less, so efforts should be made to determine whether this feat is achievable. Due to the current design of this chamber wherein protein solution is forced to undergo a 90° change in direction twice before exiting the cell, fluidic turbulence may produce backpressure at higher flow rates and cause leakage within the cell assembly. Evidence of this possibility was observed with minor leakage beyond the gasket window at 100 μL/min. In order to accommodate higher flow rates, a linear flow cell design that facilitates laminar flow would be desirable for achieving a significantly more rapid exposure.

Reactive hotspots tended to reside in the center of tryptic peptides, while similarly reactive residues were also present near peptide ends delimited by tryptic cleavage sites. However, the analytical difficulty with pinpointing these sites was described in the previous chapter, wherein the tendency for peak intensities to trickle off at the low and high ends of the mass-to-charge range reduced the presence of fragment ions to specifically diagnose oxidation occurring near peptide ends. Therefore, in order to expand the range of easily detectable sites, it would be advantageous to explore proteases with variable cleavage specificity to overlap regions that occur near tryptic endpoints and thus provide fairer sampling across the primary sequence. This approach may also provide insight into peptide bias, which may result from a variety of factors, including length, hydrophobicity, and charge.

Upon examination of the ion chromatograms resulting from the higher concentration salt pulses in 2D LC-MS/MS experiments, it appeared that fewer peaks eluted relative to the first pulse. This implies that there may be room for improvement on the 2D methodology to maximize separations, perhaps with the use of more salt pulses and/or varied increments like

10%, 40%, 75%, and 100% of 500 mM ammonium acetate instead of the 25%, 50%, and 100% scheme used in these studies. Finer tuning of salt pulse strengths utilized in multidimensional chromatography may be further benefited by reversed-phase gradient optimization. The observation that most ubiquitin peptides eluted before the solvent B contribution reached the 50% mark during a 25 to 75% ramp spurred an experiment with a slower 25 to 50% ramp. While the new gradient produced more variation in total spectral counts among tryptic peptides, more oxidized spectral counts were detected at low abundance within the control protein.

6.2 Specific Research Targets for Optimization

Based on findings of extensive background oxidation in commercially available proteins, it is advised that future work to demonstrate the application of electrochemical protein oxidation focus on solving (or minimizing) this issue. This problem may be especially pronounced in proteins exposed to oxidizing environments, such as blood proteins that are chronically exposed to oxygen. Along these lines, some proteins may be structurally sensitive to oxidative damage, such as β-lactoglobulin A which reportedly unfolds with even a single oxidation event. (Sharp, J.S., presented at ASMS Covalent Labeling Workshop in 2008) Therefore, care should be taken to sensitively measure the intact protein ensemble for structural integrity prior to oxidative experimentation. With respect to MS measurements, the intact protein solution should be screened at relatively low concentrations to avoid overwhelming the dynamic range, in order to detect lower abundance satellite peaks. This also means that attention should be given to minimal loading of the ion trap during MS analysis. However, based on the analysis of ubiquitin, it appears that a significant portion of the background oxidation may be accumulated

during sample handling, so efforts should be made to minimize the chance for oxidation to occur. These endeavors include using antioxidants in the reaction solution, degassing the solution with an inert gas prior to proteolytic digestion, and minimizing exposure to fluorescent lighting. One last concern with commercially available stock proteins is that some are produced by pooling samples from various donor sources, such as the serum albumins whose variants are evident with intact MS screening. Experiments with similar protein stocks would likely suffer informatically from sequence variants that are not contained in the reference database, as well as PTMs such as glycosylation that may occur *in vivo*.

Other concerns involve the experimental design for further investigation about the utility of the BDD electrode for surface mapping applications. Ideally, instead of a single structure, an oxidative mapping experiment would involve a relative measurement for two structures for the same protein. For example, oxidative labeling could be performed on protein in the presence of its ligand, with comparison of its site-specific labels to the "apo" form. This experimental scenario may suffer from contamination by unbound forms due to dissociation kinetics unless the ligand has an extremely high affinity for its binding site. Another possibility is labeling structures before and after a pH-induced conformational change like that of β-lactoglobulin A, which does not require a binding event to occur. Apart from of these possibilities for expanding the application, the next efforts to validate this approach should involve reducing the reaction timescale by increasing the flow rate and lowering the protein concentration. Success could be measured with site-specific labeling of a structurally solid protein like ubiquitin in relatively good agreement with SASA values calculated from MD simulations

6.3 Outlook

One method that may be fruitful in uncovering the source of background oxidation is differential isotopic labeling with normal ^{16}O and heavy ^{18}O oxygen isotopes. This approach may be complicated by virtue of the multiple reaction routes by which oxidation can proceed, such that a thorough investigation would require the use of both heavy isotope water and oxygen gas. By isolating use of the heavy isotope to either the electrochemical cell exposure or the proteolytic digestion in different experiments (with controls), it may be possible to delineate the source of induced and background oxidation. If successful, this approach also could be advantageous for development of an experimental strategy that combines over-oxidation of a denatured "control" sample (heavy – ^{18}O water/gas) along with a moderately labeled "experimental" sample (light – ^{16}O water/gas). When paired with informatic support capable of handling the combinatorial load imposed by two isotopes, analysis of these samples combined at 1:1 in solution would provide an intriguing relative comparison. In this situation, natively exposed residues would appear as a doublet while buried sites would appear as a heavy singlet within the MS1 scans. If perfected, this could provide an accounting of all "oxidizable" residues to reduce false negatives that can be combined 1:1 for a relative measurement of experimental versus control.

Further exploration of the BDD [•OH] production capabilities might also be accomplished with oxygen isotopes for defining its utility for oxidative surface mapping. While the flux of •OH production can reportedly be measured by decrease in absorbance of Alexa Fluor 488,[72] it is possible that this molecule could be oxidized directly and may adsorb to the electrode surface. Alternatively, the use of only ^{18}O water labeling could be used in spin-

trapping experiments, wherein the small molecule DMPO should undergo a +18 Da shift if it is modified by •OH attack, or a +16 Da shift if it is directly oxidized electrochemically followed by O_2 resolution.

Other notable areas for experimental investigation include clarification of analytical issues used to measure peptides. The increase in dynamic range afforded by 2D versus 1D could be estimated by spiking known concentrations (high/low) of two peptides in the sample volume to be loaded. Potential problems may arise from peptide sequence variation that imparts differential hydrophobicity, which could impact ionization efficiency and column binding affinity. The reason for the observed bias in peptide sampling that was reported here should be examined by testing digestion procedures different from the organic-aqueous protocol used in these studies and subjecting the resulting peptides to the same LC-MS methods. Attention should be paid to minimizing the surfaces touched by peptides to reduce losses due to adsorption. The use of multiple proteases on the same oxidized sample may provide better coverage of the protein sequence and increased confidence in oxidized sites due to measurement of overlapping peptides. However, the software used to search MS data should be equipped to handle this change in proteolytic cleavage specificity. Closer examination at the intact protein level by separation of oxidation products by top-down mass spectrometry could result in the detection of lower abundance oxidized products to improve integration with bottom-up searches in efforts to counter the vast combinatorial challenge posed by oxidative labeling. Finally, alternate peptide parent ion fragmentation methods should be explored, such as electron-transfer dissociation (ETD), which has been shown to improve detection of oxidized sites. [148-150]

6.4 Perspective

The research presented here is the first to describe use of electrochemistry for the purpose of indirectly oxidizing proteins with the goal of mapping their dynamically exposed surfaces. Among multiple sources of hydroxyl radicals used for this purpose, the BDD electrode arguably offers the most approachable means to achieve controlled protein oxidation. The initial steps were taken here toward optimizing the utility of the BDD electrode for the purpose of rapid oxidative labeling. Elucidation of potential drawbacks to its usage has provided a springboard to promptly steer future experimentation toward success. In addition, the overall challenges in detection and reporting of oxidized sites were evaluated in undertaking paired chromatographic and informatic approaches. Dynamic simulations of solution-phase behavior for resolved protein structures revealed insight about transient solvent exposures, which may further improve the correlation with reactivity of amino acids within freshly purified proteins. In general, every stride toward the complete analysis of oxidative surface mapping results will advance the field, which may soon include in its toolbox the promising BDD electrode.

References

1. **Minor, D.L.,** ***The neurobiologist's guide to structural biology: A primer on why macromolecular structure matters and how to evaluate structural data.*** **Neuron, 2007. 54(4): p. 511-533.**
2. **Myles, D.A.,** ***Neutron protein crystallography: current status and a brighter future.*** **Curr Opin Struct Biol, 2006. 16(5): p. 630-7.**
3. **Bonvin, A.M.J.J., R. Boelens, and R. Kaptein,** ***NMR analysis of protein interactions.*** **Current Opinion in Chemical Biology, 2005. 9(5): p. 501-508.**
4. **Henderson, R.,** ***Realizing the potential of electron cryo-microscopy.*** **Quarterly Reviews of Biophysics, 2004. 37(1): p. 3-13.**
5. **Geerlof, A., et al.,** ***The impact of protein characterization in structural proteomics.*** **Acta Crystallogr D Biol Crystallogr, 2006. 62(Pt 10): p. 1125-36.**
6. **Koch, M.H.J., P. Vachette, and D.I. Svergun,** ***Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution.*** **Quarterly Reviews of Biophysics, 2003. 36(2): p. 147-227.**
7. **Fontana, A., et al.,** ***Probing the partly folded states of proteins by limited proteolysis.*** **Fold Des, 1997. 2(2): p. R17-26.**
8. **Hubbard, S.J.,** ***The structural aspects of limited proteolysis of native proteins.*** **Biochim Biophys Acta, 1998. 1382(2): p. 191-206.**
9. **Konermann, L., X. Tong, and Y. Pan,** ***Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches.*** **Journal of Mass Spectrometry, 2008. 43(8): p. 1021-1036.**
10. **Clayton, A.H. and W.H. Sawyer,** ***Site-specific tryptophan fluorescence spectroscopy as a probe of membrane peptide structure and dynamics.*** **Eur Biophys J, 2002. 31(1): p. 9-13.**
11. **Engelborghs, Y.,** ***The analysis of time resolved protein fluorescence in multi-tryptophan proteins.*** **Spectrochim Acta A Mol Biomol Spectrosc, 2001. 57(11): p. 2255-70.**
12. **Haas, E.,** ***The study of protein folding and dynamics by determination of intramolecular distance distributions and their fluctuations using ensemble and single-molecule FRET measurements.*** **Chemphyschem, 2005. 6(5): p. 858-70.**
13. **Kelly, S.M., T.J. Jess, and N.C. Price,** ***How to study proteins by circular dichroism.*** **Biochim Biophys Acta, 2005. 1751(2): p. 119-39.**
14. **Miles, A.J., L. Whitmore, and B.A. Wallace,** ***Spectral magnitude effects on the analyses of secondary structure from circular dichroism spectroscopic data.*** **Protein Sci, 2005. 14(2): p. 368-74.**
15. **Liu, H.L. and J.P. Hsu,** ***Recent developments in structural proteomics for protein structure determination.*** **Proteomics, 2005. 5(8): p. 2056-2068.**
16. **Berman, H.M., et al.,** ***The Protein Data Bank.*** **Acta Crystallographica Section D-Biological Crystallography, 2002. 58: p. 899-907.**
17. **Qu, X., et al.,** ***A guide to template based structure prediction.*** **Curr Protein Pept Sci, 2009. 10(3): p. 270-85.**
18. **Floudas, C.A.,** ***Computational methods in protein structure prediction.*** **Biotechnology and Bioengineering, 2007. 97(2): p. 207-213.**
19. **Dill, K.A., et al.,** ***The protein folding problem: when will it be solved?*** **Curr Opin Struct Biol, 2007. 17(3): p. 342-6.**
20. **Chance, M.R., et al.,** ***High-throughput computational and experimental techniques in structural genomics.*** **Genome Res, 2004. 14(10B): p. 2145-54.**

21. **Misura, K.M.S. and D. Baker, *Progress and challenges in high-resolution refinement of protein structure models.* Proteins-Structure Function and Bioinformatics, 2005. 59(1): p. 15-29.**
22. **Fogolari, F., et al., *MM/PBSA analysis of molecular dynamics simulations of bovine beta-lactoglobulin: free energy gradients in conformational transitions?* Proteins, 2005. 59(1): p. 91-103.**
23. **Karplus, M. and J.A. McCammon, *Molecular dynamics simulations of biomolecules.* Nature Structural Biology, 2002. 9(9): p. 646-652.**
24. **Brooks, B.R., et al., *Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations.* Journal of Computational Chemistry, 1983. 4(2): p. 187-217.**
25. **Nelson, M.T., et al., *NAMD: A parallel, object oriented molecular dynamics program.* International Journal of Supercomputer Applications and High Performance Computing, 1996. 10(4): p. 251-268.**
26. **MacKerell, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* Journal of Physical Chemistry B, 1998. 102(18): p. 3586-3616.**
27. **Sanbonmatsu, K.Y. and C.S. Tung, *High performance computing in biology: Multimillion atom simulations of nanoscale systems.* Journal of Structural Biology, 2007. 157(3): p. 470-480.**
28. **Mendoza, V.L. and R.W. Vachet, *Probing protein structure by amino acid-specific covalent labeling and mass spectrometry.* Mass Spectrom Rev, 2009. 28(5): p. 785-815.**
29. **Sharp, J.S., et al., *Measurement of multisite oxidation kinetics reveals an active site conformational change in Spo0F as a result of protein oxidation.* Biochemistry, 2006. 45(20): p. 6260-6.**
30. **Venkatesh, S., K.B. Tomer, and J.S. Sharp, *Rapid identification of oxidation-induced conformational changes by kinetic analysis.* Rapid Communications in Mass Spectrometry, 2007. 21(23): p. 3927-3936.**
31. **Watson, C., et al., *Pulsed electron beam water radiolysis for submicrosecond hydroxyl radical protein footprinting.* Anal Chem, 2009. 81(7): p. 2496-505.**
32. **Guan, J.Q. and M.R. Chance, *Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry.* Trends Biochem Sci, 2005. 30(10): p. 583-92.**
33. **Takamoto, K. and M.R. Chance, *Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes.* Annu Rev Biophys Biomol Struct, 2006. 35: p. 251-76.**
34. **Smedley, J.G., et al., *Probing the pH-dependent prepore to pore transition of Bacillus anthracis protective antigen with differential oxidative protein footprinting.* Biochemistry, 2008. 47(40): p. 10694-10704.**
35. **Sharp, J.S. and K.B. Tomer, *Analysis of the oxidative damage-induced conformational changes of apo- and holocalmodulin by dose-dependent protein oxidative surface mapping.* Biophys J, 2007. 92(5): p. 1682-92.**
36. **Stocks, B.B. and L. Konermann, *Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-induced Oxidative Labeling and Mass Spectrometry.* Analytical Chemistry, 2009. 81(1): p. 20-27.**
37. **Xu, G., et al., *Structural allostery and binding of the transferrin*receptor complex.* Mol Cell Proteomics, 2005. 4(12): p. 1959-67.**

38. **Wong, J.W.H., S.D. Maleknia, and K.M. Downard, *Study of the ribonuclease-S-protein-peptide complex using a radical probe and electrospray ionization mass spectrometry.* Analytical Chemistry, 2003. 75(7): p. 1557-1563.**
39. **Wong, J.W.H., S.D. Maleknia, and K.M. Downard, *Hydroxyl radical probe of the calmodulin-melittin complex interface by electrospray ionization mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2005. 16(2): p. 225-233.**
40. **Buxton, G.V., et al., *Critical-Review of Rate Constants for Reactions of Hydrated Electrons, Hydrogen-Atoms and Hydroxyl Radicals (.Oh/.O-) in Aqueous-Solution.* Journal of Physical and Chemical Reference Data, 1988. 17(2): p. 513-886.**
41. **Stadtman, E.R., *Oxidation of free amino acids and amino acid residues in proteins by radiolysis and by metal-catalyzed reactions.* Annu Rev Biochem, 1993. 62: p. 797-821.**
42. **Dolinova, J., et al., *Oxidation of aromatic and aliphatic hydrocarbons by OH radicals photochemically generated from H2O2 in ice.* Environmental Science & Technology, 2006. 40(24): p. 7668-7674.**
43. **Sharp, J.S., et al., *Photochemical surface mapping of C14S-Sml1p for constrained computational modeling of protein structure.* Anal Biochem, 2005. 340(2): p. 201-12.**
44. **Hambly, D.M. and M.L. Gross, *Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale.* J Am Soc Mass Spectrom, 2005. 16(12): p. 2057-63.**
45. **Aye, T.T., T.Y. Low, and S.K. Sze, *Nanosecond laser-induced photochemical oxidation method for protein surface mapping with mass spectrometry.* Analytical Chemistry, 2005. 77(18): p. 5814-5822.**
46. **Sharp, J.S., J.M. Becker, and R.L. Hettich, *Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry.* Anal Biochem, 2003. 313(2): p. 216-25.**
47. **Johansson, T., L. Weidolf, and U. Jurva, *Mimicry of phase I drug metabolism--novel methods for metabolite characterization and synthesis.* Rapid Commun Mass Spectrom, 2007. 21(14): p. 2323-31.**
48. **Brillas, E., et al., *Degradation of the herbicide 2,4-DP by anodic oxidation, electro-Fenton and photoelectro-Fenton using platinum and boron-doped diamond anodes.* Chemosphere, 2007. 68(2): p. 199-209.**
49. **Maleknia, S.D., et al., *Determination of macromolecular folding and structure by synchrotron x-ray radiolysis techniques.* Anal Biochem, 2001. 289(2): p. 103-15.**
50. **Kiselar, J.G., et al., *Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry.* Int J Radiat Biol, 2002. 78(2): p. 101-14.**
51. **Xu, G. and M.R. Chance, *Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting.* Anal Chem, 2005. 77(14): p. 4549-55.**
52. **Sharp, J.S. and K.B. Tomer, *Effects of anion proximity in peptide primary sequence on the rate and mechanism of leucine oxidation.* Anal Chem, 2006. 78(14): p. 4885-93.**
53. **Maleknia, S.D., M.R. Chance, and K.M. Downard, *Electrospray-assisted modification of proteins: a radical probe of protein structure.* Rapid Commun Mass Spectrom, 1999. 13(23): p. 2352-8.**
54. **Wong, J.W., S.D. Maleknia, and K.M. Downard, *Study of the ribonuclease-S-protein-peptide complex using a radical probe and electrospray ionization mass spectrometry.* Anal Chem, 2003. 75(7): p. 1557-63.**

55. **Maleknia, S.D., J.W. Wong, and K.M. Downard, *Photochemical and electrophysical production of radicals on millisecond timescales to probe the structure, dynamics and interactions of proteins.* Photochem Photobiol Sci, 2004. 3(8): p. 741-8.**
56. **Shum, W.K., S.D. Maleknia, and K.M. Downard, *Onset of oxidative damage in alpha-crystallin by radical probe mass spectrometry.* Anal Biochem, 2005. 344(2): p. 247-56.**
57. **Brabec, V., Mornstein, V., *Electrochemical behaviour of proteins at graphite electrodes. I. Electrooxidation of proteins as a new probe of protein structure and reactions.* Biochem Biophys Acta, 1980. 625(1): p. 43-50.**
58. **Brabec, V., Mornstein, V., *Electrochemical behaviour of proteins at graphite electrodes II. Electrooxidation of amino acids.* Biophys Chem, 1980. 12(2): p. 159-65.**
59. **Fedurco, M., *Redox reactions of heme-containing metalloproteins: dynamic effects of self-assembled monolayers on thermodynamics and kinetics of cytochrome c electron-transfer reactions.* Coordination Chemistry Reviews, 2000. 209: p. 263-331.**
60. **Alessandrini, A., S. Corni, and P. Facci, *Unravelling single metalloprotein electron transfer by scanning probe techniques.* Physical Chemistry Chemical Physics, 2006. 8(38): p. 4383-4397.**
61. **Bridgewater, J.D., J. Lim, and R.W. Vachet, *Using metal-catalyzed oxidation reactions and mass spectrometry to identify amino acid residues within 10 angstrom of the metal in Cu-binding proteins.* Journal of the American Society for Mass Spectrometry, 2006. 17(11): p. 1552-1559.**
62. **Richards, P.G., et al., *Electrochemical modification of lysozyme: anodic reaction of tyrosine residues.* Enzyme Microb Technol, 1994. 16(9): p. 795-801.**
63. **Kendall, G., et al., *Specific electrochemical nitration of horse heart myoglobin.* Arch Biochem Biophys, 2001. 392(2): p. 169-79.**
64. **Matters, D., et al., *Mass spectrometry in demonstrating the site-specific nitration of hen egg white lysozyme by an improved electrochemical method.* Anal Biochem, 2006. 356(2): p. 171-81.**
65. **Davies, M.J., *The oxidative environment and protein damage.* Biochimica Et Biophysica Acta-Proteins and Proteomics, 2005. 1703(2): p. 93-109.**
66. **Halliwell, B., *Oxidative stress and cancer: have we moved forward?* Biochemical Journal, 2007. 401: p. 1-11.**
67. **Kapalka, A., G. Foti, and C. Comninellis, *The importance of electrode material in environmental electrochemistry Formation and reactivity of free hydroxyl radicals on boron-doped diamond electrodes.* Electrochimica Acta, 2009. 54(7): p. 2018-2023.**
68. **Panizza, M. and G. Cerisola, *Application of diamond electrodes to electrochemical processes.* Electrochimica Acta, 2005. 51(2): p. 191-199.**
69. **Zhi, J.F., et al., *Electrochemical incineration of organic pollutants on boron-doped diamond electrode. Evidence for direct electrochemical oxidation pathway.* Journal of Physical Chemistry B, 2003. 107(48): p. 13389-13395.**
70. **Ito, T., et al., *Carbon- and nitrogen-centered radicals produced from L-lysine by radiation-induced oxidation: A pulse radiolysis study.* Chemical Physics Letters, 2008. 462(1-3): p. 116-120.**
71. **Bridgewater, J.D. and R.W. Vachet, *Metal-catalyzed oxidation reactions and mass spectrometry: the roles of ascorbate and different oxidizing agents in determining Cu-protein-binding sites.* Anal Biochem, 2005. 341(1): p. 122-30.**

72. Xu, G. and M.R. Chance, *Hydroxyl radical-mediated modification of proteins as probes for structural proteomics.* Chem Rev, 2007. 107(8): p. 3514-43.
73. Cech, N.B. and C.G. Enke, *Practical implications of some recent studies in electrospray ionization fundamentals.* Mass Spectrom Rev, 2001. 20(6): p. 362-87.
74. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem, 2001. 73(23): p. 5683-90.
75. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nat Biotechnol, 2001. 19(3): p. 242-7.
76. McClintock, C., V. Kertesz, and R.L. Hettich, *Development of an electrochemical oxidation method for probing higher order protein structure with mass spectrometry.* Analytical Chemistry, 2008. 80(9): p. 3304-3317.
77. Tanner, S., et al., *InsPecT: identification of posttranslationally modified peptides from tandem mass spectra.* Anal Chem, 2005. 77(14): p. 4626-39.
78. Bern, M., Y. Cai, and D. Goldberg, *Lookup peaks: a hybrid of de novo sequencing and database search for protein identification by tandem mass spectrometry.* Anal Chem, 2007. 79(4): p. 1393-400.
79. Saladino, J., et al., *Aliphatic peptidyl hydroperoxides as a source of secondary oxidation in hydroxyl radical protein footprinting.* J Am Soc Mass Spectrom, 2009. 20(6): p. 1123-6.
80. Van Berkel, G.J., et al., *Efficient analyte oxidation in an electrospray ion source using a porous flow-through electrode emitter.* J Am Soc Mass Spectrom, 2004. 15(12): p. 1755-66.
81. Sandra, K., et al., *Highly efficient peptide separations in proteomics - Part 1. Unidimensional high performance liquid chromatography.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2008. 866(1-2): p. 48-63.
82. Douglas, D.J., A.J. Frank, and D.M. Mao, *Linear ion traps in mass spectrometry.* Mass Spectrometry Reviews, 2005. 24(1): p. 1-29.
83. Stafford, G., *Ion trap mass spectrometry: A personal perspective.* Journal of the American Society for Mass Spectrometry, 2002. 13(6): p. 589-596.
84. Perry, R.H., R.G. Cooks, and R.J. Noll, *Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications.* Mass Spectrometry Reviews, 2008. 27(6): p. 661-699.
85. Marshall, A.G. and C.L. Hendrickson, *High-Resolution Mass Spectrometers.* Annual Review of Analytical Chemistry, 2008. 1: p. 579-599.
86. Chen, M. and K.D. Cook, *Oxidation artifacts in the electrospray mass spectrometry of Abeta Peptide.* Anal Chem, 2007. 79(5): p. 2031-6.
87. Strader, M.B., et al., *Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic-aqueous solvent systems.* Anal Chem, 2006. 78(1): p. 125-34.
88. Tabb, D.L., et al., *DBDigger: reorganized proteomic database identification that improves flexibility and speed.* Anal Chem, 2005. 77(8): p. 2464-74.
89. Eng, J.K., McCormack, A. L., Yates, J. R. III, *An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database.* J Am Soc Mass Spectrom, 1994. 5: p. 976-989.
90. Narasimhan, C., et al., *MASPIC: intensity-based tandem mass spectrometry scoring scheme that improves peptide identification at high confidence.* Anal Chem, 2005. 77(23): p. 7581-93.
91. Bern, M. and D. Goldberg, *Improved ranking functions for protein and modification-site identifications.* J Comput Biol, 2008. 15(7): p. 705-19.

92. **Charvatova, O., et al.,** ***Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1.*** **Journal of the American Society for Mass Spectrometry, 2008. 19(11): p. 1692-1705.**
93. **Peng, J.M., et al.,** ***Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome.*** **Journal of Proteome Research, 2003. 2(1): p. 43-50.**
94. **Whitmore, L. and B.A. Wallace,** ***DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data.*** **Nucleic Acids Res, 2004. 32(Web Server issue): p. W668-73.**
95. **Fraczkiewicz, R., Braun, W. , J Comp Chem, 1998. 19: p. 319.**
96. **Berman, H.M., et al.,** ***The Protein Data Bank.*** **Nucleic Acids Research, 2000. 28(1): p. 235-242.**
97. **Permentier, H.P., et al.,** ***Electrochemical oxidation and cleavage of peptides analyzed with on-line mass spectrometric detection.*** **Rapid Commun Mass Spectrom, 2003. 17(14): p. 1585-92.**
98. **Permentier, H.P. and A.P. Bruins,** ***Electrochemical oxidation and cleavage of proteins with on-line mass spectrometric detection: development of an instrumental alternative to enzymatic protein digestion.*** **J Am Soc Mass Spectrom, 2004. 15(12): p. 1707-16.**
99. **Brabec, V. and I. Schindlerova,** ***Electrochemical-Behavior of Proteins at Graphite-Electrodes .3. The Effect of Protein Adsorption.*** **Bioelectrochemistry and Bioenergetics, 1981. 8(4): p. 451-458.**
100. **Wang, D., et al.,** ***Effect of chemical structure on the adsorption of amino acids with aliphatic and aromatic substitution groups: In situ STM study.*** **Journal of Physical Chemistry B, 2003. 107(33): p. 8474-8478.**
101. **Hager, G. and A.G. Brolo,** ***Adsorption/desorption behaviour of cysteine and cystine in neutral and basic media: electrochemical evidence for differing thiol and disulfide adsorption to a Au(111) single crystal electrode.*** **Journal of Electroanalytical Chemistry, 2003. 550: p. 291-301.**
102. **Liu, Z. and G. Wu,** ***The electro-oxidative activity of cysteine on the Au electrode as evidenced by surface enhanced Raman scattering.*** **Spectrochim Acta A Mol Biomol Spectrosc, 2006. 64(1): p. 251-4.**
103. **Spataru, N., et al.,** ***Voltammetric determination of L-cysteine at conductive diamond electrodes.*** **Anal Chem, 2001. 73(3): p. 514-9.**
104. **Oliveira, R.T., et al.,** ***Electrochemical oxidation of benzene on boron-doped diamond electrodes.*** **Chemosphere, 2007. 66(11): p. 2152-8.**
105. **Panizza, M., et al.,** ***Electrochemical oxidation of phenol at boron-doped diamond electrode. Application to electro-organic synthesis and wastewater treatment.*** **Ann Chim, 2002. 92(10): p. 995-1006.**
106. **Canizares, P., et al.,** ***Electrochemical oxidation of phenolic wastes with boron-doped diamond anodes.*** **Water Res, 2005. 39(12): p. 2687-703.**
107. **Zhu, X.P., et al.,** ***Electrochemical oxidation characteristics of p-substituted phenols using a boron-doped diamond electrode.*** **Environmental Science & Technology, 2007. 41(18): p. 6541-6546.**
108. **Komatsu, M., T.N. Rao, and A. Fujishima,** ***Detection of hydroxyl radicals formed on an anodically polarized diamond electrode surface in aqueous media.*** **Chemistry Letters, 2003. 32(4): p. 396-397.**

109. **Troster, I., et al., *Electrochemical advanced oxidation process using DiaChem (R) electrodes.* Water Science and Technology, 2004. 49(4): p. 207-212.**

110. **Hervey, W.J.t., M.B. Strader, and G.B. Hurst, *Comparison of Digestion Protocols for Microgram Quantities of Enriched Protein Samples.* J Proteome Res, 2007. 6(8): p. 3054-3061.**

111. **Vaisar, T. and J. Urban, *Probing the proline effect in CID of protonated peptides.* Journal of Mass Spectrometry, 1996. 31(10): p. 1185-1187.**

112. **Pan, C. and R.L. Hettich, *Multipole-storage-assisted dissociation for the characterization of large proteins and simple protein mixtures by ESI-FTICR-MS.* Analytical Chemistry, 2005. 77(10): p. 3072-3082.**

113. **Freund, J., Kalbitzer, H. R., *Physiological buffers for NMR spectroscopy.* J Biomol NMR, 1995. 5: p. 321-322.**

114. **Compilation, Notre Dame Radiation Laboratory Chemistry Data Center: p. http://www.rcdc.nd.edu/compilations/Hydroxyl/OH.htm.**

115. **Hofmeister, F., Arch Exp Pathol Pharmacol, 1888. 24: p. 247-260.**

116. **Ramirez-Alvarado, M., F.J. Blanco, and L. Serrano, *Elongation of the BH8 beta-hairpin peptide: Electrostatic interactions in beta-hairpin formation and stability.* Protein Science, 2001. 10(7): p. 1381-1392.**

117. **Gao, J., et al., *Changes in the protein expression of yeast as a function of carbon source.* Journal of Proteome Research, 2003. 2(6): p. 643-649.**

118. **Liu, H.B., R.G. Sadygov, and J.R. Yates, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics.* Analytical Chemistry, 2004. 76(14): p. 4193-4201.**

119. **Thompson, M.R., D.K. Thompson, and R.L. Hettich, *Systematic assessment of the benefits and caveats in mining microbial post-translational modifications from shotgun proteomic data: the response of Shewanella oneidensis to chromate exposure.* J Proteome Res, 2008. 7(2): p. 648-58.**

120. **Washburn, M.P., D. Wolters, and J.R. Yates, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature Biotechnology, 2001. 19(3): p. 242-247.**

121. **Xu, G., et al., *Secondary reactions and strategies to improve quantitative protein footprinting.* Anal Chem, 2005. 77(10): p. 3029-37.**

122. **Aitken, A. and M. Learmonth, *Protein Determination by UV Absorption.* The Protein Protocols Handbook, 1996: p. 3-6.**

123. **Snook, M.E. and G.A. Hamilton, *Oxidation and Fragmentation of Some Phenyl-Substituted Alcohols and Ethers by Peroxydisulfate and Fentons Reagent.* Journal of the American Chemical Society, 1974. 96(3): p. 860-869.**

124. **Sethna, S.M., *The Elbs Persulfate Oxidation.* Chemical Reviews, 1951. 49(1): p. 91-101.**

125. **Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. 157(1): p. 105-32.**

126. **Tong, X., J.C. Wren, and L. Konermann, *Effects of protein concentration on the extent of gamma-ray-mediated oxidative Labeling studied by electrospray mass spectrometry.* Analytical Chemistry, 2007. 79(16): p. 6376-6382.**

127. **Burke, T.W., et al., *Strong cation-exchange high-performance liquid chromatography of peptides. Effect of non-specific hydrophobic interactions and linearization of peptide retention behaviour.* J Chromatogr, 1989. 476: p. 377-89.**

128. **Mirzaei, H. and F. Regnier, *Identification and quantification of protein carbonylation using light and heavy isotope labeled Girard's P reagent.* Journal of Chromatography A, 2006. 1134(1-2): p. 122-133.**
129. **Mirzaei, H. and F. Regnier, *Enrichment of carbonylated peptides using Girard P reagent and strong cation exchange chromatography.* Analytical Chemistry, 2006. 78(3): p. 770-778.**
130. **Zhang, Y., et al., *Effect of Dynamic Exclusion Duration on Spectral Count Based Quantitative Proteomics.* Anal Chem, 2009.**
131. **Bridgewater, J.D., et al., *The effect of histidine oxidation on the dissociation patterns of peptide ions.* Journal of the American Society for Mass Spectrometry, 2007. 18(3): p. 553-562.**
132. **Boys, B.L., M.C. Kuprowski, and L. Konermann, *Symmetric behavior of hemoglobin alpha- and beta- subunits during acid-induced denaturation observed by electrospray mass spectrometry.* Biochemistry, 2007. 46(37): p. 10675-84.**
133. **Cech, N.B. and C.G. Enke, *Relating electrospray ionization response to nonpolar character of small peptides.* Anal Chem, 2000. 72(13): p. 2717-23.**
134. **Sayre, R.M., J.C. Dowdy, and M. Poh-Fitzpatrick, *Dermatological risk of indoor ultraviolet exposure from contemporary lighting sources.* Photochem Photobiol, 2004. 80: p. 47-51.**
135. **Parker, N.R., et al., *Protein-bound kynurenine is a photosensitizer of oxidative damage.* Free Radical Biology and Medicine, 2004. 37(9): p. 1479-1489.**
136. **Keller, A., et al., *Experimental protein mixture for validating tandem mass spectral analysis.* Omics, 2002. 6(2): p. 207-12.**
137. **Leitner, A. and W. Lindner, *Current chemical tagging strategies for proteome analysis by mass spectrometry.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2004. 813(1-2): p. 1-26.**
138. **Braisted, J.C., et al., *The APEX Quantitative Proteomics Tool: Generating protein quantitation estimates from LC-MS/MS proteomics results.* Bmc Bioinformatics, 2008. 9: p. -.**
139. **Sandra, K., et al., *Highly efficient peptide separations in proteomics Part 2: Bi- and multidimensional liquid-based separation techniques.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2009. 877(11-12): p. 1019-1039.**
140. **Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomed Mass Spectrom, 1984. 11(11): p. 601.**
141. **Olsen, J.V., S.E. Ong, and M. Mann, *Trypsin cleaves exclusively C-terminal to arginine and lysine residues.* Mol Cell Proteomics, 2004. 3(6): p. 608-14.**
142. **Wuhrer, M., A.M. Deelder, and C.H. Hokke, *Protein glycosylation analysis by liquid chromatography-mass spectrometry.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2005. 825(2): p. 124-133.**
143. **Morelle, W., et al., *The use of mass spectrometry for the proteomic analysis of glycosylation.* Proteomics, 2006. 6(14): p. 3993-4015.**
144. **Tsur, D., et al., *Identification of post-translational modifications via blind search of mass-spectra.* Proc IEEE Comput Syst Bioinform Conf, 2005: p. 157-66.**
145. **Wilmarth, P.A., et al., *Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility?* J Proteome Res, 2006. 5(10): p. 2554-66.**
146. **Zwickl, P., D. Voges, and W. Baumeister, *The proteasome: a macromolecular assembly designed for controlled proteolysis.* Philos Trans R Soc Lond B Biol Sci, 1999. 354(1389): p. 1501-11.**

147. **Hanson, L.A., et al.,** ***Protective factors in milk and the development of the immune system.*** **Pediatrics, 1985. 75(1 Pt 2): p. 172-6.**

148. **Srikanth, R., et al.,** ***Improved sequencing of oxidized cysteine and methionine containing peptides using electron transfer dissociation.*** **Journal of the American Society for Mass Spectrometry, 2007. 18(8): p. 1499-1506.**

149. **Srikanth, R., J. Wilson, and R.W. Vachet,** ***Correct identification of oxidized histidine residues using electron-transfer dissociation.*** **Journal of Mass Spectrometry, 2009. 44(5): p. 755-762.**

150. **Wiesner, J., T. Premsler, and A. Sickmann,** ***Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications.*** **Proteomics, 2008. 8(21): p. 4466-4483.**

Appendix

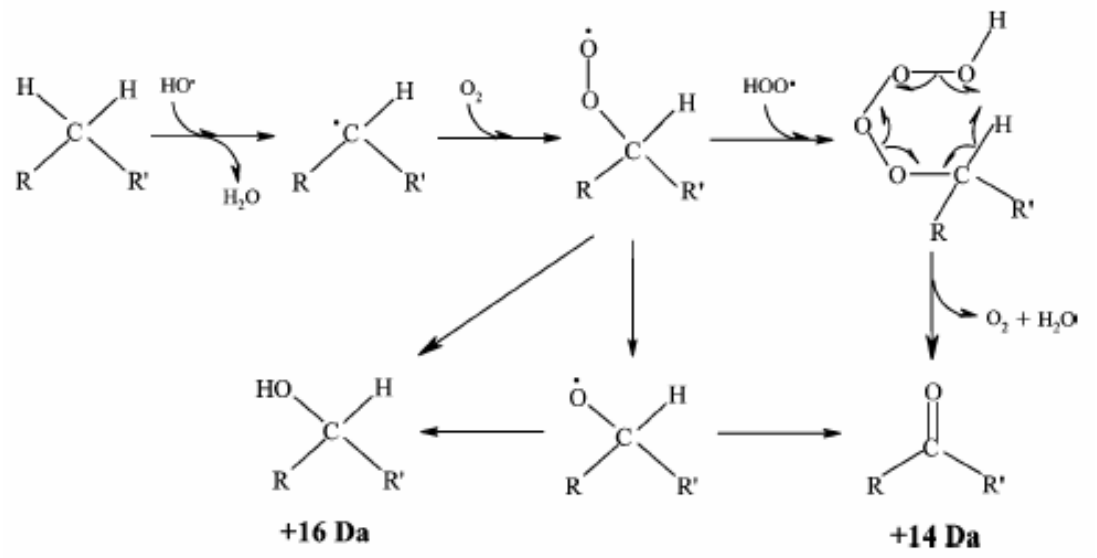
Amino Acid Oxidation Rates and Mechanisms

The following table and figures were taken directly from **Xu, G. and M.R. Chance, *Hydroxyl radical-mediated modification of proteins as probes for structural proteomics.* Chem Rev, 2007. 107(8): p. 3514-43. [72]** Please refer to this critical review article for a detailed discussion of the protein oxidation chemistry depicted below. General mechanism schemes are followed by specific amino acid oxidation schemes in order of increasing •OH reactivity.

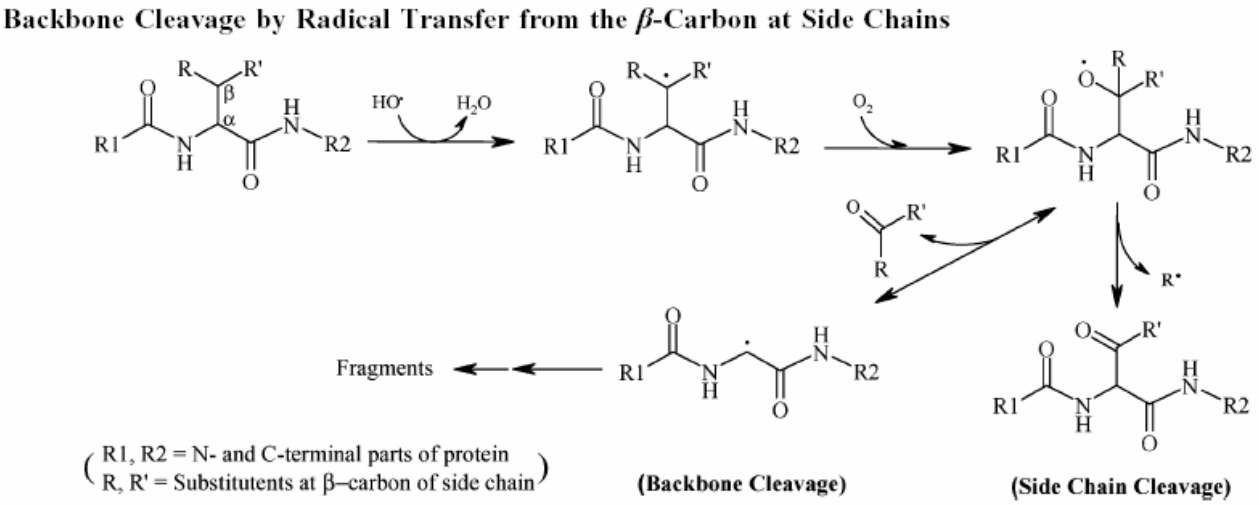
Table A-9. Rate Constants for Reaction of Free Amino Acids with Hydroxyl Radical

substrate	rate (M^{-1} s^{-1})	pH
Cys	3.5×10^{10}	7.0
Trp	1.3×10^{10}	6.5−8.5
Tyr	1.3×10^{10}	7.0
Met	8.5×10^{9}	6−7
Phe	6.9×10^{9}	7−8
His	4.8×10^{9}	7.5
Arg	3.5×10^{9}	6.5−7.5
cystine	2.1×10^{9}	6.5
Ile	1.8×10^{9}	6.6
Leu	1.7×10^{9}	~6
Val	8.5×10^{8}	6.9
Pro	6.5×10^{8}	6.8
Gln	5.4×10^{8}	6.0
Thr	5.1×10^{8}	6.6
Lys	3.5×10^{8}	6.6
Ser	3.2×10^{8}	~6
Glu	2.3×10^{8}	6.5
Ala	7.7×10^{7}	5.8
Asp	7.5×10^{7}	6.9
Asn	4.9×10^{7}	6.6
Gly	1.7×10^{7}	5.9

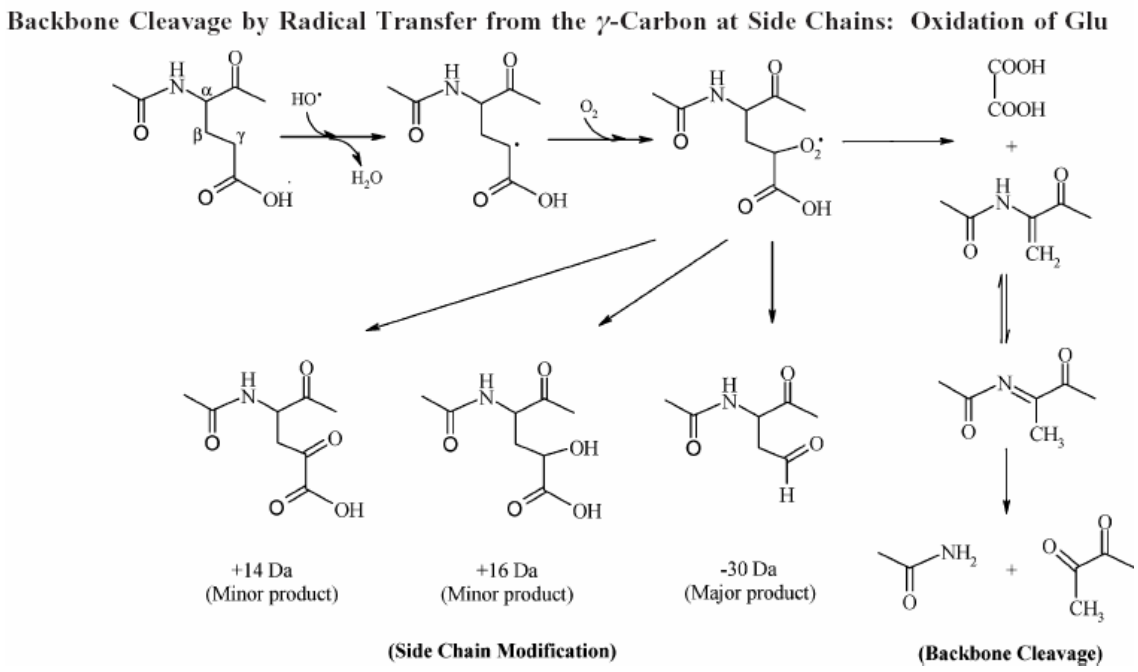
Scheme A-1: General Mechanism for Oxidation of Hydrocarbon Side Chains



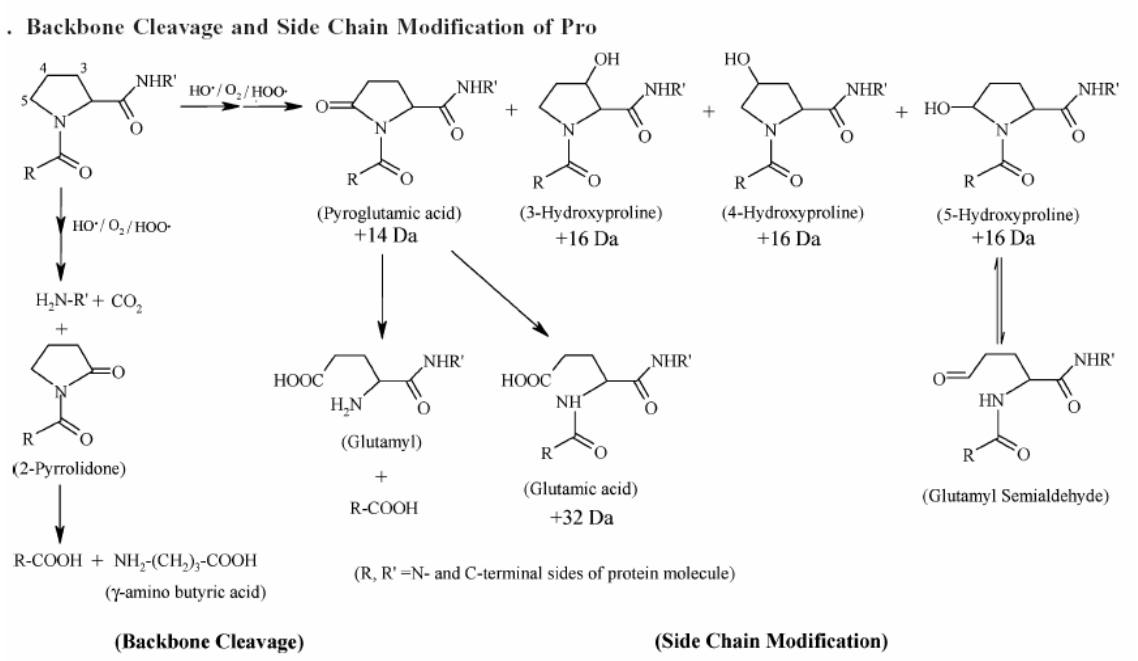
Scheme A-2: Oxidative Peptide Backbone Cleavage and Side Chain Oxidation Products



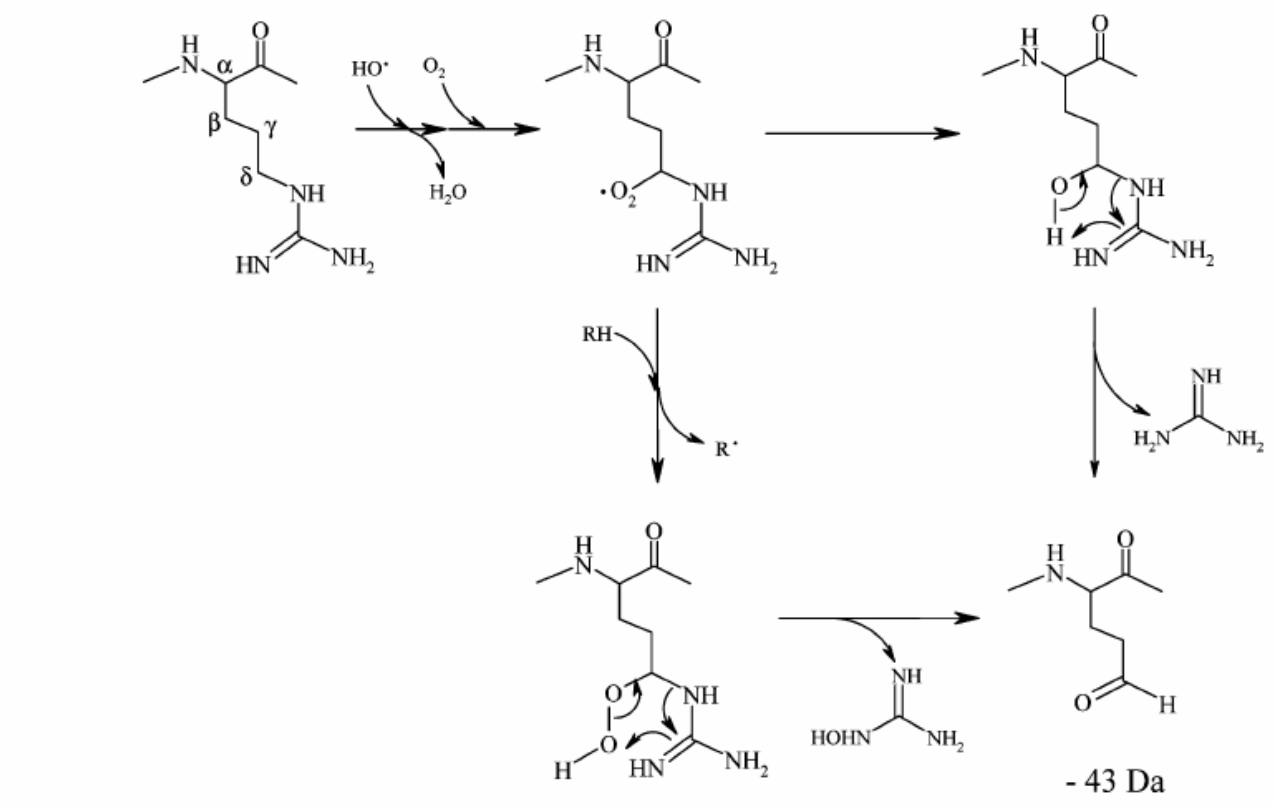
Scheme A-3: Glutamic Acid Oxidation Products



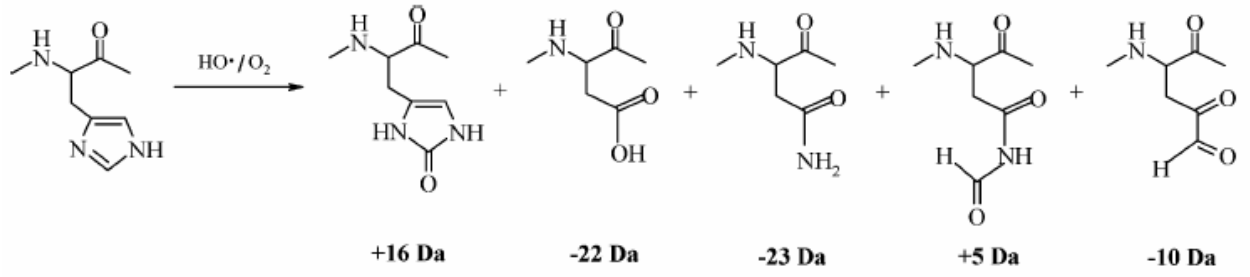
Scheme A-4: Proline Oxidation Products



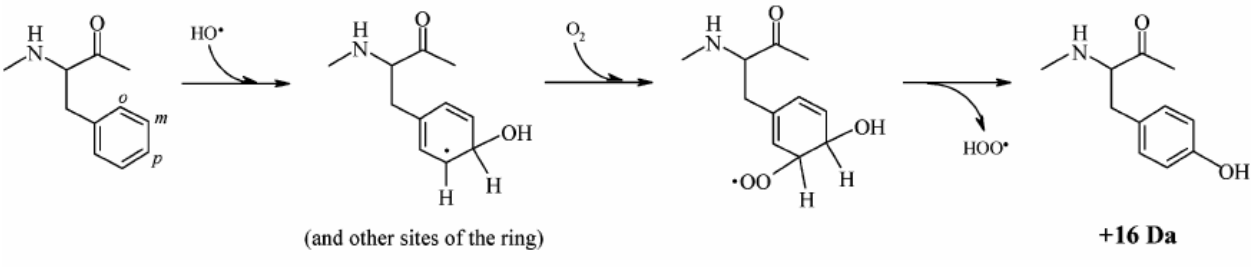
Scheme A-5: Arginine Oxidation Products



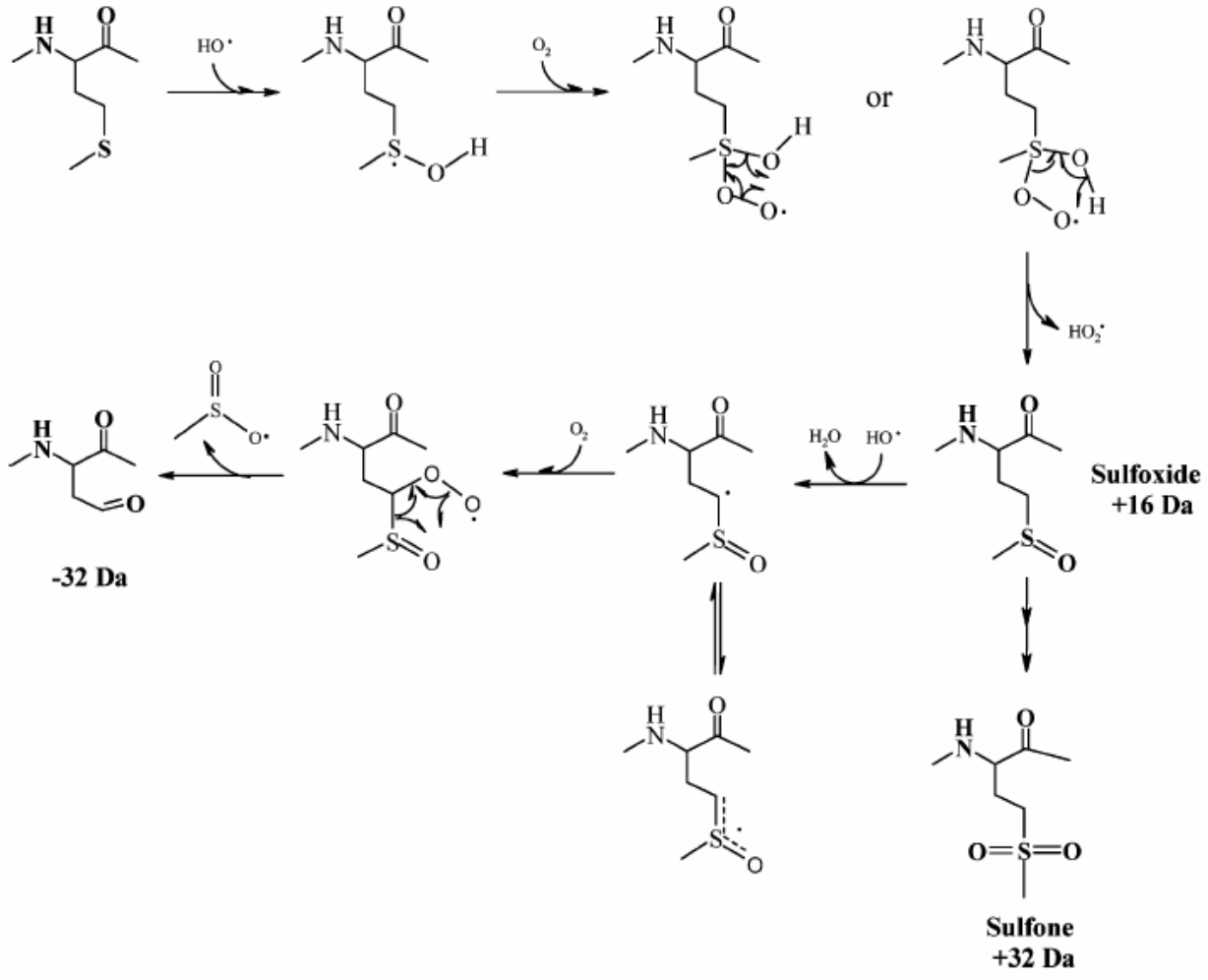
Scheme A-6: Histidine Oxidation Products



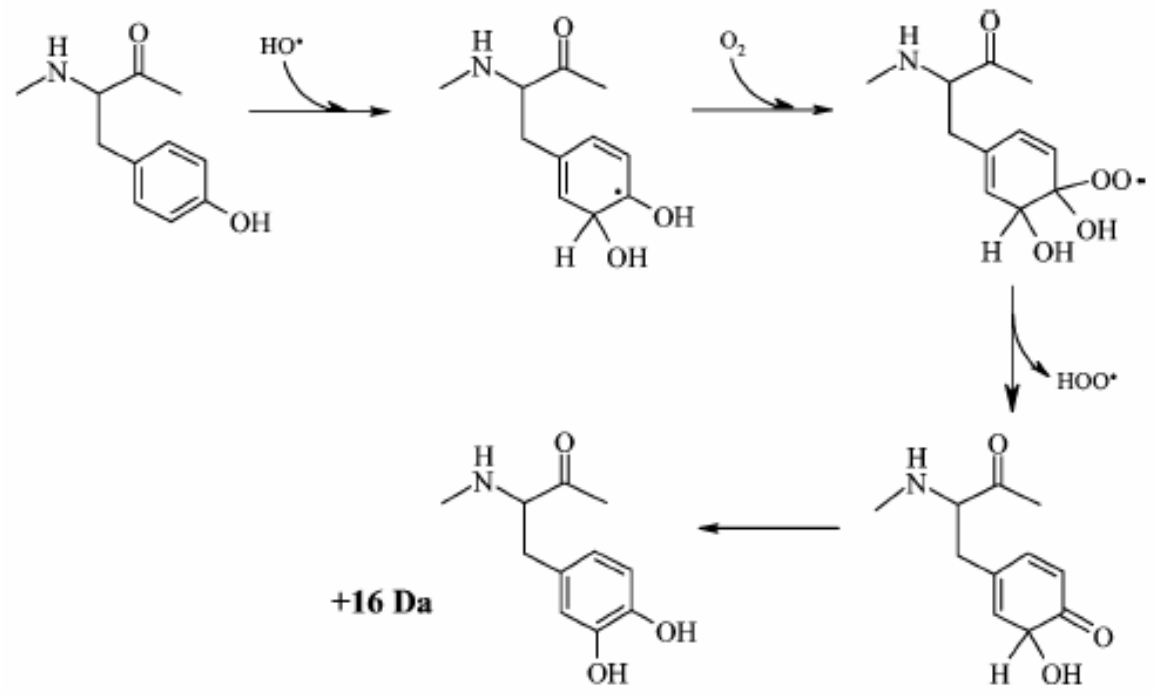
Scheme A-7: Phenylalanine Oxidation Products



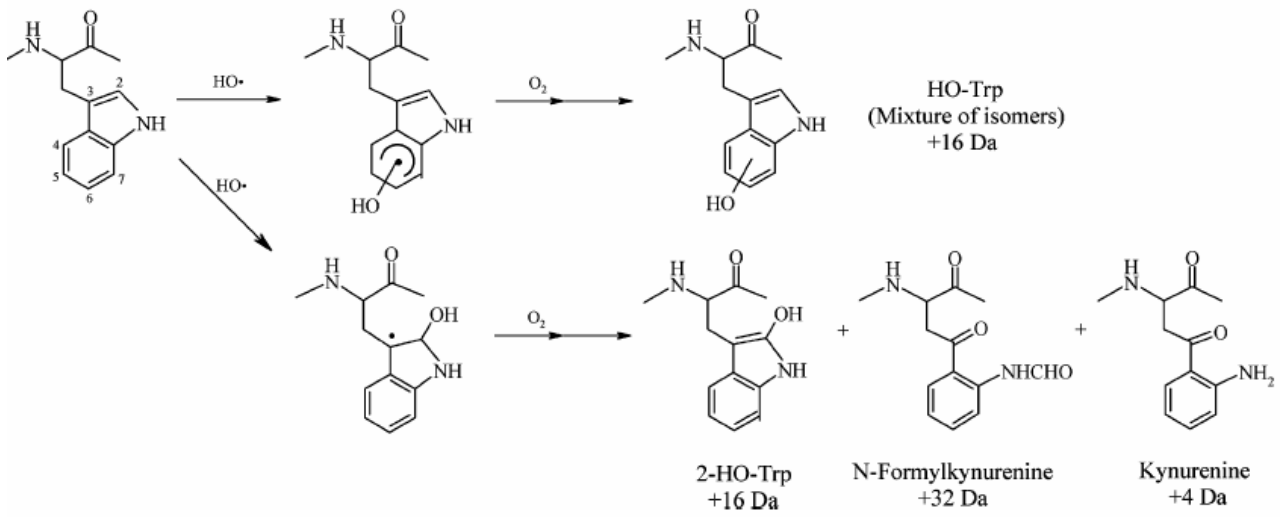
Scheme A-8: Methionine Oxidation Products



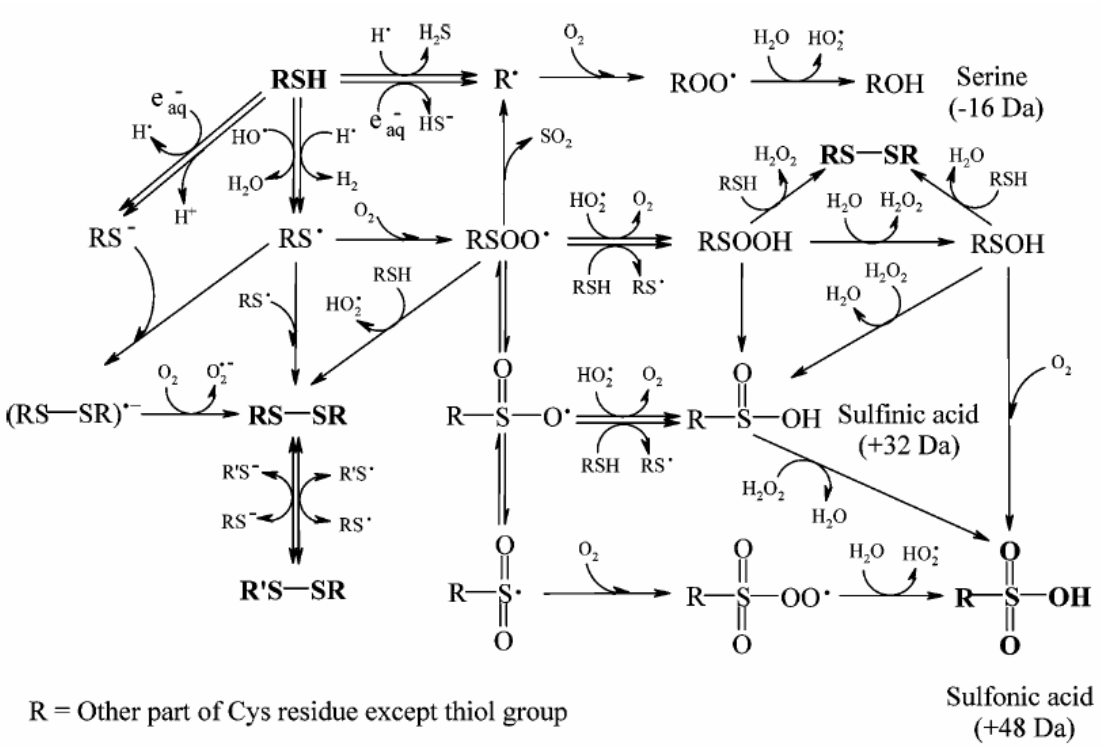
Scheme A-9: Tyrosine Oxidation Products



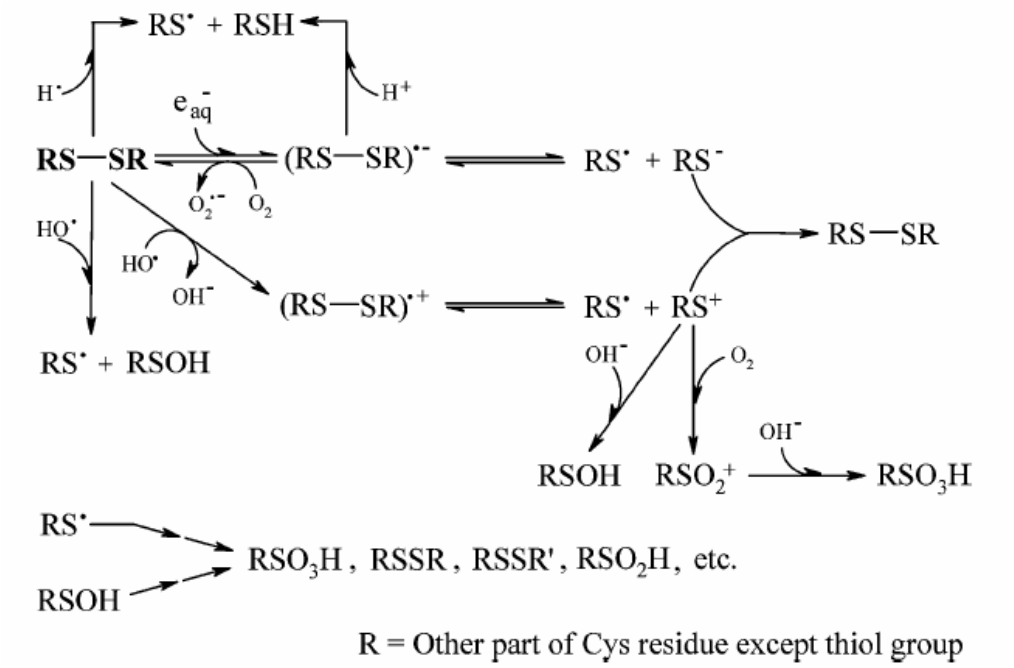
Scheme A-10: Tryptophan Oxidation Products



Scheme A-11: Cysteine Oxidation Products



Scheme A-12: Disulfide Oxidation Products



Vita

Carlee McClintock, maiden name Patterson, completed her Bachelor of Science degree in Neurobiology in 2004 and moved to Knoxville shortly after her marriage to David McClintock near their alma mater, the University of Texas at Austin. Longhorn orange still burns vibrantly at their home nestled in the East Tennessee woods only a few miles away from the bright orange lights of Volunteer territory.