1-1-2011

Mass Spectrometry for Determination of Conformation and Dynamics of Proteins and Structure and Biosynthesis of Bacterial Peptidoglycan

Jiawei Chen
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Chemistry

Dissertation Examination Committee:
Michael L. Gross, Chair
Robert Blankenship
Cynthia Lo
Jacob Schaefer
Xiong Su
John-Stephen A. Taylor

Mass Spectrometry for Determination of Conformation and Dynamics of Proteins and Structure and Biosynthesis of Bacterial Peptidoglycan

by

Jiawei Chen

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2011

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Mass Spectrometry for Determination of Conformation and Dynamics of Proteins and Structure and Biosynthesis of Bacterial Peptidoglycan

By

Jiawei Chen

Doctor of Philosophy in Chemistry

Washington University in St. Louis, 2011

Professor Michael L. Gross, Chair

Mass spectrometry (MS) has emerged as an important tool for analyzing and characterizing large biomolecules. In this dissertation, two aspects of the development and application of MS-based approaches are presented; they include (1) protein conformation and folding dynamics (in Chapters 2 to 5) and bacterial peptidoglycan (PG) structure and biosynthesis (Chapters 6 to 8). Chapter 1 serves as the introduction for both aspects.

Part I of the dissertation focuses on the development of analytical methods combining fast photochemical oxidation of proteins (FPOP) and mass spectrometry analysis. In chapter 2 to 4, we discuss protein folding with sub-millisecond time resolution by a new pump/probe procedure. Perturbations in protein structure are by temperature jump of the protein solution, followed by fast photochemical oxidation of proteins (FPOP) as the probe. The hydroxyl radical lifetime was predicted by a

dosimeter experiment (Chapter 2). The T jump-induced folding constant was measured at the global protein level (Chapter 3), and the residue level detail was revealed by proteolysis and liquid chromatography-mass spectrometry (LC-MS) (Chapter 4). Chapter 5 discusses the development of a new FPOP reagent, iodobenzoic acid, and its application on studying conformational differences between apo- and holo- proteins.

Part II of the thesis discusses the development and application of MS-based methods to investigate bacterial peptidoglycan. Chapter 6 focuses on the methodology of the bottom-up MS to characterize the fine structure of *enterococcus faecium* (*E. faecium*) peptidoglycan. Furthermore, we developed a time-dependent isotopic labeling strategy and applied it to *E. faecium* during the cell wall growth to determine quantitatively the percentage of heavy isotope incorporation into different muropeptides through peptidoglycan growth cycles, discussed in chapter 7. The results are important for understanding tertiary structure and designing novel drugs for antibiotic-resistant pathogens. In chapter 8, we applied the above approaches to investigate methicillin-resistant *staphylococcus aureus* (*S. aureus*) and its *fem*-mutants. We emphasize the peptidoglycan composition, fine structures, and biosynthesis.

ACKNOWLEDGEMENTS

I thank my advisor, Prof. Michael Gross, for his support during my entire thesis research. He gave me freedom to explore on my own, and at the same time guided me through difficulties. He has been not only a terrific mentor but also a caring father figure to me. My gratitude to him can never be overstated.

I thank my colleagues, Don Rempel, Henry Rohrs, Daryl Giblin, Ilan Vidavsky, Weidong Cui, Joyce Neff, Jun Zhang, Lisa Jones, Manolo Plasencia, Bo Wang, Sergio Carstro, Justin Sperry, Adam Bruskern, Dian Su, Brian Gau, Hao Zhang, Richard Huang, Yining Huang, Yuetian Yan and Ying Zhang for their kind support, and their effort to create a lab environment that is so enjoyable. Thanks to Brian, Weidong, Daryl for helping with my research projects, and Henry, for encouraging me when I fail, helping me when I need, and most importantly, sharing a friendship during the past five years. A special thanks to Don for his persistence and support in my project and the endless supply of chocolates and nuts during working nights – something to be memorized.

I also would like to acknowledge my collaborators, Profs. Jacob Schaefer, John-Stephen Taylor and Drs. Gary Patti, Sung Joon Kim, Shasaad Sharif, Manmilan Singh and Rongsheng Wang for the productive and pleasant collaboration. Thanks to my committee members, Professors Schaefer, Taylor, Blankenship, Lo and Su for their advice and discussions regarding my thesis research.

This thesis research is supported by the National Institute of Health and National Center for Research Resources (Grant No. 2P41RR000954).

I thank my friends for sharing my happiness and sadness, helping me through the difficult times, and cheering me up whenever I am down. Life would be tough without them.

A special thanks to my parents for having me, supporting me and understanding me. Although we are thousands of miles apart, I know that their unconditional love is always right here with me. Finally, thanks to Xin for loving me for who I am and letting me be who I want to be.

Life is a journey. Thanks to everyone that I know/knows me for being part of it.

Table of Contents

Table of Figures

Table of Tables

Table of Schemes

1. Introduction

Part I: Mass-Spectrometry-Based Studies of Protein Conformation and Dynamics

1.1 Protein Structure and Conformation Analysis

1.1.1 General Methods

The determination of the structure and conformation of a protein is essential to understanding its biological function. The approaches include the high resolution methods of X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, the low resolution methods of circular dichroism spectroscopy and fluorescence spectroscopy, and the intermediate-level methods of mass spectrometry, which is more recently developed.

X-ray crystallography and NMR are the two main approaches that provide structural information of macromolecules at atomic resolution. X-ray crystallography has been a vital tool for studying protein structure since the late 1950s (*1*). It produces a three-dimensional picture of the density of electrons within the protein crystal, thus revealing positions of the atoms. It is now used routinely to determine protein structure and protein-protein or protein-ligand interactions (*2*). It is limited, however, by the difficulty of getting some proteins, especially intrinsic membrane proteins, to form crystals. The nearest competing method, NMR, is carried out in solution, thus eliminating the crystallization problem (*3*). Furthermore, NMR provides additional information about protein folding, dynamics and conformational equilibria (*4-6*). Given that there are problems in resolving overlapping peaks in larger proteins, NMR spectroscopy has been restricted to relatively small proteins (*7*). NMR also requires

significant quantities (milligram) of protein samples at high concentration (~100 μM); under conditions of high concentration, proteins can aggregate and precipitate.

Several additional resolution methods have been developed over the last decades to obtain general, less-detailed information about protein structure and conformation. CD arises from differential absorbance of polarized light by the sample, and is the most common method for providing a general picture of secondary structure (*8, 9*). Fluorescence and Förster resonance energy transfer (FRET) are used to investigate protein conformational changes and protein-protein/protein-ligand interactions (*10-12*). They can be highly sensitive, cost-efficient, and quantitative. Other low resolution methods include FTIR, Raman, and cryo-electron microscopy.

1.1.2 Mass Spectrometry-based Footprinting

Mass spectrometry has emerged during the last two decades as an important tool for analyzing proteins and peptides, spurred by the groundbreaking invention of electrospray ionization (*13*) and matrix-assisted laser desorption ionization (MALDI) in the 1980s. Both of these ionization methods allow the transfer of biomolecules into the gas phase as intact species usually bearing some excess charge from protons, making them amenable to detection by MS (*14*). Despite its high sensitivity (femtomole level), MS has other advantages owing to its capability to analyze not only intact proteins but also constituent peptides after proteolysis. By incorporating MS/MS methods, information at the residue level can be obtained without using the time-consuming method of site-directed mutagenesis.

Mass spectrometry-based footprinting serves as an intermediate level approach in studying protein structure and conformation. Although it does not provide atomic-level definition, protein footprinting assays protein conformation by either selectively modifying or cleaving amino acid residues. It uses changes in chemical reactivity that is based on sensitivity of reactivity to solvent accessibility of various amino-acid residues. A typical footprinting experiment determines those regions of a protein that undergo a change in solvent accessibility between two or more states (e.g., apo- *vs.* holo- , folded *vs.* unfolded). The readout, usually a mass shift, can be measured by MS often in combination with proteolysis to give peptides, which can be separated and definitively analyzed in an LC/MS experiment.

Hydrogen deuterium exchange (HDX) is a footprinting technique that probes the solvent accessibility and hydrogen bonding environment of amide backbone hydrogens (*15, 16*). By monitoring the global deuterium uptake, the net change of a protein upon unfolding, rearrangement, or ligand-binding can be detected by MS measurement in comparison with a control protein sample. By proteolysis and LC-MS, deuterium uptake at the peptide level can be revealed. Recently, with the development of electron-transfer dissociation (ETD) and electron-capture dissociate (ECD) as fragmentation method, one can minimize any amide backbone hydrogen scrambling (*17*) and obtain information at the residue-level (*18, 19*).

1.2 Hydroxyl Radical Footprinting

1.2.1 Principles and Methods

The hydroxyl radical is emerging as a widely-applicable footprinting reagent owing to its easy-formation, high reactivity and low selectivity (*20*). Hydroxyl radical probes solvent accessibility of amino-acid sidechains by oxidizing them, and such modification is preserved throughout the following proteolysis and LC-MS/MS experiments, allowing time for peptide and amino acid level data acquisition. Its advantage over H/DX is that the modification is usually irreversible (i.e., there is likely to be no scrambling).

There are different pathways for hydroxyl radical labeling on sidechains, among which the incorporation of oxygen is the major one (*21*-*24*). Table 1.1 lists the common modifications observed in hydroxyl-radical labeling experiments. Although hydroxyl radicals are non-specific reactants, they do not label sidechains with equal efficiency, as shown by the wide range of rate constants that underlie their reactivity, as indicated in the table.

amino acid	$k_{\bullet OH}$ (M^{-1} sec^{-1})	common modifications (Da)				
Cys	3.5×10^{10}	–15.9772	+31.9898	+47.9847		
Trp	1.3×10^{10}	+3.9949	+15.9949	+31.9898	+47.9847	
Tyr	1.3×10^{10}	+15.9949	+31.9898	+47.9847		
Met	8.5×10^{9}	–32.0085	+15.9949	+31.9898		
Phe	6.9×10^{9}	+15.9949	+31.9898	+47.9847		
His	4.8×10^{9}	–23.0160	–22.0320	–10.0320	+4.9789	+15.9949
Arg	3.5×10^{9}	–43.0534	+13.9793	+15.9949		
Ile	1.8×10^{9}	+13.9793	+15.9949			
Leu	1.7×10^{9}	+13.9793	+15.9949			
Val	8.5×10^{8}	+13.9793	+15.9949			
Pro	6.5×10^{8}	+13.9793	+15.9949			
Gln	5.4×10^{8}	+13.9793	+15.9949			
Thr	5.1×10^{8}	–2.0157	+15.9949			
Lys	3.5×10^{8}	+13.9793	+15.9949			
Ser	3.2×10^{8}	–2.0157	+15.9949			
Glu	2.3×10^{8}	–30.0106	+13.9793	+15.9949		
Ala	7.7×10^{7}	+15.9949				
Asp	7.5×10^{7}	–30.0106	+15.9949			
Asn	4.9×10^{7}	+15.9949				
Gly	1.7×10^{7}	n.d.				

Table 1.1: Initial •OH-amino acid side-chain reaction rates and common mass spectrometry-observed products of •OH-mediated protein footprinting (*24*, *25*).

There are many ways of generating hydroxyl radicals for protein footprinting, and most were reviewed by Chance et al. (*24*). They include electron-pulse radiolysis, synchrotron radiolysis, use of the Fenton reaction, and laser photolysis. The synchrotron radiolysis of water with X-ray photons involves photoionization, where the energy of a photon is transferred to water, causing its ionization to make a radical cation, which loses H_3O^+ to give ·OH (eq 1) (*26*).

$$H_2O + \text{ionizing radiation} \longrightarrow H_2O^{\cdot +} + e^-_{dry} \xrightarrow{H_2O} H_3O^+ + {}^{\cdot}OH + e^-_{aq} \quad (1)$$

The advantage of this method is that no reagents need to be added to the solutions. It is limited, however, by the availability of a synchrotron radiation source and the cost of operation. The Fenton reaction is one of the inexpensive ways to generate hydroxyl radicals. In this reaction, hydroxyl radicals are produced from hydrogen peroxide by oxidizing EDTA-chelated Fe^{2+} to Fe^{3+}. A major disadvantage is that it is a continuous labeling approach and susceptible to overlabeling. It is used as part of a standard method for studying DNA and RNA interactions (*27*).

1.2.2 Fast Photochemical Oxidation of Proteins (FPOP)

Hambly and Gross (*28, 29*) developed a fast-labeling, pulsed method, FPOP, utilizing a 248 nm laser beam to dissociate hydrogen peroxide into hydroxyl radicals (eq 2).

$$H_2O_2 \xrightarrow{h\nu} 2\,{}^{\cdot}OH \quad (2)$$

The protein solution mixed with hydrogen peroxide and glutamine flows through a fused silica capillary and is irradiated by an excimer laser at a frequency that ensures all sample protein is exposed to the beam only once. The pulse rate is adjusted so that there is a measurable exclusion fraction (Figure 1.1) (i.e., a region that is not exposed to the laser and serves to isolate the plugs of solution that are oxidatively modified by the radicals.

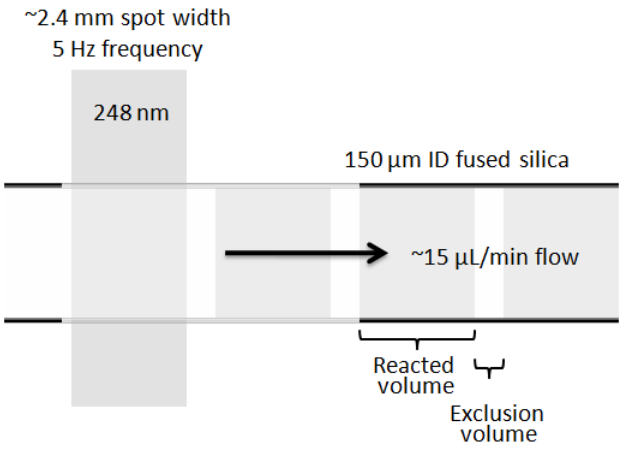


Figure 1.1: Schematic of the FPOP reaction region with normal settings. The laser beam intersects the flow tube (ID = 150 µm) at right angles. Its frequency and the flow rate are adjusted to insure that the protein is only exposed once to radicals and to provide an exclusion region or buffer between two successively irradiated plugs.

One advantage of FPOP is that the reaction time is controlled by adding glutamine as a radical scavenger. Kinetic analysis reveals that by one microsecond, the free [•OH] is almost all consumed (*28*). Thus, for proteins that are structurally sensitive to oxidation, any FPOP-perturbed conformations are not sampled by the labeling. Gau et

al. (*30*) further demonstrated that the protein after FPOP under proper control shows that the distribution of modified states (i.e., 0, +16, +32…) fit a Poisson distribution, indicating that the protein population shares an invariant conformation during the oxidative modification. In the absence of the glutamine scavenger, the protein distribution is shifted to higher states of modification and is not well described by a Poisson distribution, indicating that partial unfolding can occur during the timescale of modification if proper controls are not made.

1.3 Protein Folding

1.3.1 Principles

The folding of a protein to its native compact structure is a fundamental biological process. The correct fold that leads to a proper structure is critical for a protein to function effectively. The process is far from perfect, however, and sometimes results in a fraction of proteins failing to fold correctly. Cells have mechanisms to eliminate this problem by either degrading the incorrect structures or motivating refolding into correct structures. When the cells fail to do so, the accumulation of misfolded proteins can lead to some serious diseases including Alzheimer's, Huntington's, Parkinson's, and prion-related diseases (*31, 32*). Investigating the mechanisms of protein folding is crucial to understanding the misfolding problem from a fundamental viewpoint. Moreover, more knowledge about the folding mechanisms may provide insight for protein structure prediction.

It is believed that the native state of a protein usually corresponds to the most thermodynamically stable structure (*33*). Although the classic view of protein folding is

based on models involving folding intermediates in a structural way and sometimes are regarded as a linear sequence of events, a new view replaces the "folding pathway" idea with an "energy landscape and folding funnels", emphasizing the ensemble nature of protein conformations. Thus, folding can be viewed as a multi-pathway, diffusion-like process (*34*, *35*). Studies of a series of small proteins with 60-100 amino acid residues has suggested that the fundamental mechanism of protein folding involves the interaction of a relatively small number of residues to form a folding nucleus, around which the rest of the structure rapidly condenses (*36*). The incorporation of experimental measurements into computer simulation shows that these transition states, although highly heterogeneous, share a similar overall topology as that of the native structure, thus indicating that the final folding to the native structure happens invariably after the correct topology is formed (*37*). For larger proteins, one or more intermediates can be often observed during the folding process (*38*). In such cases, proteins fold in modules and establish folds within local regions before the fully native structures are formed (*39*).

Protein folding can be as rapid as microseconds to milliseconds, thus presenting a challenge for analysis. To investigate the kinetics and the structures of transition states in such fast-folding events, several techniques have been developed to study fast folding by accommodating the folding speed with high time and temporal resolution.

1.3.2 Protein Folding Studied at Atomic Resolution

Experimental techniques have been developed to study protein folding since the 1950s. Folding probes generally include low resolution ones such as CD and fluorescence, and high resolution ones that are emphasized in this section.

NMR spectroscopy plays an importance role in studying protein folding. Early NMR investigations focused on the equilibrium conversion between the unfolded state and the native state under severe conditions; this approach can provide thermodynamics information (*40*). Additionally, NMR spectroscopy of non-native states determines the rates of conversion between different states of folding, although the system is at equilibrium. Slow folding on the second timescale and above can be determined directly by NMR exchange spectroscopy, whereas fast folding on a microsecond to millisecond timescale are accessible by line shape analyses of NMR (*41*). Non-equilibrium reactions can directly be followed by NMR combined with stopped-flow device to observe real-time protein folding (*42, 43*). The recent development of real-time 1D and 2D NMR techniques allow simultaneous observation of reaction kinetics for a large number of nuclear sites along the polypeptide chain of a protein with an unprecedented time resolution of a few seconds (*44*). To complement solution NMR on studying inhomogeneous unfolded or partially folded states, solid-state NMR methods have been developed and applied to frozen solutions where non-equilibrium, transient states are trapped. It was demonstrated that solid-state NMR can provide information about folding processes that is not obtained by other methods (*45*).

The only experimental technique available for fine structure characterization of protein folding transition states is the Φ value analysis. It uses site-directed mutagenesis to alter the direct interaction of a sidechain of a protein with a ligand or with other residues in the protein, perturbing the interaction. Then by comparing the rate and equilibrium of the native protein and the mutant, it can be determined whether the

interaction is formed in an intermediate (*46*). The Φ value is calculated following equation 3.

$$\Phi = \Delta\Delta G_{\ddagger\text{-D}} / \Delta\Delta G_{\text{N-D}} \qquad (3)$$

$\Delta\Delta G_{\ddagger\text{-D}}$ and $\Delta\Delta G_{\text{N-D}}$ are the change of free energy of activation and the free energy of folding caused by mutation (*47*). A Φ value of 1 implies that the structure at the mutated site is native-like, while a Φ value of 0 means the structure at the mutated site is denatured. Φ values in between imply that either partial folding is reached in the transition or a mixture of structure is formed (*48*). This technique has high spatial resolution, yet it is relatively time and cost-consuming. The sites to mutate must be selected carefully to avoid structure perturbation at other regions.

1.3.3 Rapid Folding Initiation Techniques

One approach to study protein folding or unfolding is to introduce a sudden change in the solvent environment of the protein to shift rapidly the equilibrium constant. The response of the protein is then measured by the techniques described above.

The earliest and also the most widely used technique for folding initiation is rapid mixing (*49-51*). In a stop-flow mixing device, a solution containing protein and denaturant is combined with a buffer to induce refolding, and the consequence of mixing is then detected by techniques such as circular dichroism spectrometer and fluorescence spectrometer. The time resolution of the conventional mixing device is limited by insufficient turbulence to a few milliseconds. Later, with a number of innovative designs, the dead time can be pushed down to microseconds. These designs include: 1. wider tubes for transporting the two streams to the tip-shaped mixing chamber, at which site is a

mixing sphere as small as a few micrometers (*52*). 2. a mixer with narrower cross-section (<100 μm) incorporated into the classic mixing device (*53*), 3. a combination of a mixing sphere and a continuous-flow mixing head of high stability (*54*, *55*).

Temperature jump can also be applied to study protein fast folding/unfolding and peptide secondary structure formation (*56*-*59*). There are two major T jump methods, electrical-discharge-induced T jump with rise times of 500 ns to 10 μs[54], and laser-induced T jump on the nanosecond time scale (*60*). Given that the fastest possible secondary structure formation is 100 ns for helices (*58*) and 6 μs for β-sheets (*61*), the T jump approach is required for investigating such events. Besides the time resolution, T-jump studies are advantageous because they are usually carried out under native-like conditions without high denaturant content.

Photochemical triggering methods, including laser flash photolysis and electron-trannsfer-induced folding/refolding, can also be used to study proteins with heme or other prosthetic/binding groups. One example is the photodissociation of CO-bound cytochrome *c* initiated by laser flash photolysis, which allows refolding to be investigated on the nanosecond time scale (*62*). The same protein has also been investigated by laser-initiated electron-transfer reaction (*63*).

1.3.4 Protein Folding Probed by Mass Spectrometry

Protein folding/unfolding can be studied by MS in various ways. One strategy is to monitor the changes in the positive-ion charge-state distributions of protein ions induced by changes in solvent composition. Wang et al. (*64*) studied the stability of myoglobin and apomyoglobin in different solvent conditions by analyzing charge state

distribution observed in the mass spectra and demonstrated that apomyoglobin has a large partially unfolded portion. Later, Konermann et al. (*65, 66*) investigated kinetic studies of protein-refolding/unfolding reactions using a similar approach with a continuous flow mixing technique and time-resolved measurements. To aid the data analysis on protein conformational isomers coexisting in solution, Dobo et al. (*67*) introduced and evaluated a procedure of detecting and characterizing such isomers by analyzing charge-state distribution of protein ions in ESI mass spectra.

The combination of HDX and MS to probe protein conformational change was first published by Katta and Chait (*68*). Amide hydrogens located in the hydrophobic core of a protein where access to solvent is limited are slowly replaced by deuterons, while those located on the protein surface exchange faster with deuterons. Thus, the extent of the HDX of these amide hydrogens reveals local solvent accessibility, which often changes during folding/unfolding process. Following the initial work, Miranker et al. (*69*) developed a variation of this technique, in which amide proton HDX during the folding of lysozyme was monitored by ESI-MS and NMR. In such an experiment, ESI-MS distinguishes the populations of protein with various masses, thus can be used to detect the existence of multiple pathways. A sub-second time resolution was achieved.

Other chemical labeling approaches have been applied to study folding/unfolding. Their basis is that residues in the folded areas are more protected and would suffer little labeling whereas residues in the unfolded regions are more exposed to solvent and have higher tendency to be labeled. Similar to HDX, the readout of such experiments is the mass shift created by the chemical labeling, and the extent of labeling can also be quantified by MS measurements. One example is the protein folding on the millisecond

time scale probed by combining mutagenesis and pulsed cysteine labeling, followed by global MS analysis (*70*). The kinetics of change in cysteine accessibility during refolding was monitored for ten different mutant variants of barstar. This provides, with high temporal resolution, information on structure formation at 10 different locations during folding.

FPOP also allows protein folding be probed in the millisecond time scale, taking advantage of the fast labeling time of OH· in the presence of the radical scavenger (*30*). FPOP can be combined with rapid mixing to characterize short-lived unfolding intermediates as was reported for myoglobin (*71*) and the folding of cytochrome *c* (*72*). Spatial resolution was achieved by monitoring the ratio of unmodified tryptic peptides. Lately, Pan et al. (*73*) reported the application of pulsed oxidative labeling for deciphering the folding mechanism of a membrane protein, bacteriorhodopsin, from the SDS-denatured state to the membrane and retinal bound state. The time-dependent changes in solvent accessibility were monitored by measuring the extent of oxidative modifications at methionine sidechains. This technique has also been expanded from studying single-chain folding to investigating protein assembly. It was demonstrated that hydroxyl radical labeling tracks time-dependent accessibility changes during folding and assembly of the S100A11 homodimer, reveals the mechanism of folding and dimerization, and identifies short-lived intermediates during this process (*74*). A time limit of several milliseconds can be achieved in such experiments. Given that OH· footprinting is an irreversible labeling process, the folding/unfolding intermediates is "captured" for MS analysis to afford both global protein information and site-specific and

region-specific information from proteolysis followed by LC/MS analysis. Locating sites of modification gives spatial or site-specific resolution on top of the time resolution.

1.4 Dissertation Topics in Part I

1.4.1 Chapter 2: A Dosimetry Experiment for the Quantitation of Hydroxyl Radical Produced by Laser Photolysis in Solution

This is a short yet important chapter describing an experiment measuring the initial concentration of hydroxyl radical produced in solution in an FPOP experiment. Using phenylalanine as a dosimeter, we measured the radical concentration to be 0.42 ± 0.1 mM. Numerical simulation shows that the concentration of hydroxyl radical is decreased to less than 10^{-7} M in 1 μs in the presence of glutamine, the radical scavenger. The fast disappearance of radicals ensures that FPOP footprinting is faster than protein unfolding induced by oxidative labeling. It is also suggested that FPOP is a promising new approach for detecting protein conformational change in the time scale of ms to μs.

1.4.2 Chapter 3: Temperature Jump and FPOP Probe Submillisecond Protein Folding

The focus of this work is to expand FPOP to study protein folding dynamics on the submillisecond time scale. Taking advantage of the microsecond FPOP reaction, we coupled the FPOP probe with a temperature jump that initiates protein folding/unfolding within ns by heating the solution. By varying the delay time between the T jump and the probe, we labeled the protein at different stages of folding during the first 1-2 milliseconds. The readout is the global protein mass spectra showing different oxidative

state distributions as a function of folding extent. We demonstrated this pump-probe method using a model protein, barstar, and showed that it gradually becomes less solvent accessible during the first 1 ms, and reaches a partially folded state after then. This study has been published in *Journal of the American Chemical Society (75).*

1.4.3 Chapter 4: Watching Amino Acid Residues Change during Protein Fast Folding

In this chapter we use the pump-probe method described in Chapter 3, followed by proteolysis and LC-MS/MS to study barstar early folding at the amino-acid residue level and characterize the intermediate state. Our results show that among the four helices and three beta sheets, the first helix shows an increasing trend of protection during the first 1 ms of folding. Although other regions do not show an obvious trend, some individual residues are modified to a different extent in the intermediate state than in the unfold condition. We propose that in the early folding stage, barstar forms a partially solvent-accessible hydrophobic core consisting of several residues that interact with remote residues in the protein sequence. While it is not realistic to correlate directly modification with secondary structure formation, our data are consistent with previous observation that barstar fast folding follow the nucleation-condensation mechanism with the nucleus centered on $helix_1$ formed in the folding intermediate.

1.4.4 Chapter 5: Fast Photochemical Iodination Footprinting Combined with Top-down and Bottom-up Mass Spectrometric Analysis Reveals Protein Conformation

In this chapter, we report a new FPOP-based protein footprinting method utilizing the iodine radical as the reagent. The radicals are generated by photolysis of iodobenzoic acid at 248 nm. We find that iodine radical reacts principally with tyrosine and histidine residues; thus, it is a site-specific labeling reagent. To prove that this method is capable of differentiating protein conformational change, we studied myoglobin and apomyoglobin as a pair and observed both global and residue-level difference in term of iodination extent. The modified residues are mapped into the structure of the protein, and the differences reveal changes of solvent accessibility upon heme binding. When applying this footprinting method to carbonic anhydrase and apo-carbonic anhydrase, we found no significant difference of modification extent as a function of protein state. This is consistent with previous findings that the structures of the apo and holo protein states are similar to each other (*76*).

1.5 Bacterial Peptidoglycan and the Analysis

1.5.1 Structure and Function

Peptidoglycan, also known as murein (*77*), is a polymer consisting of polysaccharides and amino acids that forms layers outside of the bacterial membrane, and is the main component of the gram-positive bacteria cell wall. The linear glycan chains are comprised of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) (see Figure 1.2). The average chain length varies between 10 and 65 disaccharide units (*78*). The structure of the glycan chain is largely conserved in all bacteria although there are minor variations such as *O*-acetylation. Attached to the *N*-acetylmuramic acid is a short peptide chain which can form cross-links to another peptide

chain of another strand. The peptide moiety contains *L* and *D* amino acids, usually with *L*-alanine bound to the muramic acid, followed by *D*-glutamic acid, which is connected by its γ-carboxyl group to an *L*-diamino acid with *D*-alanine attached. Sometimes there is an additional alanine at the C terminus. The peptide linkage is formed between the amino group of the *L*-diamino acid and the C terminal *D*-alanine of another peptide stem. In the formation of peptidoglycan, two types of enzymatic activities are involved: transglycosylase catalyzes the formation of glycan chains and transpeptidase cross-links the peptide stems (*79*, *80*).

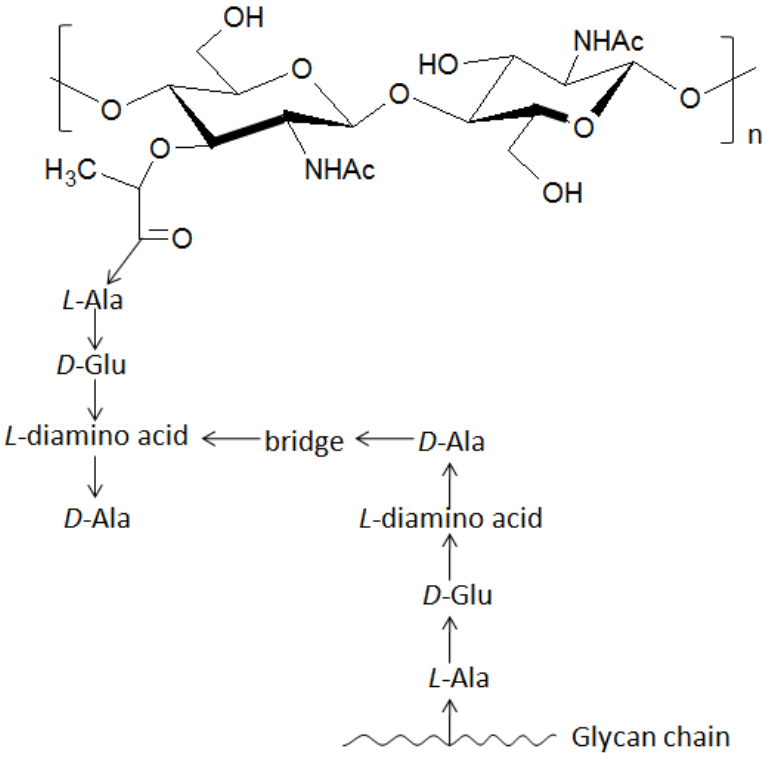


Figure 1.2: A repeating unit of a peptidoglycan structure.

The three-dimensional structure of peptidoglycan is an important, yet unsolved problem. There are controversial models proposed for the arrangement of peptidoglycan. One model is that the glycan chains are parallel to the cell surface (*81*), whereas the other model proposes that the chains grow perpendicularly to the surface (*82*). The former model, however, is generally more preferred.

The two main functions of peptidoglycan in bacteria are to maintain the shape of the cell and to counteract the osmotic pressure of the bacterial protoplast. Hence, the peptidoglycan is a target for most of the antibiotics that inhibit bacterial cell-wall synthesis.

1.5.2 Antibiotics and their relation to peptidoglycan

Antibiotics, also known as antibacterials, are drugs that are used to treat infections caused by bacteria. There are several types of antibiotics such as beta-lactam antibiotics; examples are glycopeptides, aminoglycosides, macrolides, fluoroquinolones, and tetracyclines.

Glycopeptides are a class of antibiotics composed of glycosylated cyclic or polycyclic nonribosomal peptides. They slow the formation of a bacterial cell wall by inhibiting the synthesis of peptidoglycan. Some common glycopeptides antibiotics include vancomycin (Figure 1.3), teicoplanin, telavancin, bleomycin, ramoplanin and decaplanin. The mechanism of inhibition of cell-wall biosynthesis involves binding to amino acids of the peptidylglycam, specifically the acyl-*D*-alanyl-*D*-alanine, to prevent the addition of new units (Figure 1.3). As a result of antibiotic use, resistance towards antibiotics becomes more common and presents a clinical challenge. The primary cause

of resistance is the genetic mutation in bacteria. For example, bacteria that are resistant to glycopeptides antibiotics (e.g., vancomycin) have their peptidoglycan altered from *D*-Ala-*D*-Ala to *D*-Ala-*D*-Lac, which has a 1,000 fold lower binding constant to vancomycin (*83, 84*).

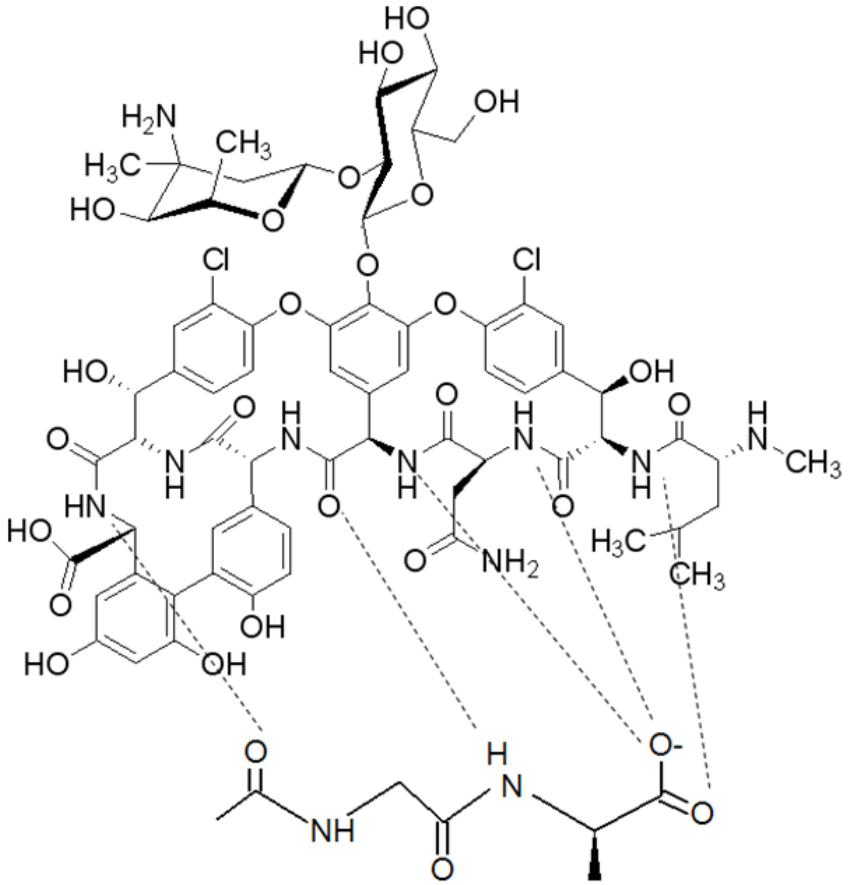


Figure 1.3. Binding of vancomycin to the *D*-Ala-*D*-Ala terminus of a peptidoglycan stem.

As bacteria resistance emerges, the need for the development of new antibiotics arises. Generally, there are two strategies to look for new antibiotics. One is to screen a large chemical library in which a reference drug as a starting point and many random structural variations based on this drug are present. Another approach is to design the drugs based on the understanding of the structure and biochemistry of the targets. In this context, understanding the detailed structure of peptidoglycan and the modes of action of glycopeptides is crucial to the drug development.

1.6 Bacterial Peptidoglycan Analysis

1.6.1 Solid-State NMR Applied to Bacterial Peptidoglycan

Solid-state NMR is an ideal method to characterize bacterial peptidoglycan because it is compatible with insoluble and largely heterogeneous samples. The idea was first introduced in 1983 by Schaefer and coworkers whereby cross-linking in the peptidoglycan was quantified using cross-polarization magic-angle spinning (CPMAS) technique (*85*). This paper also showed that the modes of action of antibiotics such as penicillin could be investigated using solid-state NMR. The paper did not receive much attention, however, until the late 1980s when antibiotic resistant was recognized as a serious problem.

Rotational-echo double-resonance (REDOR), introduced by Schaefer et al. in 1989 (*86*), is capable of providing distance measurement between isotopes incorporated into the peptidoglycan and bound antibiotics. Using REDOR, Schaefer et al. (*87, 88*) demonstrated in several publications that vancomycin inhibits transglycosylation in *Staphylococcus aureus*. Other vancomycin derivatives can also be characterized by this

approach (*89-91*). REDOR, typically performed on solid samples undergoing magic-angle spinning (MAS), directly measures the distance-dependent heteronuclear dipolar coupling between spin pairs. The experiment is comprised of two parts. In the first part, the dephasing channel is off, and MAS spatially averages chemical shift and dipolar anisotropic interactions to produce an echo of full intensity. In the second part, a train of π pulses are applied to interfere with the dipolar coupling. This results in an observed dephasing that is related to distance by the dipolar-coupling strength.

In a recent application relevant to this thesis research, Patti et al. (*92*) investigated the structure of *Enterococcus faecium* peptidoglycan and the modes of action of vancomycin and oritavancin in the organism by using REDOR (*92, 93*). One goal of this dissertation is to develop mass spectrometry as a complementary method to solid-state NMR in studying *Enterococcus faecium* peptidoglycan.

1.6.2 Mass Spectrometry-based Methods

As indicated above, peptidoglycan is highly heterogeneous and insoluble, thus presenting challenges for mass spectrometric analysis. This problem can be solved in part by digesting the peptidoglycan to disaccharide units with the peptides attached; these are called muropeptides. To accomplish this, muramidase such as lysozyme and mutanolysis can be used to give digests that are suitable for bottom-up analysis. Lysozyme has its lytic effect upon susceptible bacteria through hydrolysis of glycosidic bonds (*94-96*), specifically the β-1,4 linkages between *N*-acetylmuramic acid and *N*-acetylglucosamine. Resistance to such lytic action arises owing either to the presence of other cell walls and cell membrane polymers (*97*) or the modification of peptidoglycan

structures, such as *O*-acetylation on the glycan (*98-102*). Thus, mutanolysin, also a muramidase cleaving the *N*-acetylmuramyl-β(1,4)-*N*-acetylglucosamine linkage but less perturbed by the modification on the glycan, is often used as a complementary enzyme to lysozyme (*103, 104*).

The digested murein samples were usually separated by HPLC before mass spectrometric analysis (*105-109*). Previously, fast atom bombardment mass spectrometry (FAB-MS) and collisionally activated dissociation tandem mass spectrometry (FAB-CAD-MS/MS) were used to characterize the fragments produced by digestion (*110, 111*). Unfortunately, this approach usually fails to provide information for trimers and higher oligomers. With the development of electrospray ionization (*112*) and matrix-assisted laser desorption ionization (MALDI) (*113*) , peptidoglycan analysis by MS became more feasible and complete (*114, 115*). Given the heterogeneity and complexity of the peptidoglycan structure, HPLC fraction collection and off-line MS analysis, however, has the tendency of losing information on the minor components. There have been several reports of using liquid chromatography-mass spectrometry (LC-MS) to separate and identify peptidoglycan fragments, but they did not address detailed muropeptide structural variations presumably owing to the lack of sensitivity in detection and mass accuracy for composition determination (*116*).

Advances in mass-spectrometry instrumentation and on-line separation have provided instruments with high mass resolving power and sensitivity. Examples of MS instrumentation include Fourier-transform ion cyclotron resonance (FTICR) mass spectrometers, Orbitrap mass spectrometers and some quadrupole time-of-flight (Q-TOF) mass spectrometers. In this dissertation, we take advantage of these relatively new

capabilities to develop an LC-MS based method to identify and semi-quantify individual muropeptide species including minor structural variations. This cannot be accomplished readily by solid state NMR, but takes advantage of accurate mass measurement, MS/MS, and data-processing strategies including the ability to extract ion chromatograms.

1.7 Dissertation Topics in Part II

1.7.1 Chapter 6: LC/MS Method Development for Characterizing Peptidoglycan Structures

In Chapter 6 and as a collaboration with G. Patti and J. Schaefer, we describe the LC/MS-based method development for peptidoglycan characterization, starting from enzymatic digestion to data analysis, using *Enterococcus faecium* as a model. We identified ~50 muropeptides including 16 basic ones and their variations. Their formulas were confirmed by accurate mass measurement, and structural assignments were made by MS/MS. We approximated the relative amount of each by assuming equal response in ionization and calculated the fractions of monomers, dimers and trimers and the cross-linking value. These findings have biological implications with regard to peptidoglycan growth, maturation, and antibiotic resistance. We propose that this LC/MS-based method is widely applicable to bacterial peptidoglycan characterization and that it may have reasonable throughput. Part of this chapter was published in the *Journal of American Society for Mass Spectrometry* (*117*).

1.7.2 Chapter 7: Time-Dependent Isotopic Labeling and Quantitative Liquid-Chromatography/Mass Spectrometry Reveals Bacterial Cell-Wall Architecture

In Chapter 7 and in collaboration with G. Patti, we expanded the steady-state analysis of bacterial peptidoglycan to the study of the growth dynamics by applying time-dependent isotopic labeling while the bacteria are growing. This allows us to track the fate of the peptidoglycan precursor during maturation by comparing the isotopic enrichment of muropeptide along the growth curve. We apply the approach to *Enterococcus faecium* and find the sequence of cell-wall maturation. The results also reveal some information on the cell-wall organization and tertiary structure. Part of this chapter was published in *Analytical Chemistry* (*118*).

1.7.3 Chapter 8: MS Investigation of Peptidoglycan in *Fem*-deletion Mutants of Methicillin-resistant *Staphylococcus aureus*

We applied, as described in this chapter, the bottom-up MS method to investigate methicillin-resistant *Staphylococcus aureus (S. aureus)* and its *fem* (for factor essential for methicillin resistance)-deletion mutants. Fine structures of muropeptides are characterized and compared among these three strains. The average bridge length, a characteristic feature, is calculated to be 4.3, 0.9 and 2.2 for parent, FemA and FemB, respectively, by quantifying all the relevant species. The muropeptide with three glycines as the bridge is detected in FemA. A time-dependent, isotopic-labeling experiment conducted on FemA reveals the fate of peptidoglycan and also has implications on its tertiary structure. The MS-based methods again serve as a complementary tool to NMR, and prove to be applicable to a variety of bacteria.

1.8 Conclusion

The method developments described in this thesis have implications in two important areas of biochemistry and biophysics. The FPOP-based methods add in valuable tools for structural biology. The iodination method is widely applicable in detecting protein conformational change and mapping protein-ligand or protein-protein interface. The T jump and FPOP combination brings a wealth of structural information about protein folding within reach, and has considerable flexibility. These methods could permit the elucidation of protein structure and dynamics for systems where NMR or X-ray is not applicable. The bottom-up MS methods for peptidoglycan analysis contribute to the understanding of bacterial cell wall structure and biosynthesis. They are useful not only in determining the peptidoglycan structure of a variety of gram-positive bacteria but also in investigating the antibiotic modes of action.

1.8 References

1. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis, *Nature 181*, 662-666.
2. Scapin, G. (2006) Structural biology and drug discovery, *Current pharmaceutical design 12*, 2087-2097.
3. Wuthrich, K. (1990) Protein structure determination in solution by NMR spectroscopy, *The Journal of biological chemistry 265*, 22059-22062.

4. Dyson, H. J., and Wright, P. E. (1996) Insights into protein folding from NMR, *Annual review of physical chemistry 47*, 369-395.

5. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1997) Characterization of the backbone dynamics of folded and denatured states of an SH3 domain, *Biochemistry 36*, 2390-2402.

6. Kay, L. E. (1998) Protein dynamics from NMR, *Nature structural biology 5 Suppl*, 513-517.

7. Clore, G. M., and Gronenborn, A. M. (1998) Determining the structures of large proteins and protein complexes by NMR, *Trends in biotechnology 16*, 22-34.

8. Fasman, G. D. (1996) *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York.

9. Greenfield, N. J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, *Analytical biochemistry 235*, 1-10.

10. Sharma, A., and Schulman, S. G. (1999) *Introduction to Fluorescence Spectroscopy*, Wiley-interscience, New York.

11. Stryer, L., and Haugland, R. P. (1967) Energy transfer: a spectroscopic ruler, *Proceedings of the National Academy of Sciences of the United States of America 58*, 719-726.

12. Schuler, B., and Eaton, W. A. (2008) Protein folding studied by single-molecule FRET, *Current opinion in structural biology 18*, 16-26.

13. Pascher, T., Chesick, J. P., Winkler, J. R., and Gray, H. B. (1996) Protein folding triggered by electron transfer, *Science (New York, N.Y 271*, 1558-1560.

14. Griffiths, J. (2008) A brief history of mass spectrometry, *Analytical chemistry 80,* 5678-5683.

15. Kaltashov, I. A., and Eyles, S. J. (2005) *Mass Spectrometry in Biophysics*, John Wiley and Sons, Inc.

16. Englander, S. W. (2006) Hydrogen exchange and mass spectrometry: A historical perspective, *Journal of the American Society for Mass Spectrometry 17,* 1481-1489.

17. Rand, K. D., Adams, C. M., Zubarev, R. A., and Jorgensen, T. J. (2008) Electron capture dissociation proceeds with a low degree of intramolecular migration of peptide amide hydrogens, *Journal of the American Chemical Society 130,* 1341-1349.

18. Pan, J., Han, J., Borchers, C. H., and Konermann, L. (2009) Hydrogen/deuterium exchange mass spectrometry with top-down electron capture dissociation for characterizing structural transitions of a 17 kDa protein, *Journal of the American Chemical Society 131,* 12801-12808.

19. Rand, K. D., Zehl, M., Jensen, O. N., and Jorgensen, T. J. (2009) Protein hydrogen exchange measured at single-residue resolution by electron transfer dissociation mass spectrometry, *Analytical chemistry 81*, 5577-5584.

20. Heyduk, T., Baichoo, N., and Heyduk, E. (2001) Hydroxyl radical footprinting of proteins using metal ion complexes, *Metal ions in biological systems 38,* 255-287.

21. Xu, G., and Chance, M. R. (2004) Radiolytic modification of acidic amino acid residues in peptides: probes for examining protein-protein interactions, *Analytical chemistry 76,* 1213-1221.

22. Xu, G., and Chance, M. R. (2005) Radiolytic modification and reactivity of amino acid residues serving as structural probes for protein footprinting, *Analytical chemistry 77*, 4549-4555.

23. Xu, G., and Chance, M. R. (2005) Radiolytic modification of sulfur-containing amino acid residues in model peptides: fundamental studies for protein footprinting, *Analytical chemistry 77*, 2437-2449.

24. Xu, G., and Chance, M. R. (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics, *Chemical reviews 107*, 3514-3543.

25. Buxton, G. V., Greenstock, C. L., Helman, W. P., and Ross, A. B. (1988) Critical Review of Rate Constants for Reactions of Hydrated Electrons, Hydrogen Atoms and Hydroxyl Radicals (*OH/*O-) in Aqueous Solution., *J. Phys. Chem. Ref. Data 17*.

26. Davies, M. J., and Dean, R. T. (1997) Radical-Mediated Protein Oxidation: From Chemistry to Medicine, Oxford University Press.

27. Tullius, T. D., and Greenbaum, J. A. (2005) Mapping nucleic acid structure by hydroxyl radical cleavage, *Current opinion in chemical biology 9*, 127-134.

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

29. Hambly, D. M., and Gross, M. L. (2007) Laser Flash Photochemical Oxidation to Locate Heme Binding and Conformational Change in Myoglobin, *Int. J. Mass Spectrom. 259*, 124-129.

30. Gau, B. C., Sharp, J. S., Rempel, D. L., and Gross, M. L. (2009) Fast photochemical oxidation of protein footprints faster than protein unfolding, *Analytical chemistry 81*, 6563-6571.

31. Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900-904.

32. Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annual review of biochemistry 75*, 333-366.

33. Dobson, C. M., Sali, A., and Karplus, M. (1998) Protein Folding: A Perspective from Theory and Experiment, *Angewandte Chemie International Edition 37*, 868-893.

34. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nature structural biology 4*, 10-19.

35. Baldwin, R. L. (1994) Protein folding. Matching speed and stability, *Nature 369*, 183-184.

36. Fersht, A. R. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism, *Proceedings of the National Academy of Sciences of the United States of America 97*, 1525-1529.

37. Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2001) Three key residues form a critical contact network in a protein folding transition state, *Nature 409*, 641-645.

38. Roder, H., and Colon, W. (1997) Kinetic role of early intermediates in protein folding, *Current opinion in structural biology 7*, 15-28.

39. Vendruscolo, M., Paci, E., Karplus, M., and Dobson, C. M. (2003) Structures and relative free energies of partially folded states of proteins, *Proceedings of the National Academy of Sciences of the United States of America 100*, 14817-14821.

40. Dobson, C. M., Evans, P. A., and Williamson, K. L. (1984) Proton NMR studies of denatured lysozyme, *FEBS letters 168*, 331-334.

41. Huang, G. S., and Oas, T. G. (1995) Submillisecond folding of monomeric lambda repressor, *Proceedings of the National Academy of Sciences of the United States of America 92*, 6878-6882.

42. McGee, W. A., and Parkhurst, L. J. (1990) A combined nuclear magnetic resonance and absorbance stopped-flow apparatus for biochemical studies, *Analytical biochemistry 189*, 267-273.

43. Frieden, C., Hoeltzli, S. D., and Ropson, I. J. (1993) NMR and protein folding: equilibrium and stopped-flow studies, *Protein Sci 2*, 2007-2014.

44. Schanda, P., Forge, V., and Brutscher, B. (2007) Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy, *Proceedings of the National Academy of Sciences of the United States of America 104*, 11257-11262.

45. Hu, K. N., and Tycko, R. (2010) What can solid state NMR contribute to our understanding of protein folding?, *Biophysical chemistry 151*, 10-21.

46. Fersht, A. R., Leatherbarrow, R. J., and Wells, T. N. (1987) Structure-activity relationships in engineered proteins: analysis of use of binding energy by linear free energy relationships, *Biochemistry 26*, 6030-6038.

47. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, *Journal of molecular biology 224*, 771-782.

48. Nolting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) The folding pathway of a protein at high resolution from microseconds to seconds, *Proceedings of the National Academy of Sciences of the United States of America 94*, 826-830.

49. Tsong, T. Y. (1976) Ferricytochrome c chain folding measured by the energy transfer of tryptophan 59 to the heme group, *Biochemistry 15*, 5467-5473.

50. Luchins, J., and Beychok, S. (1978) Far-ultraviolet stopped-flow circular dichroism, *Science (New York, N.Y 199*, 425-426.

51. Kiefhaber, T., Schmid, F. X., Willaert, K., Engelborghs, Y., and Chaffotte, A. (1992) Structure of a rapidly formed intermediate in ribonuclease T1 folding, *Protein Sci 1*, 1162-1172.

52. Regenfuss, P., and Clegg, R. M. (1987) Diffusion-controlled association of a dye, 1-anilinonaphthalene-8-sulfonic acid, to a protein, bovine serum albumin, using a fast-flow microsecond mixer and stopped-flow, *Biophysical chemistry 26*, 83-89.

53. Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) Folding of cytochrome c initiated by submillisecond mixing, *Nature structural biology 4*, 44-50.

54. Sauder, J. M., MacKenzie, N. E., and Roder, H. (1996) Kinetic mechanism of folding and unfolding of Rhodobacter capsulatus cytochrome c2, *Biochemistry 35*, 16852-16862.

55. Park, S. H., O'Neil, K. T., and Roder, H. (1997) An early intermediate in the folding reaction of the B1 domain of protein G contains a native-like core, *Biochemistry 36*, 14277-14283.

56. Nolting, B., Golbik, R., and Fersht, A. R. (1995) Submillisecond events in protein folding, *Proceedings of the National Academy of Sciences of the United States of America 92*, 10668-10672.

57. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) Direct observation of fast protein folding: the initial collapse of apomyoglobin, *Proceedings of the National Academy of Sciences of the United States of America 93*, 5759-5764.

58. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1996) Fast events in protein folding: helix melting and formation in a small peptide, *Biochemistry 35*, 691-697.

59. Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1997) Fast events in protein folding: relaxation dynamics of secondary and tertiary structure in native apomyoglobin, *Proceedings of the National Academy of Sciences of the United States of America 94*, 3709-3713.

60. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) Observation of distinct nanosecond and microsecond protein folding events, *Nature structural biology 3*, 923-926.

61. Munoz, V., Thompson, P. A., Hofrichter, J., and Eaton, W. A. (1997) Folding dynamics and mechanism of beta-hairpin formation, *Nature 390*, 196-199.

62. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) Fast events in protein folding initiated

by nanosecond laser photolysis, *Proceedings of the National Academy of Sciences of the United States of America 90*, 11860-11864.

63. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R., and Gray, H. B. (1996) Cytochrome c folding triggered by electron transfer, *Chemistry & biology 3*, 491-497.

64. Wang, F., and Tang, X. (1996) Conformational heterogeneity of stability of apomyoglobin studied by hydrogen/deuterium exchange and electrospray ionization mass spectrometry, *Biochemistry 35*, 4069-4078.

65. Konermann, L., Collings, B. A., and Douglas, D. J. (1997) Cytochrome c folding kinetics studied by time-resolved electrospray ionization mass spectrometry, *Biochemistry 36*, 5554-5559.

66. Sogbein, O. O., Simmons, D. A., and Konermann, L. (2000) Effects of pH on the kinetic reaction mechanism of myoglobin unfolding studied by time-resolved electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry 11*, 312-319.

67. Dobo, A., and Kaltashov, I. A. (2001) Detection of multiple protein conformational ensembles in solution via deconvolution of charge-state distributions in ESI MS, *Analytical chemistry 73*, 4763-4773.

68. Katta, V., and Chait, B. T. (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry, *Rapid Commun Mass Spectrom 5*, 214-217.

69. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., and Dobson, C. M. (1993) Detection of transient protein folding populations by mass spectrometry, *Science (New York, N.Y 262*, 896-900.

70. Jha, S. K., and Udgaonkar, J. B. (2007) Exploring the cooperativity of the fast folding reaction of a small protein using pulsed thiol labeling and mass spectrometry, *The Journal of biological chemistry 282*, 37479-37491.

71. Stocks, B. B., and Konermann, L. (2009) Structural characterization of short-lived protein unfolding intermediates by laser-induced oxidative labeling and mass spectrometry, *Analytical chemistry 81*, 20-27.

72. Stocks, B. B., and Konermann, L. (2010) Time-dependent changes in side-chain solvent accessibility during cytochrome c folding probed by pulsed oxidative labeling and mass spectrometry, *Journal of molecular biology 398*, 362-373.

73. Pan, Y. B. L., and Konermann, L. (2011) Kinetic Folding Mechanism of an Integral Membrane Protein Examined by Pulsed Oxidative Labeling and Mass Spectrometry *Journal of molecular biology 410*, 146-158.

74. Stocks, B. B., Rezvanpour, A., Shaw, G. S., and Konermann, L. (2011) Temporal Development of Protein Structure during S100A11 Folding and Dimerization Probed by Oxidative Labeling and Mass Spectrometry *Journal of molecular biology 409*, 669-679.

75. Chen, J., Rempel, D. L., and Gross, M. L. (2010) Temperature jump and fast photochemical oxidation probe submillisecond protein folding, *Journal of the American Chemical Society 132*, 15502-15504.

76. Hakansson, K., Carlsson, M., Svensson, L. A., and Liljas, A. (1992) Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes, *J Mol Biol 227*, 1192-1204.

77. Weidel, W., and Pelzer, H. (1964) Bagshaped Macromolecules--a New Outlook on Bacterial Cell Walls, *Advances in enzymology and related areas of molecular biology 26*, 193-232.

78. Ghuysen, J. M. (1968) Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism, *Bacteriological reviews 32*, 425-464.

79. Tipper, D. J., and Strominger, J. L. (1968) Biosynthesis of the peptidoglycan of bacterial cell walls. XII. Inhibition of cross-linking by penicillins and cephalosporins: studies in Staphylococcus aureus in vivo, *The Journal of biological chemistry 243*, 3169-3179.

80. Schleifer, K. H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications, *Bacteriological reviews 36*, 407-477.

81. Vollmer, W., and Holtje, J. V. (2004) The architecture of the murein (peptidoglycan) in gram-negative bacteria: vertical scaffold or horizontal layer(s)?, *Journal of bacteriology 186*, 5978-5987.

82. Dmitriev, B. A., Toukach, F. V., Holst, O., Rietschel, E. T., and Ehlers, S. (2004) Tertiary structure of Staphylococcus aureus cell wall murein, *Journal of bacteriology 186*, 7141-7148.

83. Uttley, A. H., Collins, C. H., Naidoo, J., and George, R. C. (1988) Vancomycin-resistant enterococci, *Lancet 1*, 57-58.

84. Bugg, T. D., Wright, G. D., Dutka-Malen, S., Arthur, M., Courvalin, P., and Walsh, C. T. (1991) Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA, *Biochemistry 30*, 10408-10415.

85. Jacob, G. S., Schaefer, J., and Wilson, G. E., Jr. (1983) Direct measurement of peptidoglycan cross-linking in bacteria by 15N nuclear magnetic resonance, *The Journal of biological chemistry 258*, 10824-10826.

86. Guillion, T., and Schaefer, J. (1989) Detection of Weak Heteronuclear Dipolar Coupling by Rotational-Echo Double-Resonance Nuclear Magnetic Resonance, *Adv. Magn. Reson. 13*, 57-83.

87. Kim, S. J., Cegelski, L., Studelska, D. R., O'Connor, R. D., Mehta, A. K., and Schaefer, J. (2002) Rotational-echo double resonance characterization of vancomycin binding sites in Staphylococcus aureus, *Biochemistry 41*, 6967-6977.

88. Cegelski, L., Kim, S. J., Hing, A. W., Studelska, D. R., O'Connor, R. D., Mehta, A. K., and Schaefer, J. (2002) Rotational-echo double resonance characterization of the effects of vancomycin on cell wall synthesis in Staphylococcus aureus, *Biochemistry 41*, 13053-13058.

89. Cegelski, L., Steuber, D., Mehta, A. K., Kulp, D. W., Axelsen, P. H., and Schaefer, J. (2006) Conformational and quantitative characterization of oritavancin-peptidoglycan complexes in whole cells of Staphylococcus aureus by in vivo 13C and 15N labeling, *Journal of molecular biology 357*, 1253-1262.

90. Kim, S. J., Cegelski, L., Stueber, D., Singh, M., Dietrich, E., Tanaka, K. S., Parr, T. R., Jr., Far, A. R., and Schaefer, J. (2008) Oritavancin exhibits dual mode of

action to inhibit cell-wall biosynthesis in Staphylococcus aureus, *Journal of molecular biology 377*, 281-293.

91. Kim, S. J., Matsuoka, S., Patti, G. J., and Schaefer, J. (2008) Vancomycin derivative with damaged D-Ala-D-Ala binding cleft binds to cross-linked peptidoglycan in the cell wall of Staphylococcus aureus, *Biochemistry 47*, 3822-3831.

92. Patti, G. J., Kim, S. J., and Schaefer, J. (2008) Characterization of the peptidoglycan of vancomycin-susceptible Enterococcus faecium, *Biochemistry 47*, 8378-8385.

93. Patti, G. J., Kim, S. J., Yu, T. Y., Dietrich, E., Tanaka, K. S., Parr, T. R., Jr., Far, A. R., and Schaefer, J. (2009) Vancomycin and oritavancin have different modes of action in Enterococcus faecium, *Journal of molecular biology 392*, 1178-1191.

94. Grula, E. A., and Hartsell, S. E. (1954) Lysozyme and morphological alterations induced in Micrococcus lysodeikticus, *Journal of bacteriology 68*, 171-177.

95. Salton, M. R. (1952) Cell wall of Micrococcus lysodeikticus as the substrate of lysozyme, *Nature 170*, 746-747.

96. Goodman, H., Pollock, J. J., Iacono, V. J., Wong, W., and Shockman, G. D. (1981) Peptidoglycan loss during hen egg white lysozyme-inorganic salt lysis of Streptococcus mutans, *Journal of bacteriology 146*, 755-763.

97. Horne, D., and Tomasz, A. (1980) Lethal effect of a heterologous murein hydrolase on penicillin-treated Streptococcus sanguis, *Antimicrobial agents and chemotherapy 17*, 235-246.

98. Brumfitt, W. (1959) The mechanism of development of resistance to lysozyme by some gram-positive bacteria and its results, *British journal of experimental pathology 40*, 441-451.

99. Brumfitt, W., Wardlaw, A. C., and Park, J. T. (1958) Development of lysozyme-resistance in Micrococcus lysodiekticus and its association with an increased O-acetyl content of the cell wall, *Nature 181*, 1783-1784.

100. Logardt, I. M., and Neujahr, H. Y. (1975) Lysis of modified walls from Lactobacillus fermentum, *Journal of bacteriology 124*, 73-77.

101. Work, E. (1967) Factors affecting the susceptibility of bacterial cell walls to the action of lysozyme, *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character 167*, 446-447.

102. Holden, J. T., Grant, W., and Degroot, J. (1979) Lysozyme sensitivity of pantothenate-deficient Lactobacillus plantarum grown with exogenous fatty acids, *Biochimica et biophysica acta 574*, 173-176.

103. Lammler, C., and Frede, C. (1988) Mutanolysin-induced lysis of actinomyces pyogenes determined by aggregometry, *Zentralblatt fur Bakteriologie, Mikrobiologie, und Hygiene 269*, 447-453.

104. Fliss, I., Emond, E., Simard, R. E., and Pandian, S. (1991) A rapid and efficient method of lysis of Listeria and other gram-positive bacteria using mutanolysin, *BioTechniques 11*, 453, 456-457.

105. Quintela, J. C., Pittenauer, E., Allmaier, G., Aran, V., and de Pedro, M. A. (1995) Structure of peptidoglycan from Thermus thermophilus HB8, *Journal of bacteriology 177*, 4947-4962.

106. Quintela, J. C., Garcia-del Portillo, F., Pittenauer, E., Allmaier, G., and de Pedro, M. A. (1999) Peptidoglycan fine structure of the radiotolerant bacterium Deinococcus radiodurans Sark, *Journal of bacteriology 181*, 334-337.

107. Singh, P. D., and Johnson, J. H. (1984) Muraceins--muramyl peptides produced by Nocardia orientalis as angiotensin-converting enzyme inhibitors. II. Isolation and structure determination, *The Journal of antibiotics 37*, 336-343.

108. Pittenauer, E., Schmid, E. R., Allmaier, G., Pfanzagl, B., Loffelhardt, W., Fernandez, C. Q., de Pedro, M. A., and Stanek, W. (1993) Structural characterization of the cyanelle peptidoglycan of Cyanophora paradoxa by 252Cf plasma desorption mass spectrometry and fast atom bombardment/tandem mass spectrometry, *Biological mass spectrometry 22*, 524-536.

109. Pfanzagl, B., Zenker, A., Pittenauer, E., Allmaier, G., Martinez-Torrecuadrada, J., Schmid, E. R., De Pedro, M. A., and Loffelhardt, W. (1996) Primary structure of cyanelle peptidoglycan of Cyanophora paradoxa: a prokaryotic cell wall as part of an organelle envelope, *Journal of bacteriology 178*, 332-339.

110. Martin, S. A., Rosenthal, R. S., and Biemann, K. (1987) Fast atom bombardment mass spectrometry and tandem mass spectrometry of biologically active peptidoglycan monomers from Neisseria gonorrhoeae, *The Journal of biological chemistry 262*, 7514-7522.

111. de Jonge, B. L., Chang, Y. S., Gage, D., and Tomasz, A. (1992) Peptidoglycan composition of a highly methicillin-resistant Staphylococcus aureus strain. The role of penicillin binding protein 2A, *The Journal of biological chemistry 267*, 11248-11254.

112. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science (New York, N.Y 246*, 64-71.

113. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Analytical chemistry 60*, 2299-2301.

114. Xu, N., Huang, Z. H., de Jonge, B. L., and Gage, D. A. (1997) Structural characterization of peptidoglycan muropeptides by matrix-assisted laser desorption ionization mass spectrometry and postsource decay analysis, *Analytical biochemistry 248*, 7-14.

115. Bacher, G., Korner, R., Atrih, A., Foster, S. J., Roepstorff, P., and Allmaier, G. (2001) Negative and positive ion matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and positive ion nano-electrospray ionization quadrupole ion trap mass spectrometry of peptidoglycan fragments isolated from various Bacillus species, *J Mass Spectrom 36*, 124-139.

116. Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., van Heijenoort, J., and Gutmann, L. (1996) Peptidoglycan structure of Enterococcus faecium expressing vancomycin resistance of the VanB type, *The Biochemical journal 313 (Pt 3)*, 711-715.

117. Patti, G. J., Chen, J., Schaefer, J., and Gross, M. L. (2008) Characterization of structural variations in the peptidoglycan of vancomycin-susceptible Enterococcus faecium: understanding glycopeptide-antibiotic binding sites using mass spectrometry, *Journal of the American Society for Mass Spectrometry 19*, 1467-1475.

118. Patti, G. J., Chen, J., and Gross, M. L. (2009) Method revealing bacterial cell-wall architecture by time-dependent isotope labeling and quantitative liquid chromatography/mass spectrometry, *Analytical chemistry 81*, 2437-2445.

2. A Dosimetry Experiment for the Quantitation of Hydroxyl Radical Produced by Laser Photolysis in Solution

2.1 Abstract

Like other chemical footprinting methods, fast photochemical oxidation of proteins (FPOP) should be done so that only the native conformation is modified. The current FPOP design utilizes a radical scavenger that limits the hydroxyl-radical exposure time to approximately 1 μs, as determined from kinetics calculations that assume an initial concentration of OH radicals produced upon photolysis. If this time is correct, it is sufficiently short that protein unfolding induced by the oxidative modifications is avoided. In this chapter, we report an experimental determination of the initial concentration of ·OH by using phenylalanine as a dosimeter molecule for determining the initial concentration of hydroxyl radical upon laser photolysis is 0.42 ± 0.1 mM for a 14 ns laser pulse of 45 mJ. Numerical simulations using this initial concentration reveal that in less than 1 μs, the concentration of ·OH decreases to less than 10^{-7} M in the presence of the scavenger. The fast disappearance of radicals in this well-controlled footprinting technique suggests that misleading oxidative modifications do not occur when using the FPOP approach.

2.2 Introduction

Protein footprinting monitors protein conformation by selectively labeling or cleaving residues. Hydroxyl radical oxidation, a class of footprinting methods, probes solvent accessibility by modifying amino acid side chains. The advantages of hydroxyl radical footprinting are first, it is a non-specific labeling method that modifies a significant portion of amino acid residues, thus providing higher coverage compared to residue-specific labeling methods; and second, it imparts to the protein irreversible covalent modifications that are insensitive to demanding sample handling after labeling. As implied in the introduction of this thesis, there are several approaches of hydroxyl radical footprinting including electron-pulse radiolysis, synchrotron radiolysis, use of the Fenton reaction, and laser photolysis (*1*). These methods have been developed and used to map protein surface, as well as footprint DNA:protein and protein:ligand interactions (*2-5*).

Recently Hambly and Gross (*6*) developed the method of fast photochemical oxidation of proteins (FPOP) that uses a pulsed laser to photolyze hydrogen peroxide to hydroxyl radicals; a similar approach was introduced by Aye and coworkers (*7*). It was demonstrated in the Hambly experiment that in the absence of a scavenger glutamine, the protein is heavily and perhaps excessively oxidized. In the presence of 20 mM glutamine, however, the radicals are almost gone within 1 µs, calculated by using the initial concentration of radical (estimated to be 1mM) and the known rate constant for reaction of glutamine with the radical. This timescale is faster than the second step in protein unfolding, a step that re-admits water and hydroxyl radicals to oxidize the protein

in an uncontrolled fashion (*8*). Thus, we have assumed that it is principally the native protein that is probed by FPOP.

To validate experimentally the claim by Hambly that FPOP oxidizes protein faster than its unfolding, Gau et al. (*9*) applied FPOP to three oxidation-sensitive proteins and found that the distribution of oxidation modification states is Poisson in the presence of a radical scavenger, consistent with a model that holds that only a single conformation of the is oxidatively modified. In the absence of the scavenger, the model no longer pertains. The result verifies that FPOP does occur on a timescale enough to sample a single conformational state of the protein. This approach does not provide, however, a lifetime for the radicals.

The high speed of oxidative modifications can placed on firmer ground by measuring the initial concentration of hydroxyl radical upon laser photolysis instead of estimating the number based on the quantum yield. In this chapter, we use a chemical dosimeter to measure directly the amount of radicals produced in the experiment; this concentration is an input for any kinetic analysis of the radical decay.

Radiation dosimetry is a means of measuring the absorbed dose of radiation in matter and/or tissue; the dose originates from the exposure to indirectly and directly ionizing radiation. As an ionizing gamma or x-ray passes through water, it leaves a trail of superoxide and hydroxyl radicals created by water ionization and splitting: $H_2O \rightarrow H_2O^{+\cdot} \rightarrow \cdot OH + H^+$. The radical will then react with the dosimeter molecule and generate an easily measurable change in the target molecule. One dosimeter that has been used to measure hydroxyl radicals generated by Fenton reaction is terephthalic acid (THA or

"para-carboxy benzoate") and has a symmetrical structure (*10, 11*). It is an appropriate dosimeter because it reacts with hydroxyl radical to yield only one mono-hydroxylated isomer. In this chapter, we chose phenylalanine over THA as the dosimeter because phenylalanine is an amino acid, the simplest unit of the proteins we wish to study, and its oxidation products are more amenable to MS analysis than are THA and monohydroxy THA.

2.3 Experimental

2.3.1 Reagents

Phenylalanine, tyrosine, acetonitrile and phosphate buffered saline (PBS) were purchased from the Sigma Aldrich Chemical Company (St. Louis, MO). Purified water (18 MΩ) was obtained from an in-house Milli-Q Synthesis system (Millipore, Billerica, MA).

2.3.2 Dosimetry experiment

Each 50 μL sample of protein was prepared in PBS (10 mM phosphate buffer, 138 mM NaCl, 2.7 mM KCl, pH 7.4 at 25 °C) with 2 mM phenylalanine. Hydrogen peroxide was added to a final concentration of 15 mM just before infusion into the FPOP region. The flowing sample solution, after laser exposure, was collected in a 0.6 mL microcentrifuge tube containing an additional 20 μL of 100 nM catalase and 70 mM methionine in PBS to quench any oxidation that might occur following FPOP.

FPOP was conducted as described previously except the ID of the fused silica was 150 μm (Polymicro Technologies, Phoenix, AZ)(*6*). The 2.45 mm beam width was

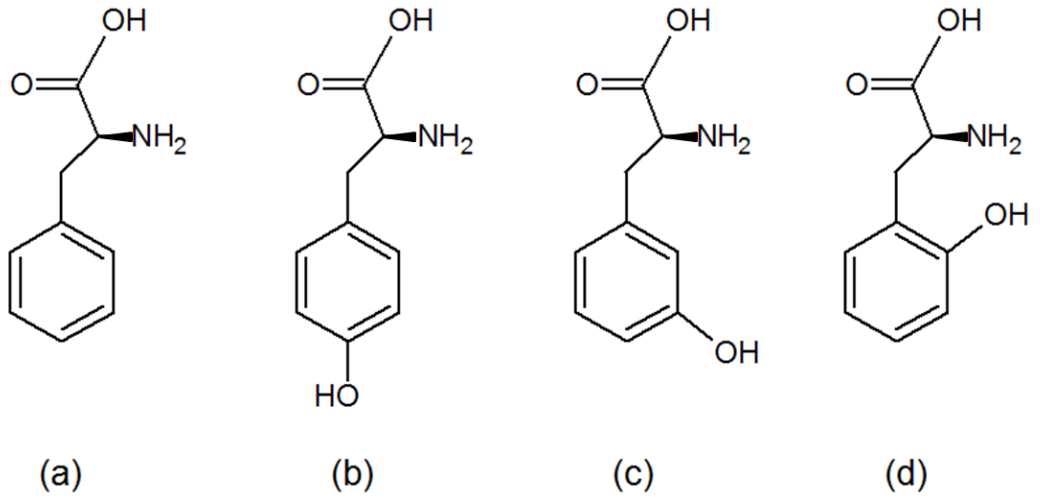
measured from a 20-shot burn pattern on label tape affixed to a temporary beam stop placed in the plane of the flow cell. Sample plugs representing an oxidation with an approximate 15% exclusion volume fraction were infused at a rate of 15.28 ± 0.04 μL/min, and the excimer pulse frequency was set to 5.00 ± 0.02 Hz.

Calibration solutions were prepared to contain everything in the above solution except for a mixture of phenylalanine and tyrosine (used as a calibrant) at a total concentration of 2 mM. The fractions of tyrosine were from 5.0% to 50.0% in increments of 5.0%. Calibration solutions were analyzed fresh after preparation.

2.3.3 Mass spectrometry

Samples were diluted 20 fold with 50% acetonitrile and directly injected for positive-ion ESI analysis on a Bruker maXis^{TM} mass spectrometer. Full mass spectra from *m/z* 150 to 1000 were acquired at a mass resolving power of 40,000 at *m/z* 1000. Samples were all analyzed using the same tune file on the same day as the sample submitted to FPOP. Each sample analysis was repeated three times.

2.4 Results and Discussion



Scheme 2.1. Structures of phenylalanine and its oxidation products. (a). phenylalanine. (b) tyrosine. (c) 3-hydroxyphenylalanine. (d). 2-hydroxyphenylalanine.

Experimental Design: Phenylalanine is a suitable dosimeter molecule because (1) it is highly reactive with OH radicals, with a rate constant of 6.9 x 10^9 M^{-1} sec^{-1} and (*12*) it yields relatively simple oxidation products. The proposed mechanism of such modification is that the hydroxyl radical adds rapidly without much positional selectivity to the aromatic ring of phenylalanine to form a hydroxycyclohexadienyl radical, which then reacts with oxygen and eliminates HOO· to form the final products (*1*). The major oxidation products from this experiment are oxidized phenylalanine, including tyrosine, 2-hydroxylphenylalanine and 3-hydroxylphenyalanine (Scheme 2.1). Hambly (*12*) previously used 20 mM phenylalanine as a dosimeter to capture and quantitate hydroxyl radicals but observed no oxidation products, likely owing to the limited dynamic range of

the analyzer mass spectrometer used in that work. We reduced the dosimeter concentration to 2 mM in our experiment. Samples were prepared and irradiated in PBS buffer to mimic the regular FPOP conditions.

The post-FPOP solution contains phenylalanine, its oxidation products, and the buffer salts. Usually such samples are desalted before they are subjected to mass spectrometry analysis. In our experiment, however, this purification step was skipped because the target molecules are small and must be preserved at their maximum quantity for accurate measurement. The mass spectrum of the post-FPOP solution acquired in ESI positive-ion mode after direct infusion (Figure 2.1) shows the formation of products that are represented by peaks corresponding to sodiated dimers in the mass spectrum, as labeled in the figure. Multiple hydroxylations are not observed, likely due to the presence of excess phenylalanine.

Measurement of the initial concentration of radicals: To quantify the oxidatively modified phenylalanine, a calibration curve was prepared by using standard mixtures of phenylalanine and tyrosine (the standard for oxidized phenyl alanine). The mass spectrum of one of the calibration solutions containing phenylalanine and tyrosine (Figure 2.2) provides a basis for measuring the fraction of the tyrosine by using the equation incorporated in the figure. A calibration curve was obtained by plotting the measured tyrosine fraction *vs.* the actual percentage of tyrosine in the solution, shown in Figure 2.3. The assumption here is that all the oxidation products with the same *m/z* share the same ionization efficiency, which is reasonable considering their structural similarity. The actual concentration of oxidation products after FPOP was then calculated to be 0.42 mM, according to the curve, and represented the initial OH radical

concentration. Self quenching of the radicals to give back hydrogen peroxide was not taken into account as self quenching would consume only a smaller fraction of the radicals. The measured radical concentration is less than but similar to the estimated value of 1 mM, used by Hambly (*6*).

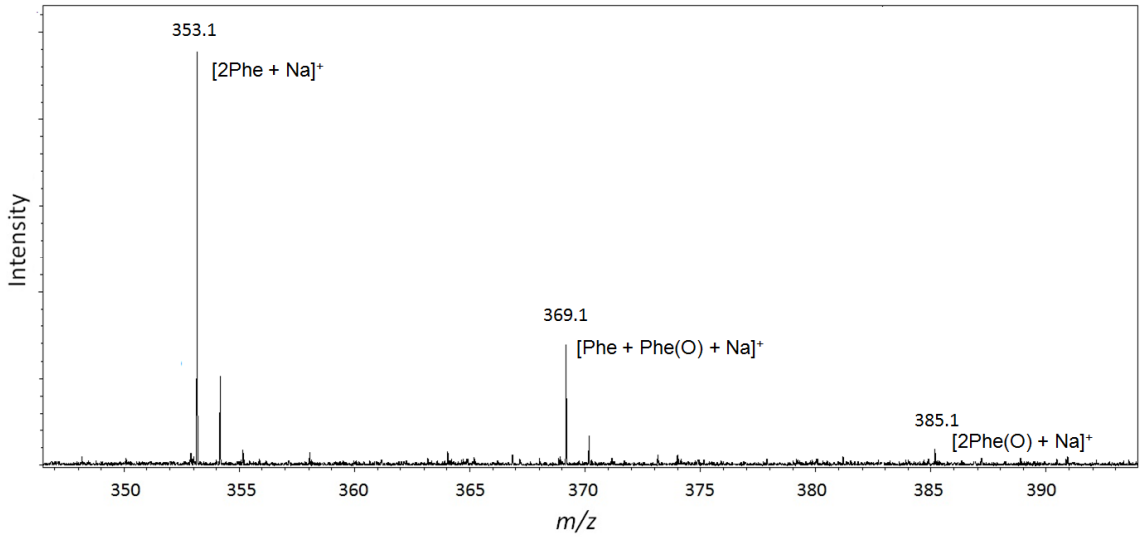


Figure 2.1: Mass spectrum of a post-FPOP sample containing phenylalanine and oxidatively modified phenylalanine.

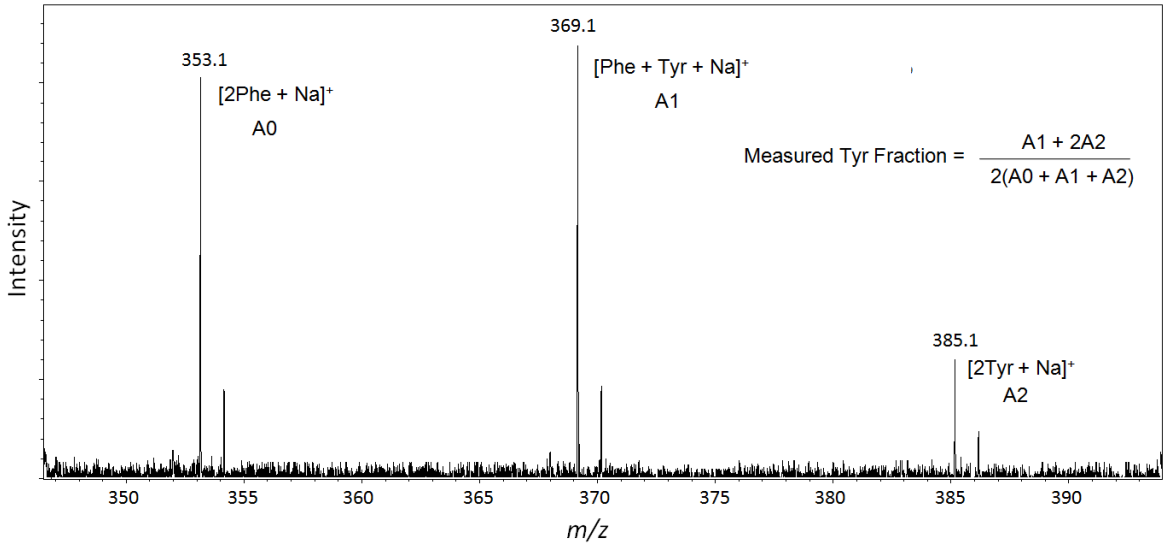


Figure 2.2: Representative mass spectrum of a calibration solution containing a mixture of phenylalanine and tyrosine.

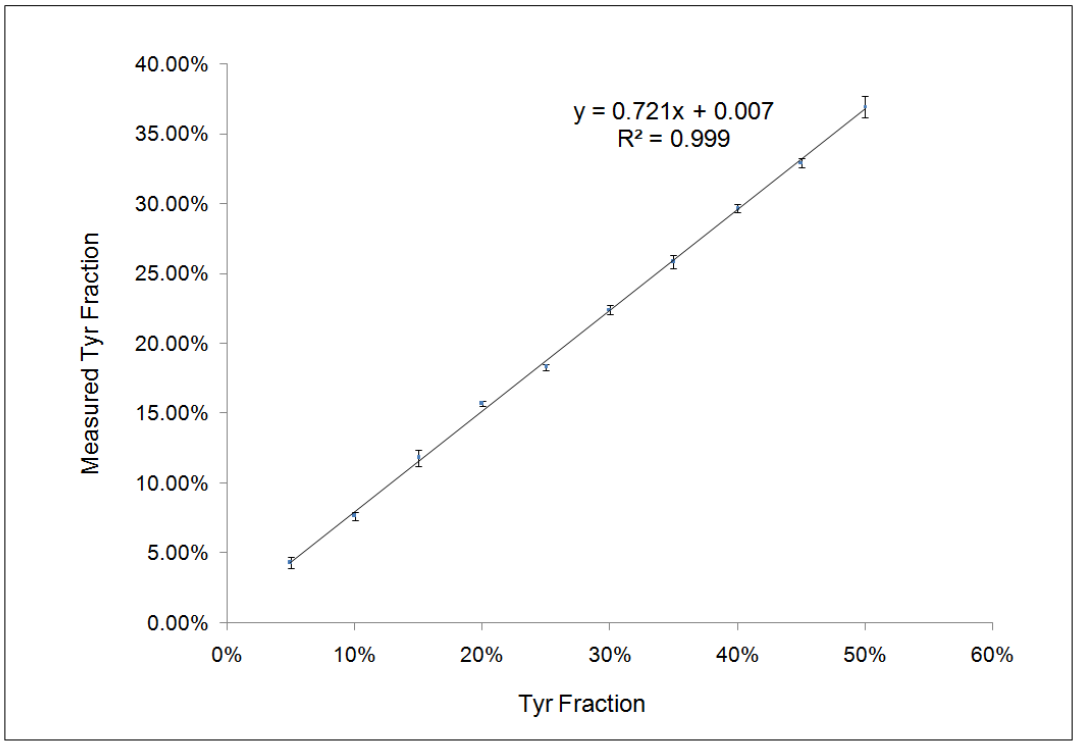


Figure 2.3: Calibration curve of measured tyrosine fraction *vs.* actual tyrosine fraction.

Lifetime of the radicals: Time courses of the hydroxyl radical concentration were obtained by numerical simulation and plotted in Figure 2.4. The simulation is based on a system of non-linear differential equations that were solved with an adaptive step-size Runge-Kutta method implemented in Mathcad version 14. The differential equations were chosen to model two second-order reactions. The first is the reaction of hydroxyl radical with itself with a rate constant of 6×10^{9} $M^{-1}s^{-1}$ and the initial concentration of 0.42 mM. The second reaction is that of the hydroxyl radical with glutamine with a rate constant of 5×10^{8} $M^{-1}s^{-1}$ and the initial concentration of glutamine as 20 mM. These equations were coupled by the common hydroxyl radical concentration. We show that in less than 1 µs, the concentration of hydroxyl radical is less than 10^{-7} M. This outcome is similar to that predicted by Hambly et al. who used an estimate for the initial concentration of ·OH (*6*).

Relation to speed of protein unfolding: The rates of folding/unfolding vary from protein to protein, depending on their size and structure. It usually takes at least two steps for a protein to fold into its native state. For some proteins, the steps are the formation of secondary structure (ns or sub µs timescale) as the early step and the formation and consolidation of tertiary structure (tens of µs and longer) as the late step (*13*). In some other proteins the steps are hydrophobic collapse (µs) and tertiary-structure formation (tens of µs and longer) (*14*). The unfolding process, which may be the reverse of folding, requires the breakdown of the tertiary structure (tens of µs and longer) and the admission of water to the structure (*15*). This timescale is significantly longer than the radical lifetime in the FPOP experiment. Although our concern is with oxidation-induced unfolding of protein rather than unfolding induced by other physical means, we can

reasonably expect that such unfolding occurs on a timescale no shorter than tens of µs in our experimental condition.

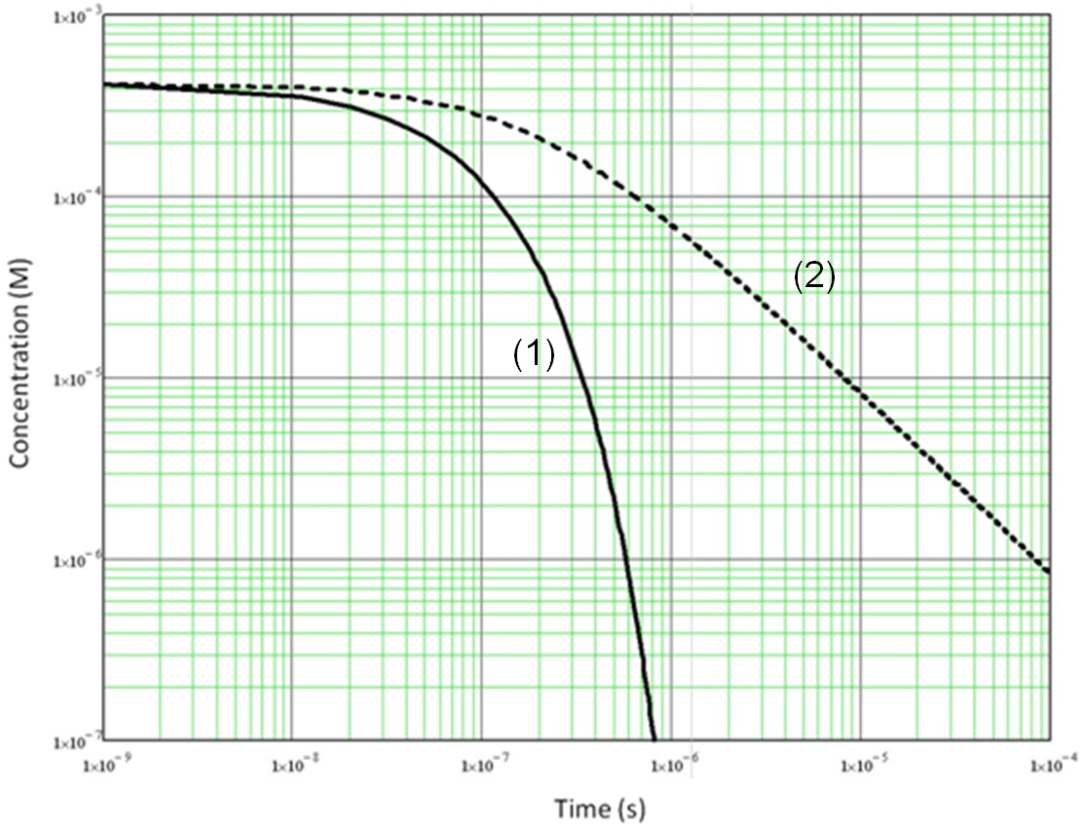


Figure 2.4: Simulation of two OH radical reactions: (1) with Gln and (2) with OH radicals themselves (self-quenching).

2.5 Conclusion

We designed a dosimetry experiment to measure the hydroxyl radical concentration produced by laser photolysis in an FPOP experiment and used that to show that the OH radical concentration is less than $1x10^{-7}$ M in less than 1 µs. This work provides further validation for the claim that FPOP modifies a protein in its native state. Furthermore, the argument that FPOP occurs in 1 µs lays the foundation for the work in the next two chapters where we apply FPOP to study protein folding on the sub ms timescale.

2.6 References

1. Xu, G., and Chance, M. R. (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics, *Chemical reviews 107*, 3514-3543.
2. Tullius, T. D., and Greenbaum, J. A. (2005) Mapping nucleic acid structure by hydroxyl radical cleavage, *Current opinion in chemical biology 9*, 127-134.
3. Maleknia, S. D., and Downard, K. (2001) Radical approaches to probe protein structure, folding, and interactions by mass spectrometry, *Mass spectrometry reviews 20*, 388-401.
4. Kiselar, J. G., Maleknia, S. D., Sullivan, M., Downard, K. M., and Chance, M. R. (2002) Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry, *International journal of radiation biology 78*, 101-114.

5. Sharp, J. S., and Tomer, K. B. (2007) Analysis of the oxidative damage-induced conformational changes of apo- and holocalmodulin by dose-dependent protein oxidative surface mapping, *Biophysical journal 92*, 1682-1692.

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

7. Aye, T. T., Low, T. Y., and Sze, S. K. (2005) Nanosecond laser-induced photochemical oxidation method for protein surface mapping with mass spectrometry, *Analytical chemistry 77*, 5814-5822.

8. Vu, D. M., Myers, J. K., Oas, T. G., and Dyer, R. B. (2004) Probing the folding and unfolding dynamics of secondary and tertiary structures in a three-helix bundle protein, *Biochemistry 43*, 3582-3589.

9. Gau, B. C., Sharp, J. S., Rempel, D. L., and Gross, M. L. (2009) Fast photochemical oxidation of protein footprints faster than protein unfolding, *Analytical chemistry 81*, 6563-6571.

10. Matthews, R. W. (1980) The radiation chemistry of the terephthalate dosimeter, *Radiation research 83*, 27-41.

11. Barreto, J. C., Smith, G. S., Strobel, N. H., McQuillin, P. A., and Miller, T. A. (1995) Terephthalic acid: a dosimeter for the detection of hydroxyl radicals in vitro, *Life sciences 56*, PL89-96.

12. Hambly, D. M., and Gross, M. L. (2007) Laser Flash Photochemical Oxidation to Locate Heme Binding and Conformational Change in Myoglobin, *Int. J. Mass Spectrom. 259*, 124-129.

13. Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1997) Fast events in protein folding: relaxation dynamics and structure of the I form of apomyoglobin, *Biochemistry 36*, 15006-15012.

14. Arai, M., Kondrashkina, E., Kayatekin, C., Matthews, C. R., Iwakura, M., and Bilsel, O. (2007) Microsecond hydrophobic collapse in the folding of Escherichia coli dihydrofolate reductase, an alpha/beta-type protein, *Journal of molecular biology 368*, 219-229.

15. Dinner, A. R., and Karplus, M. (1999) Is protein unfolding the reverse of protein folding? A lattice simulation analysis, *Journal of molecular biology 292*, 403-419.

3.* Temperature Jump and FPOP Probe Submillisecond Protein Folding

*This chapter is based on a recent publication:

Reproduced with permission from Chen, J., Rempel, D. L. & Gross, M. L. Temperature Jump and Fast Photochemical Oxidation Probe Submillsecond Protein Folding. *Journal of the American Chemical Society,* 132, 15502-15504 (**2010**)

3.1 Abstract

Understanding protein folding/unfolding is a long existing scientific problem. Here we report a new mass spectrometry-based approach to study protein-folding dynamics on the submillisecond time scale. The strategy couples a laser-induced temperature jump with FPOP whereby protein folding/unfolding is followed by changes in oxidative modifications by OH radical reactions. Using a cooled flow system containing the protein barstar as a model, we altered the protein's equilibrium conformation by applying the temperature jump and demonstrated that its reactivity with OH free radicals serves as a reporter of the conformational changes. Furthermore, we found that the time-dependent increase in mass, owing to free-radical oxidation, is a measure of the rate constant for the transition from the unfolded to the first intermediate state. This advance offers the promise that, when extended with mass-spectrometry based proteomic analysis, the sites and kinetics of folding/unfolding can also be followed at submillisecond times.

3.2 Introduction

Deciphering protein folding is essential for understanding biological processes and developing therapeutic approaches to misfolding-related diseases (*1-3*). Protein folding can be followed in equilibrium experiments, which monitor protein states as a function of temperature or denaturant concentration, and in kinetics, to give a time-dependent conformational change.

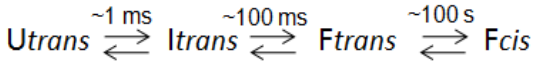
There are many techniques used to characterize protein folding/unfolding events; they include protein engineering using site-directed mutagenesis and Φ value analysis (*4-6*), time-resolved circular dichroism (CD) (*7, 8*), real-time nuclear magnetic resonance (NMR) (*9-11*), fluorescence and mass spectrometry (MS). Starting in the early 1990s, MS has emerged as an effective tool for supporting both thermodynamic and kinetic protein-folding studies. The advantage of modern MS is its ability to measure extents of protein modification and pinpoint their locations. For pulsed H/D amide exchange and other pulsed covalent labeling kinetics (*12, 13*), MS detection can now track folding events down to the millisecond timescale. Protein folding, however, often occurs more rapidly (in microseconds or even less) (*14*), which has been a difficult range to access in current MS-based studies.

Techniques for the fast initiation of protein folding include rapid mixing (*15-18*), laser-induced temperature jump (*19, 20*), optical triggering (*21*), pressure jump (*22*), dielectric relaxation and electric-field-jump (*23*). Previous mass-spectrometry-based studies of protein folding/unfolding dynamics had utilized the rapid mixing technique to create the disturbance. Konermann et al. (*24, 25*) combined rapid mixing and FPOP to observe protein folding/unfolding intermediates at times in the 10 ms range. Jha and

Udgaonkar (*26*) applied pulsed thio-labeling following the mixing of the unfolding protein and the refolding buffer, and demonstrated that millisecond folding can be assessed by global mass-spectrometry measurement.

Here we describe a new approach to investigate protein folding. We use two lasers, one to provide a temperature-jump (T-jump) and a second to generate reagents to footprint the consequences of the T-jump. This is an example of "pump/probe" (*27*, *28*), but it is distinguished by the use of chemical reactions as the structural probe rather than the usual spectroscopic (physical) approaches.

Barstar is a small protein synthesized by the bacterium *Bacillus amyloliquefaciens*. It binds to its partner barnase and inhibits the ribonuclease activity, thus countering the lethal effect of active barnase following expression (*29*). Barstar is most stable at 20-35°C, and unfolds both at low and high temperature (*30*). The folding pathway of barstar has been studies extensively and is illustrated in Scheme 3.1 (*19*, *31*, *32*). It is comprised of two intermediate states, an early one reached within ~1 ms of folding and a later intermediate established in ~10-100 ms. The folding is dominated by the slow formation of a *cis* peptidyl prolyl bond in the native protein.



Scheme 3.1. Folding pathway of barstar.

In this chapter, we will describe the instrumentation and method development using bastar as a model protein. A dosimeter experiment measuring the initial OH radical concentration upon photolysis in order to calculate the radical life time is also described.

3.3 Experimental

3.3.1 Instruments and Parts

The apparatus for the two-laser experiment consisted of a Quanta-Ray DCR Nd:YAG laser (Mountain View, CA), GAM Laser EX50/250 excimer laser (Orlando, FL), Raman shifter (to be described later), BK Precision 4001A function generator (Yorba Linda, CA), custom-built delay circuits, Thorlabs NB1-K13-Ø1'' Nd:YAG mirrors(λ=532 and 1064 nm, 0° and 45° AOI) (Newton, NJ), Thorlabs LA4184 – f = 500.0 mm - Ø1'' UV fused silica Plano-Convex lenses, Thorlabs LJ4878 – f = 75.0 mm UV fused silica Plano-Convex cylindrical lens (H = 20 mm, L = 30 mm), Thorlabs TR4 – Ø1/2''×4'' posts, Thorlabs RA90-right angle post clamps (fixed 90° adapter), Thorlabs KM100 – kinematic mirror mounts for Ø 1'' optics, Thorlabs LMR1 – lens mounts for Ø1'' optics, Thorlabs KM100C – cylindrical lens mount (kinematic), CVI Melles Griot BS1-532-67-1525-45P plate beamsplitter (Albuquerque, NM), Bruker maXis quadrupole TOF mass spectrometer (Bremen, Germany).

3.3.2 Materials

E. coli-expressed and purified barstar C82A variant was provided by Dr. Carl Frieden at a concentration of 118 μM. trifluoroacetic acid, guanidine chloride, phosphate buffered saline (PBS, pH = 7.4) and all amino acids were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO). Acetonitrile was from Honeywell Burdick and

Jackson (Muskegon, WI). Purified water (18 MΩ) was obtained from an in-house Milli-Q synthesis system (Millipore, Billerica, MA). All chemicals were used without further purification.

3.3.3 Instrument Construction

The simplified schematic of the experimental setup is shown in Figure 3.1. The 1064 nm beam produced by the Nd:YAG laser was shifted to ~1900 nm by the Raman shifter filled with the hydrogen gas (300 psi, high purity). The beam was guided through a series of optics before it reached the FPOP window and initiated the T jump. A 248 nm laser flash from the excimer laser then photolyzed added hydrogen peroxide to give hydroxyl radicals in a 14 ns pulse. The time delay between two laser flashes was controlled and adjusted by the two delay circuits (Figure 3.2). The initial low temperature of the protein solution was maintained by threading the fused silica tubing through an insulated straw that had cold air flowing through.

The arrangement of optics is illustrated in Figure 3.3. The 1064 nm beam was reflected by 4 mirrors (labeled 1-4 in the figure), which are considered optically neural, then focused by two lenses (5 and 7) before and after the Raman shifter (6). The lenses were tilted to eliminate the reflection back to the laser. The beams after the Raman shifter were a mixture of 1064 nm, 1900 nm and other anti-Stokes and Stokes lines, and were filtered by the beam splitter (8). The 1900 nm beam was then focused again and shaped by the aperture (10) and cylindrical lens (11) to an approximately 2.4 mm × 0.48 mm size at the FPOP reaction window, where it spatially overlapped with the excimer

laser beam. Figure 3.4 shows the output of MathCAD calculation. It illustrates the beam envelope *vs.* beam path length.

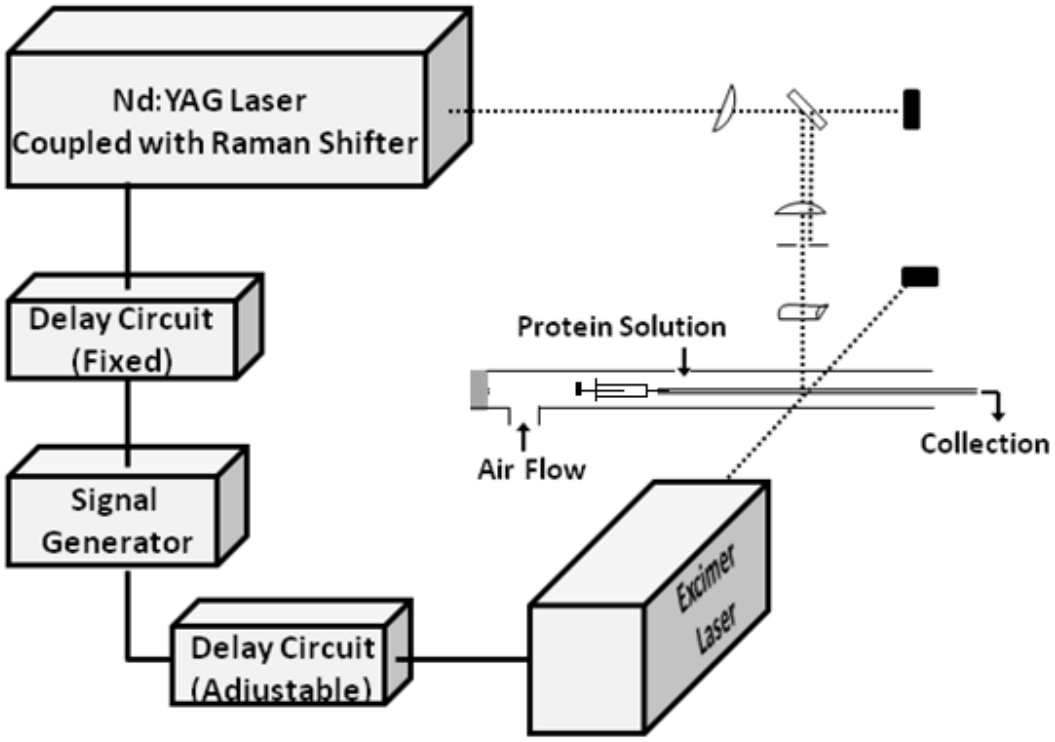


Figure 3.1: Schematic of the flow system as intersected by two laser beams at a window in the tube, as previously described for FPOP. The time between two pulses is adjustable with the "delay circuit".

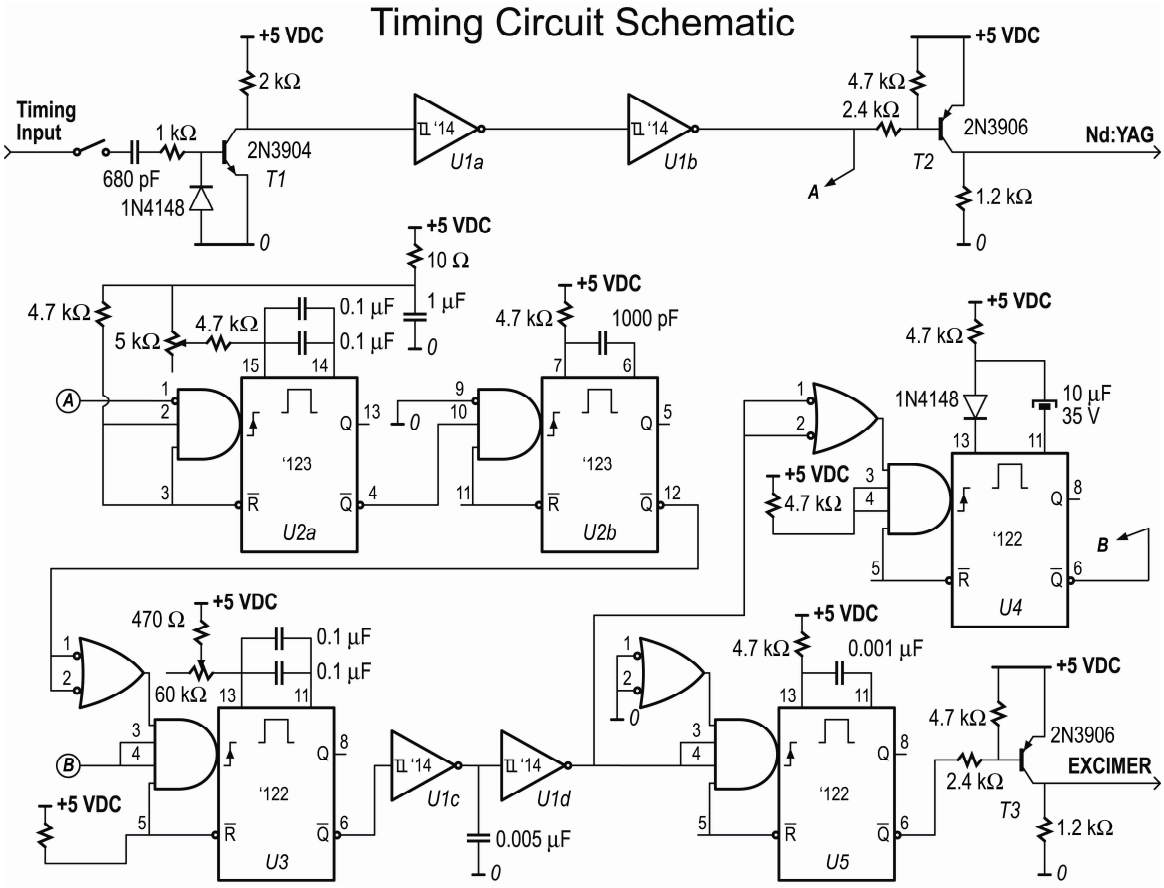


Figure 3.2: Schematic of the timing circuit. The circuit was designed and constructed by Don L. Rempel.

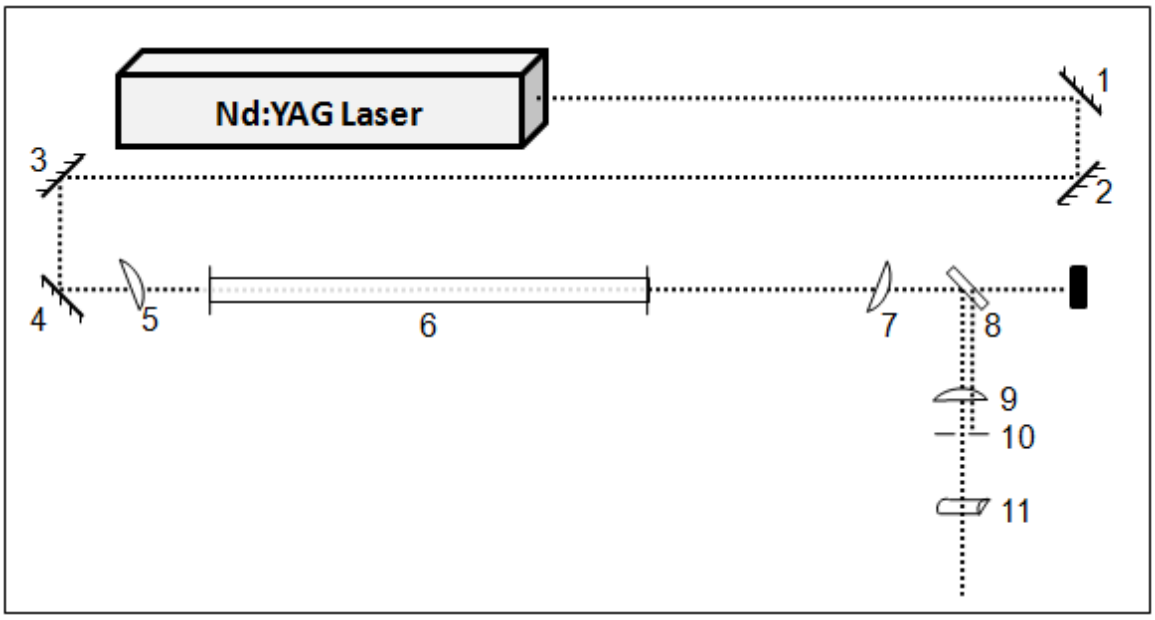


Figure 3.3: Schematic of the optic arrangement. 1-4, mirrors. 5, 7 and 9, lenses. 6, Raman shifter. 8, beam splitter. 10, aperture. 11, cylindrical lens.

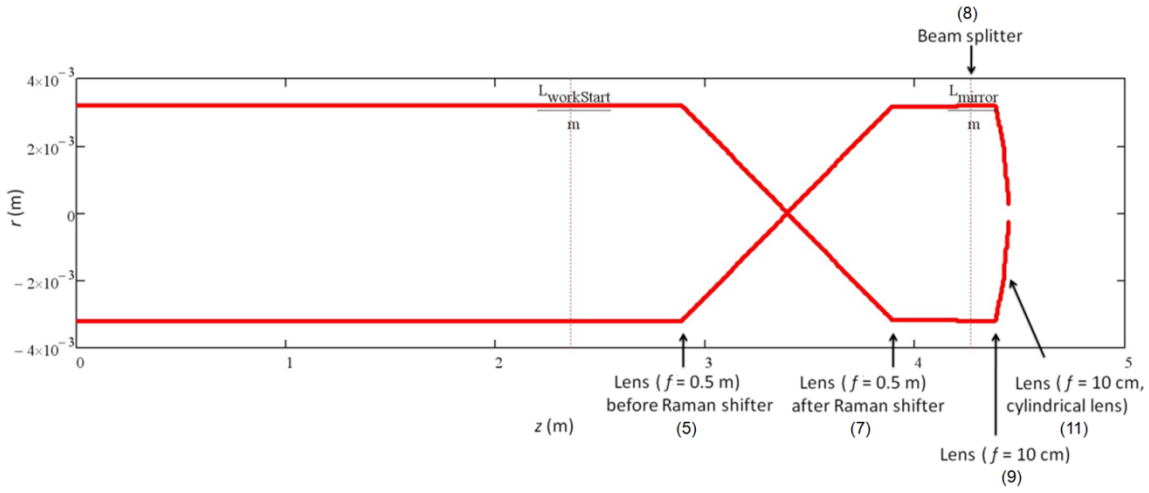


Figure 3.4: Beam envelop *vs.* beam path length. The optics are labeled as in Figure 3.3.

3.3.4 Temperature-control System

A temperature-control system was constructed for two purposes: (1) to accommodate the need of doing steady-state FPOP at different temperatures; (2) to set a specific starting temperature for the temperature-jump experiments. The system was comprised of two parts, a chamber that heats/cools the air flow and a chamber that insulates the syringe and capillary. In the first chamber, two adjustable air streams flow separately through two copper coils, one wrapped with heating strips and another immersed in ice water. These two streams meet at a junction of the T unit and then flow into the second chamber. By adjusting the flow rate at the valve before the T unit, we were able to control the temperature between 0 ± 0.5 °C and 80 ± 0.5 °C. The second chamber is shown in Figure 3.5. The box was made of insulation foam and into different layers with a removable top cover. It encloses the syringe pump, optics stand and intervening fused silica with a laser window. Force from the syringe pump is transmitted through a wooden straw to the syringe sitting inside the box. The top cover, which is not shown in the figure, has a laser window allowing the 1900 nm beam to pass perpendicular to the 248 nm beam.

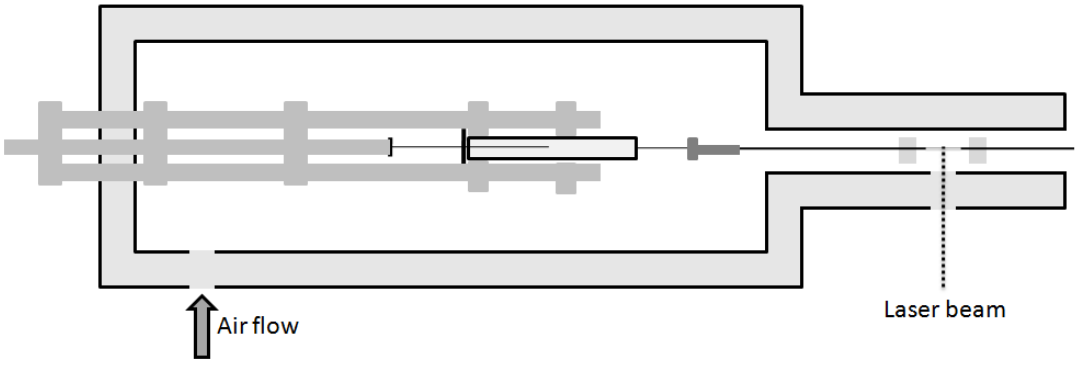


Figure 3.5: Schematic of the 2nd chamber of the temperature-control system.

3.3.5 T Jump and the FPOP Experiment

Owing to the high precision requirement of this experiment, the fine adjustment of laser beams and optics is of particular importance and should be performed with care every time before running the experiment.

First, the Raman shifter was filled with high purity hydrogen gas before the alignment. Because hydrogen gas can diffuse through stainless steel, the pressure gradually decayed with time, and the loss of hydrogen would affect the conversion efficiency. Thus, the experiment should be performed within no longer than several hours from the beginning to the end. After the Raman shifter was fixed in place, the Nd:YAG laser was tuned to its third harmonic wavelength (355 nm, green light) for the purpose of rough adjustment. The optics were adjusted to guide the beam to and align it with the capillary, tested by the reflection of the light from the capillary. Then the 248 nm excimer laser beam was aligned in the system. The two beams were adjusted to overlap with each other at the silica capillary, and this was tested with heat-sensitive paper placed at an angle of 45° to each beam. The width of the beam on the capillary was measured to be ~2.5 mm, varying slightly on different days. Both lasers were set at the frequency of 5 Hz. The flow rate was calculated based on the beam width, laser frequency and the internal diameter of the capillary (150 µm), and was approximately 15 µL/min.

Each sample was made to contain 10 µM barstar, 1.5 M GdnCl and 15 mM glutamine in PBS buffer, incubated at 0 °C for at least 2 h prior to the experiment. The

cooling system was maintained by pumping air through the copper coil that was immersed in ice water.

The time delay between the two lasers is crucial to this experiment. After each adjustment on the delay circuit, the actual delay time was measured by an oscilloscope receiving signals from the laser diode. Five μL of H_2O_2 was added to a final concentration of 15 mM just prior to FPOP infusion. The flowing sample solution was collected in an Eppendorf tube containing 20 μL of 70 mM methionine and 100 nM catalase. The sample was kept in room temperature for 10 min, allowing the breakdown of hydrogen peroxide by the catalase.

3.3.6 Mass Spectrometry Analysis

Each sample was desalted by using a $Ziptip_{C18}$ before ESI MS acquisition on a Bruker maXis^{TM} mass spectrometer. The sample was eluted in 50% acetonitrile, 0.1% formic acid in water, and then directly injected into the mass spectrometer. Full mass spectra from *m/z* 200 to 2000 were acquired at a mass resolving power of 40,000 at *m/z* 1000.

3.4 Results and Discussion

3.4.1 Global MS Measurement of Post FPOP Samples

Barstar passes through two intermediate states during its folding (*33*). The time to establish equilibration between the unfolded state and the first intermediate state is ~1 ms; between the two intermediate states, ~10-100 ms and between the second intermediate state and the native state, ~100 s. We targeted the first transition, which had not been resolved by MS methods. A close-up view of the reaction region, Figure 3.6,

shows that the barstar begins in an unfolded state (Fig. 3.6 (a)). As it passes the laser-beam intersection, it exposes to the 1900 nm laser flash and starts to refold. After a while of refolding, the 248 nm laser is triggered and photodissociates the hydrogen peroxide to hydroxyl radicals which label the solvent accessible regions of barstar in a particular folding state (Fig. 3.6 (b)). Further folding or other structural change after the FPOP labeling (Fig. 3.6 (c)) is not relevant in this experiment and will not be sampled.

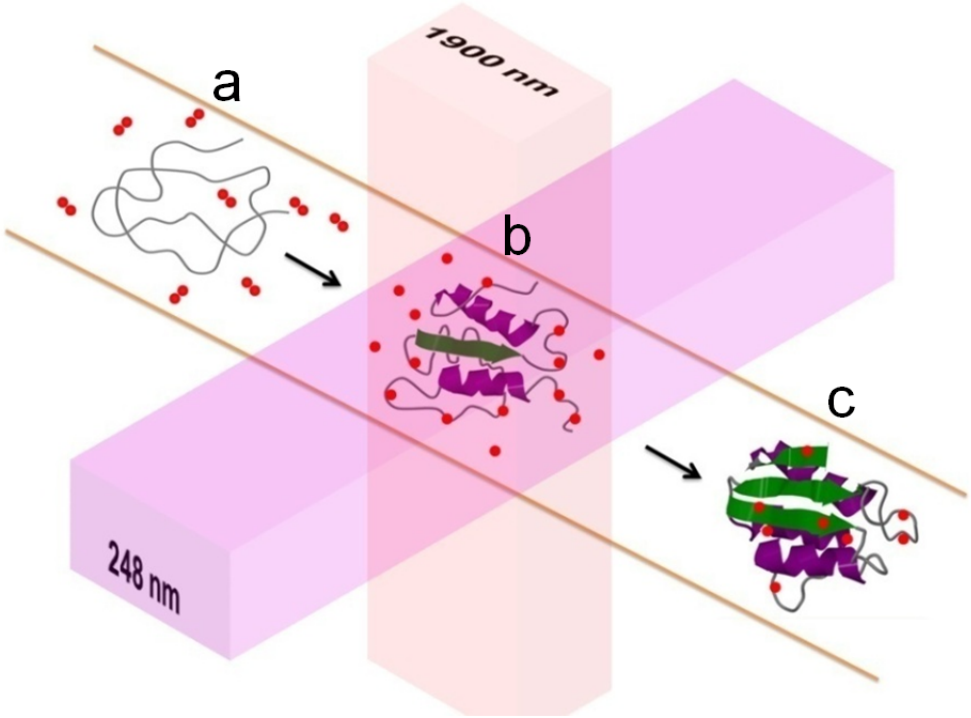


Figure 3.6: Close-view of the intersection of the two laser beams in the T-Jump Pulse, FPOP probe experiment while solution containing barstar flow through the junction.

Representative mass spectra of the most abundant charge state of barstar, post-FPOP, show the clear sensitivity of FPOP-labeling extent to the folding that occurs as a function of the delay between the two laser pulses (Fig. 3.7). As the protein folds, the extent of oxidative modification with ·OH decreases. Each oxidation state is represented by a peak, which is an envelope of isotopic clusters that are partially resolved here.

Outcomes for oxidative modification of other proteins by FPOP (*34*) are Poisson distributions for the various modified states that are dominated by +16, +32, +48…, indicating that the protein exists in a single state during modification. Here we obtain spectra of the entire protein ("global spectra") that show many more modifications, consistent with a denatured protein existing in many states. The amounts of the modified species decrease with an increase in the time difference between the heat pulse and the FPOP probe, showing clear evidence of protein refolding on the < 1 ms time scale. The extent of oxidation shown in Fig. 3.7 (e) is greater than that for folded barstar at room temperature (RT) (Fig. 3.7 (f)). This indicates that the protein at its first intermediate state possesses a structure that is more solvent-exposed compared to that of the native state and, hence, undergoes more oxidative modification than does the native state. Furthermore, some fraction of the protein may remain less folded because the heating along the laser beam is not uniform.

To illustrate the temperature jump profile, we used the basic property of water and an estimate of the power flux of the laser. The energy flux of the laser beam decreases over distance as the beam passes through the water, as described by the absorption coefficient. The equation is shown as follows:

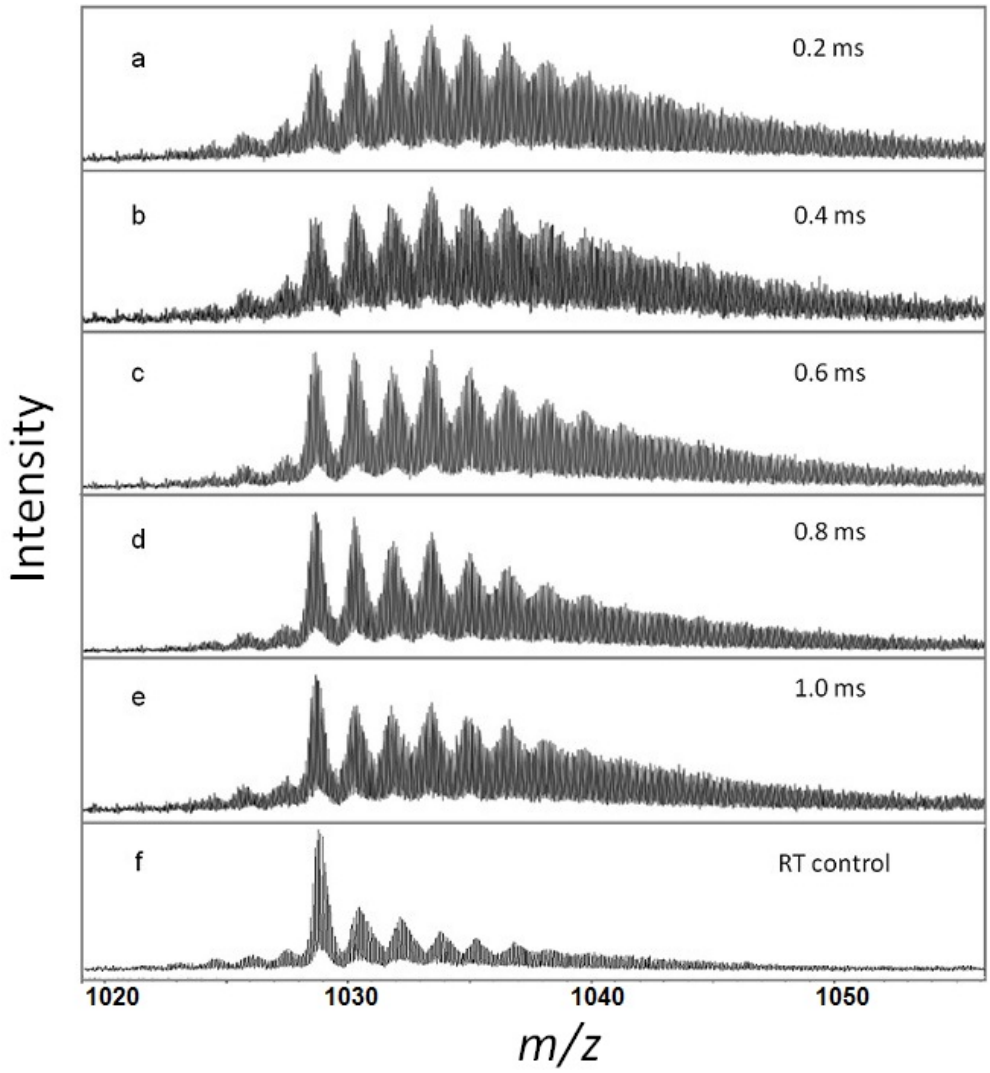


Figure 3.7: (a)-(e). Representative mass spectra of barstar post FPOP as a function of the time between the heating pulse and the FPOP probe. (f). Mass spectrum of the oxidatively modified barstar post FPOP at RT as a control.

$$E_{beam} = E_0 e^{-\alpha \cdot x}$$

The temperature jump is related to the rate of energy transfer from the beam to the water by the heat capacity and the density of water, and is shown in the following equation:

$$\Delta T(x) = \frac{E_0 \alpha e^{-\alpha \cdot x}}{C_p \cdot \rho}$$

In the calculation we used 20 mJ distributed over the 2.5 mm × 0.5 mm area as E_0. And the absorption coefficient is 80 cm^{-1} (*35*). We plotted the equation using the values in Figure 3.8. It shows a T jump of 30 K at the entrance and 10 K in the end, which is one inside diameter of the capillary (150 µm).

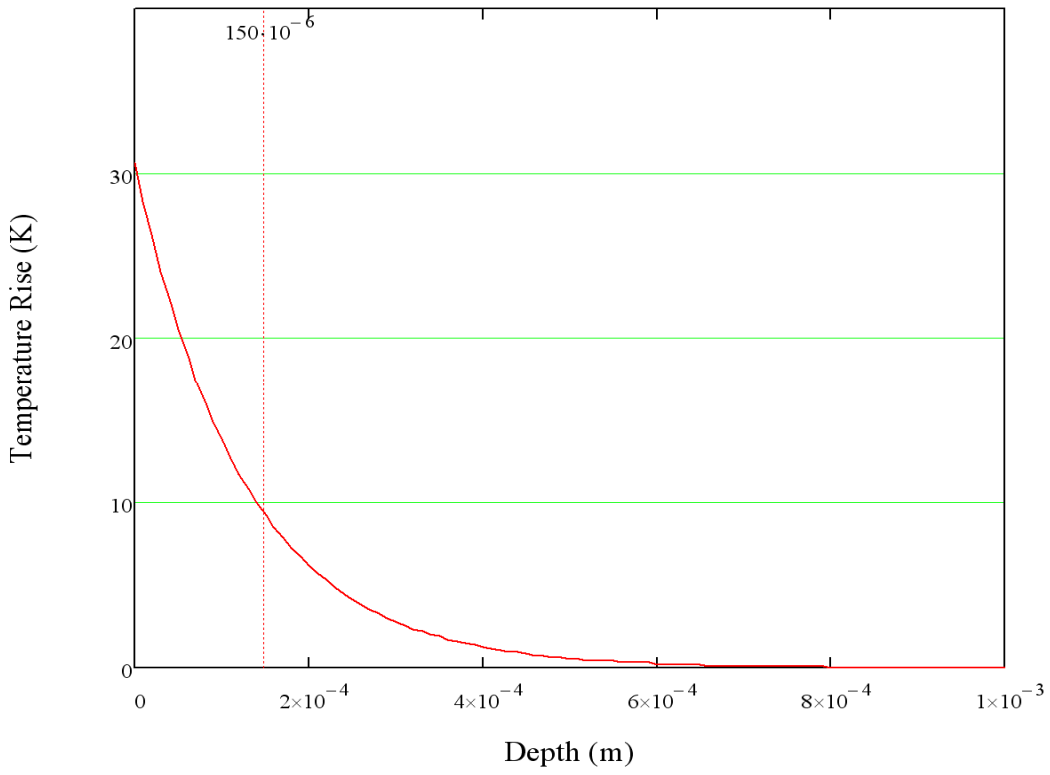


Figure 3.8. An estimate of temperature jump profile.

3.4.2 Kinetic Analysis Reveals Folding Rate Constant

We measured the rate constant for establishing folding equilibrium by using the centroid of all the oxidized states having a 10+ charge, which was the most abundant. Because the centroid shifts to higher *m/z* as the oxidation progresses, the average mass of the oxidatively modified protein decreases as folding occurs). Thus, this approach has been called "weighing protein folding" in a highlight in the journal Nature (*36*). There were two parts of the *m/z* centroid calculation of each mass spectrum. The first part estimated the baseline by taking the average spectral intensity over the interval from *m/z* 120/10 to 60/10 less than the theoretical monoisotopic *m/z*, where 10 is the magnitude of the charge state; the interval showed no discernable isotopic patterns. The second part used the centroid of the peak area that lay above the baseline in the *m/z* interval 7×16/10 (the unmodified isotopic pattern plus the first six oxidation states) starting from the theoretical monoisotopic *m/z*. The *m/z* centroid trend was fit by a single exponential function including a constant; the fitting utilized the rate constant, amplitude, and the constant as parameters.

The rate constant from the fitting is 1.5 ms^{-1} (plot in figure 3.9). Because the *m/z* centroid trend depends on the concentrations of the unfolded and the first intermediate states in a way that does not easily reveal the concentration ratio, the trend reveals only the equilibration rate constant. This rate compares with a value of 3.10 ms^{-1}, which was measured by Nolting et al (*32*), in a T-jump experiment with fluorescence detection for the transition from the unfolded state to the first intermediate state. The amplitude we find is -0.69, as normalized by the *m/z* centroid shift from 0.2 ms to a control obtained by FPOP of the folded barstar at RT without the 1900 nm heating or temperature-jump pulse

(Fig. 3.7 (f)). This outcome indicates that barstar reached a steady state that is not yet the completely folded state.

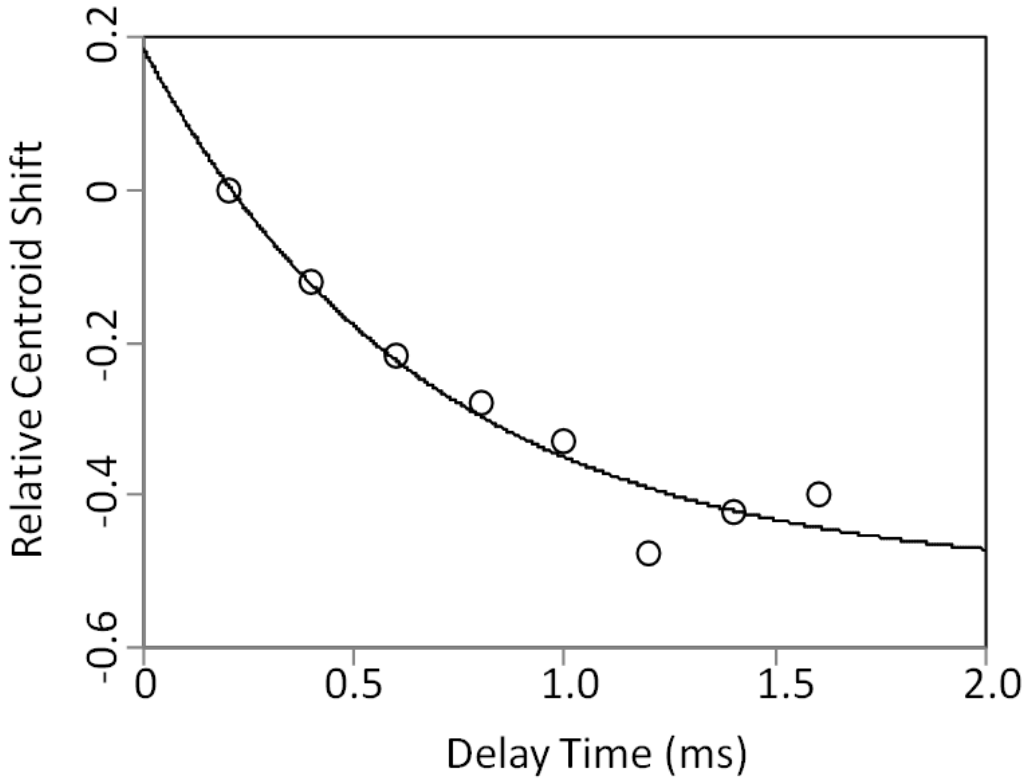


Figure 3.9: Plot of the normalized centroid mass shift *vs.* the delay time. The curve (solid line) was obtained by fitting a single exponential function to the data (open circles) to give a rate constant of 1.5 ms^{-1}. The standard deviation was estimated to be 0.06 by triplicate measurements of the centroid at one time point, and the precision of the method was estimated by repeating the kinetics study on another day to provide a k of 1.0 ms^{-1}.

3.5 Conclusion

T-jump coupled with FPOP allows sub msec protein folding kinetics to be followed by a subsequent MS measurement at the global protein level. Although we followed the folding kinetics at sub msec level for barstar, we suggest that kinetics in the µs range can be reached by taking advantage of the time scale of the two lasers, 10 ns for the Nd:YAG laser and 20 ns for the excimer laser, and the quenching time for the hydroxyl radical, which occurs in less than 1 µs. An advantage of the approach is the ability to determine modifications and follow kinetics at the amino acid level, especially for large proteins, by using proteolysis and LC-MS/MS. It can provide a broad view of the protein whereas physical measurements often depend on only a few amino-acid residues and, therefore, have low resolution or coverage. This will be expanded and discussed in the next chapter. Another advantage of the MS-based approach is that it is general and does not rely on the state of any fluorophore, or any protein engineering, thus allows direct study of protein folding/unfolding mechanisms.

This MS-based approach has considerable flexibility. Other photochemical processes can be used instead of T-jump as a pump, and other labeling reagents (reactive free radicals and radical ions) can be produced by photolysis and used as the probe. We are currently developing some of these reagent radicals. The molecular mass of the protein is not a serious limitation of the method although measuring the varying molecular weight of a protein as a function of delay time becomes more difficult as protein MW increases. If large MW becomes a limitation, one can digest the protein and follow the mass of reporter peptides instead of the mass of the protein. Lastly, we suggest that protein-protein and protein-ligand binding dynamics can also be investigated by the

two-laser approach whereby the first laser perturbs the thermodynamics and the second uses chemistry to probe the outcome (*27*).

3.6 Acknowledgements

We thank Prof. Dewey Holten for the Raman shifter and Prof. Carl Frieden and Dr. Greg DeKoster for the barstar sample. This research was supported by the NCRR of NIH (Grant No. P41RR000954).

3.7 References

1. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884-890.
2. Luheshi, L. M., Crowther, D. C., and Dobson, C. M. (2008) Protein misfolding and disease: from the test tube to the organism, *Current opinion in chemical biology 12*, 25-31.
3. Cohen, F. E., and Kelly, J. W. (2003) Therapeutic approaches to protein-misfolding diseases, *Nature 426*, 905-909.
4. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, *Journal of molecular biology 224*, 771-782.
5. Serrano, L., Matouschek, A., and Fersht, A. R. (1992) The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure, *Journal of molecular biology 224*, 805-818.

6. Matouschek, A., Serrano, L., and Fersht, A. R. (1992) The folding of an enzyme. IV. Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure, *Journal of molecular biology 224*, 819-835.

7. Johnson, W. C., Jr. (1990) Protein secondary structure and circular dichroism: a practical guide, *Proteins 7*, 205-214.

8. Nolting, B., Jung, C., and Snatzke, G. (1992) Multichannel circular dichroism investigations of the structural stability of bacterial cytochrome P-450, *Biochimica et biophysica acta 1100*, 171-176.

9. Akasaka, K., Naito, A., and Nakatani, H. (1991) Temperature-jump NMR study of protein folding: ribonuclease A at low pH, *Journal of biomolecular NMR 1*, 65-70.

10. Balbach, J., Forge, V., van Nuland, N. A., Winder, S. L., Hore, P. J., and Dobson, C. M. (1995) Following protein folding in real time using NMR spectroscopy, *Nature structural biology 2*, 865-870.

11. Hoeltzli, S. D., and Frieden, C. (1995) Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-19F-tryptophan-labeled Escherichia coli dihydrofolate reductase, *Proceedings of the National Academy of Sciences of the United States of America 92*, 9318-9322.

12. Kaltashov, I. A., and Eyles, S. J. (2005) *Mass Spectrometry in Biophysics*, John Wiley and Sons, Inc.

13. Uzawa, T., Nishimura, C., Akiyama, S., Ishimori, K., Takahashi, S., Dyson, H. J., and Wright, P. E. (2008) Hierarchical folding mechanism of apomyoglobin

revealed by ultra-fast H/D exchange coupled with 2D NMR, *Proceedings of the National Academy of Sciences of the United States of America 105*, 13859-13864.

14. Yang, W. Y., and Gruebele, M. (2003) Folding at the speed limit, *Nature 423*, 193-197.
15. Koide, S., Dyson, H. J., and Wright, P. E. (1993) Characterization of a folding intermediate of apoplastocyanin trapped by proline isomerization, *Biochemistry 32*, 12299-12310.
16. Shastry, M. C., Luck, S. D., and Roder, H. (1998) A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale, *Biophysical journal 74*, 2714-2721.
17. Kauffmann, E., Darnton, N. C., Austin, R. H., Batt, C., and Gerwert, K. (2001) Lifetimes of intermediates in the beta -sheet to alpha -helix transition of beta -lactoglobulin by using a diffusional IR mixer, *Proceedings of the National Academy of Sciences of the United States of America 98*, 6646-6649.
18. Hertzog, D. E., Michalet, X., Jager, M., Kong, X., Santiago, J. G., Weiss, S., and Bakajin, O. (2004) Femtomole mixer for microsecond kinetic studies of protein folding, *Analytical chemistry 76*, 7169-7178.
19. Nolting, B., Golbik, R., and Fersht, A. R. (1995) Submillisecond events in protein folding, *Proceedings of the National Academy of Sciences of the United States of America 92*, 10668-10672.
20. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) Direct observation of fast protein folding: the initial collapse of apomyoglobin, *Proceedings of the National Academy of Sciences of the United States of America 93*, 5759-5764.

21. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) Fast events in protein folding initiated by nanosecond laser photolysis, *Proceedings of the National Academy of Sciences of the United States of America 90*, 11860-11864.

22. Jacob, M., Holtermann, G., Perl, D., Reinstein, J., Schindler, T., Geeves, M. A., and Schmid, F. X. (1999) Microsecond folding of the cold shock protein measured by a pressure-jump technique, *Biochemistry 38*, 2882-2891.

23. Porschke, D. (1996) Electrostatics and electrodynamics of bacteriorhodopsin, *Biophysical journal 71*, 3381-3391.

24. Stocks, B. B., and Konermann, L. (2009) Structural characterization of short-lived protein unfolding intermediates by laser-induced oxidative labeling and mass spectrometry, *Analytical chemistry 81*, 20-27.

25. Stocks, B. B., and Konermann, L. (2010) Time-dependent changes in side-chain solvent accessibility during cytochrome c folding probed by pulsed oxidative labeling and mass spectrometry, *Journal of molecular biology 398*, 362-373.

26. Jha, S. K., and Udgaonkar, J. B. (2007) Exploring the cooperativity of the fast folding reaction of a small protein using pulsed thiol labeling and mass spectrometry, *The Journal of biological chemistry 282*, 37479-37491.

27. Callender, R., and Dyer, R. B. (2006) Advances in time-resolved approaches to characterize the dynamical nature of enzymatic catalysis, *Chemical reviews 106*, 3031-3042.

28. Kubelka, J. (2009) Time-resolved methods in biophysics. 9. Laser temperature-jump methods for investigating biomolecular dynamics, *Photochem Photobiol Sci 8*, 499-512.

29. Hartley, R. W. (1989) Barnase and barstar: two small proteins to fold and fit together, *Trends in biochemical sciences 14*, 450-454.

30. Agashe, V. R., and Udgaonkar, J. B. (1995) Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride, *Biochemistry 34*, 3286-3299.

31. Shastry, M. C., and Udgaonkar, J. B. (1995) The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates, *Journal of molecular biology 247*, 1013-1027.

32. Nolting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) The folding pathway of a protein at high resolution from microseconds to seconds, *Proceedings of the National Academy of Sciences of the United States of America 94*, 826-830.

33. Schreiber, G., and Fersht, A. R. (1993) The refolding of cis- and trans-peptidylprolyl isomers of barstar, *Biochemistry 32*, 11195-11203.

34. Gau, B. C., Sharp, J. S., Rempel, D. L., and Gross, M. L. (2009) Fast photochemical oxidation of protein footprints faster than protein unfolding, *Analytical chemistry 81*, 6563-6571.

35. Curcio, J. A., and Petty, C. C. (1951) The near infrared absorption spectrum of liquid water, *J Opt Soc Am 41*, 302-304

36. Gruebele, M. (2010) Analytical biochemistry: Weighing up protein folding, *Nature 468*, 640-641.

4. Watching Protein Fast Folding on the Amino-Acid Residue Level

4.1 Abstract

We report here the study of sub millisecond protein folding with amino-acid residue resolution achieved a two-laser pump/probe experiment with analysis by mass spectrometry. Barstar protein folding can be triggered by a laser-induced temperature jump (T jump) from ~0 ℃ to room temperature. Subsequent FPOP (fast photochemical oxidation of proteins) reaction at various time points after the T jump leads to oxidative modification of solvent-accessible side chains whose "protection" is changing with time. The modifications are then identified and quantified by LC-MS/MS following proteolysis. Among all the segments that form secondary structure in the native state, the $helix_1$ segment shows a decreasing trend of oxidative modification during the first 1 ms of folding while others do not exhibit a clear trend. Residues I5, H17, L20, L24 and F74 are modified less in the intermediate state than the denatured state, likely due to full or partial protection of these residues as folding occurs. We propose that in the early folding stage, barstar forms a partially solvent-accessible hydrophobic core consisting of several residues that have long-range interactions with other, more remote residues in the protein sequence. Although it is not realistic to correlate directly modification with secondary structure formation, out data are consistent with the previous conclusion that barstar fast folding follows the nucleation-condensation mechanism with the nucleus centered on $helix_1$ formed in a folding intermediate.

4.2 Introduction

The mechanism of protein folding remains a central problem of molecular biology. There are two competing models for protein folding: hydrophobic collapse and the framework model. In the former model, hydrophobic collapse, a relatively early event in the folding pathway, happens before the many secondary structures can form(*1*). In the framework model, secondary structures form during the early stages of folding, and they guide the slow development of the tertiary structure (*2*). Due to the difficulty of experimentally accessing the early folding events, most of the folding mechanism studies have been done *in silico* via molecular dynamics and simulations of folding (*3-5*). There is experimental evidence, however, that many globular proteins, such as myoglobin (*6*) and staphylococcal nuclease (*7*), follow the partial hydrophobic collapse model while folding.

One way to access folding mechanism is to characterize the protein folding intermediate, which presents challenges to analytical methods. Conventional rapid mixing techniques are adequate for studying the folding for many two-state systems and later stages of multistate systems. To detect the intermediates formed in more complex folding pathways, however, faster methods (e.g., T jump) are required (*8*). The folding kinetics are then analyzed by spectral probes, which detect the gross features of intermediates (*9, 10*). The only method so far that follows the fate of individual amino acid residues in the transition states is a Φ-value analysis that compares the folding kinetics and stability of the wild-type protein and its point mutants (*11-14*). It is an indirect method based on the assumption that there is a close relationship between protein

structure and energy, and that amino-acid substitutions (usually from mutation) do not significantly alter the folding pathway.

Barstar, an 89-residue single-domain protein containing four helices and three β-sheets, is an ideal model for folding studies. Its structure was solved by X-ray (*15*) and NMR (*16*). Barstar undergoes reversible unfolding transitions when denatured by high concentration of GdnHCl or urea, extreme pHs, or low temperature, and refolds under appropriate conditions (*17, 18*). The folding pathway of barstar is now extensively characterized (*17-19*). It was suggested by Agashe et al. (*20*) that during the first several milliseconds of barstar folding, the polypeptide chain rapidly collapses to a compact globule possessing a solvent-accessible hydrophobic core but without optically active secondary or tertiary structure. Later, the combination of temperature-jump pump coupled with a fluorescence probe revealed that barstar folds to an intermediate in about 1 ms with 50% of the surface area buried compared to the denatured state (*10*). The authors proposed that barstar folding follows the general nucleation-condensation model whereby the nucleus is centered on the first helix that is consolidated in the intermediate state. The above two proposals seem to be contradictory on whether the helix has formed during the initial folding, likely owing to different experiment conditions, but agree on the folding model of barstar. Recently, protein simulation applied to barstar shows that $helix_1$ is the first to begin forming out of all secondary structures (*5*).

Recent advances in mass spectrometry (MS)-based protein footprinting have produced methods that allow the study of folding by examining solvent accessibility as an indicator of protein conformational change, as assessed by footprinting and mass spectrometry (*21, 22*). Although the overall change of solvent accessibility is reflected

by a global mass measurement (*23*), the information on individual residues is more valuable in determining regional structure changes, which can be revealed by bottom-up proteomic analysis of such protein samples. In this chapter, we study barstar folding at the amino-acid residue level by combining temperature-jump pump and an FPOP probe, discussed in Chapter 2, followed by proteolysis and LC/MS analysis. This is a more direct method than Φ-value analysis and offers the same high-resolution.

4.3 Experimental

4.3.1 Materials

E. coli-expressed and purified barstar C82A variant was provided by Dr. C. Frieden and Dr. G. DeKoster at the concentration of 118 μM. HPLC-grade water, 30% H_2O_2, *L*-glutamine, *L*-methionine, catalase, guanidinium chloride (GdnCl), phosphate buffered saline (PBS, pH = 7.4) and proteomic-grade trypsin were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO). Acetonitrile was from Honeywell Burdick and Jackson (Muskegon, WI). All chemicals were used without further purification.

4.3.2 Equilibrium Studies

Cold and heat denaturation of barstar was studied by a circular dichroism (CD) temperature scan at 222 nm on a J-815 CD spectrometer (JASCO Analytical Instruments, Tokyo, Japan). The barstar sample in PBS buffer with 1.2 M GdnCl was incubated at 0 °C for 1 h before CD measurement. During the CD scanning, the sample was heated to 30 °C at 20 °C·hr^{-1} and then to 70 °C at 40 °C·hr^{-1}.

4.3.3 Two-color FPOP

Each protein sample was made to contain 10 μM barstar, 1.2 M GdnCl, and 15 mM glutamine in PBS buffer, incubated at 0 °C for 3 h prior to the experiment. The experiment was conducted according to the previously described method (*23*) but with the following modification. The 150 μm i.d. fused silica sits in a cooling system comprised of a thermally insulated box with two chambers abutting the FPOP apparatus. The 1st chamber contained copper tubing connected to a compressed air supply and soaked in an ice bath. The 2nd chamber, into which the copper tubing emptied, encloses the syringe pump, optics stand, and intervening fused silica, with a 2 cm^2 window for laser transmission. The temperature in the 2nd chamber was kept to less than 3°C by adjusting the air flow through the ice bath.

The Nd:YAG laser (Quanta-Ray, Mountain View, CA) was set at its full power and the KrF excimer laser power (GAM Laser Inc., Orlando, FL) was set at 45 mJ/pulse, with the pulse frequency of both lasers set to 5 Hz. The flow rate was adjusted to ensure a 25% exclusion volume to avoid repeat •OH exposure. Eight different time delays between two lasers were applied in this experiment: 0.1 ms, 0.2 ms, 0.4 ms, 0.6 ms, 0.8 ms, 1.0 ms, 1.5 ms and 2.0 ms. After each adjustment on the delay circuit, the actual delay time was measured with the oscilloscope using the signals detected by the laser diode. Two sets of control samples were submitted to FPOP at cold temperature and room temperature, respectively, without T jump. Five μL of H_2O_2 was added to a final concentration of 15 mM just prior to FPOP infusion. The flowing sample solution was collected in an Eppendorf tube containing 20 μL of 70 mM methionine and 100 nM

catalase. The sample was kept at room temperature for 10 min, allowing the breakdown of peroxide by catalase.

4.3.4 Protein Digestion and LC/MS

Each protein sample was split into two vials and dried under vacuum in a SpeedVac. One vial of sample, to be digested by trypsin, was dissolved in 100 mM ammonium bicarbonate buffer. Trypsin solution (0.4 µL) was prepared according to the manual protocol and added to each sample. Samples were stored at 37 °C for 8 h. The other vial of sample, to be digested by Glu-C, was dissolved in 25 mM ammonium carbonate buffer (pH 7.8). The incubation time was 6 h at 25 °C at a weight ratio of 1/40, enzyme:protein.

An aliquot (5 µL) of sample after 1:5 dilution, was loaded onto a custom-built silica capillary column packed with C18 reverse-phase material (Magic, 5 µm, 300 Å, Michrom, Auburn, CA). The gradient was from 2% solvent B (97% acetonitrile, 3% water, 0.1% formic acid) and 98% solvent A (97% water, 3% acetonitrile, 0.1% formic acid) to 50% solvent B over 50 min, then to 85% solvent B for 5 min at a flow rate of 260 nL/min followed by a 5 min re-equilibration step. The solution was sprayed directly from the column into an LTQ-Orbitrap mass spectrometer (Thermo Fisher, Waltham, MA) by using a PicoView PV-500 nanospray source (New Objective, Woburn, MA). A full mass spectrum of eluting peptides was recorded at high mass resolving power (100,000 for ions of *m/z* 400) with the FT mass spectrometer component while MS/MS experiments on the six most abundant ions from the eluent were performed in the LTQ at a normalized collision energy of 35% of the maximum, using a 2 Da isolation width and

wide-band activation. Ions submitted to MS/MS were placed in a dynamic exclusion list for 8 s.

4.3.5 Data Processing

LC-MS features of all acquisitions were aligned by Rosetta Elucidator (Microsoft, Bellevue, WA) peak detection and alignment software (a "feature" is the naturally occurring isotopic ensemble of one molecule eluting in time). Features were quantified by integrating the areas of all co-eluting LC-MS extracted ion chromatogram (EIC) peaks having the same monoisotopic mass within a 5 ppm resolution tolerance.

The software assigned a feature with a unique ID and associated all product-ion (MS^2) spectra with their LC-MS features by using the same unique ID nomenclature. Independent from the Elucidator analysis, the product-ion spectra were searched against a restricted database containing barstar C82A by using Mascot error-tolerant searching. An Excel-based VBA program, designed by Brian Gau, associated the Mascot calls with their LC-MS features by using the unique ID. This program matched LC-MS Mascot annotations to a theoretical FPOP-modified tryptic peptide list of barstar C82A. These matches, and over 60% of the Mascot calls, were manually validated, corrected, or rejected based on their product-ion (MS^2) spectra before the per-peptide and per-residue yield analysis.

Per-residue yields were calculated as follows:

$$residue_i\ yield = \frac{\sum peptide\ intensities\ modified\ at\ residue_i}{\sum peptide\ intensities\ with\ same\ sequence\ as\ numerator\ peptides}$$

Here the denominator includes both modified and unmodified peptides but excludes signal from missed-cleavage peptides spanning $residue_i$ if such peptides are not also

detected as modified there. Per-peptide yields were determined according to the following equation:

$$peptide_i\ yield = \frac{\sum modified\ peptide_i\ intensities}{\sum modified\ peptide_i\ intensities + unmodified\ peptide_i\ intensity}$$

4.4 Results and Discussion

4.4.1 Temperature Effect on Barstar Conformation

Barstar C82A in the presence of 1.2 M GdnCl unfolds at both high and low temperatures, as indicated by a far-UV CD measurement at 222 nm (Fig. 4.1) to monitor secondary structure. The difference in solvent accessibility between the cold-denatured structure and the folded structure was also evaluated by FPOP footprinting and mass spectrometry measurement, serving as the two end-point controls for the T jump experiment, both on the global scale and the amino-acid residue level.

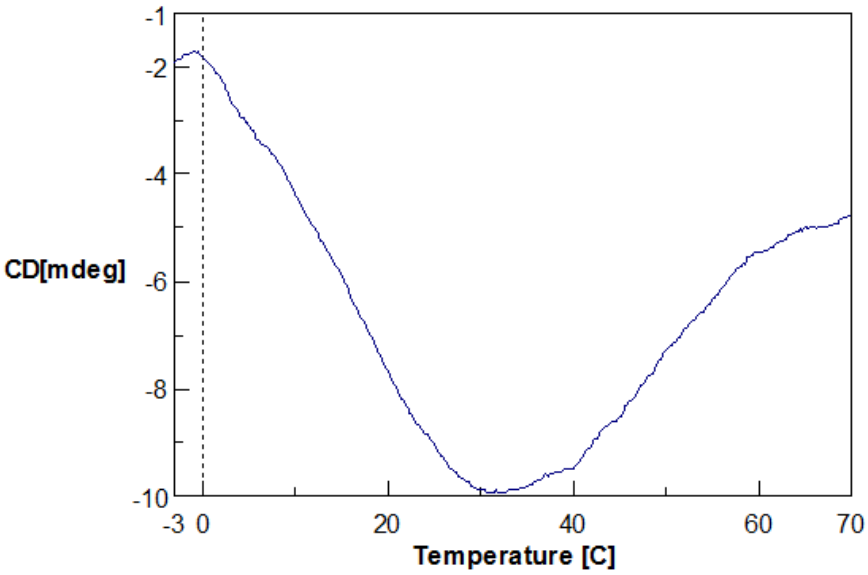


Figure 4.1: CD spectrum of barstar as function of temperature shows that the protein is unfolded at both low and high temperature. CD temperature scan was performed at 222 nm wavelength.

Barstar is not the only protein that exhibits cold unfolding. In fact, cold denaturation is a general phenomenon for all globular proteins (*24-27*). Cold unfolding results from specific and strongly temperature-dependent interaction of the non-polar groups of a protein with water. Given that cold denaturation of most proteins occurs well below the freezing point of water, a denaturant is sometimes added to shift the Gibbs free-energy function of protein to a higher temperature, allowing the cold denaturation to be completed and observed. The concentration of 1.2 M GdnCl is sufficiently low to maintain most of the barstar protein in its native state but creates a disturbance that makes a significant and measurable difference of the conformation at cold temperature vs. room temperature.

4.4.2 Peptide Mapping and Oxidative Site Determination

The global MS analysis of barstar folding dynamics was described in the previous chapter. To understand the structural details during folding, however, spatial, amino-acid residue resolution is required. There are four essential components in the experiment that identify the oxidation sites and quantify the modified products: complete proteolysis with good sequence coverage, base-line separation of the modified and unmodified peptides, high-resolving power to afford accurate mass measurement, and MS/MS. Although trypsin digestion affords a complete set of peptides covering the entire protein sequence, there is a 32-residue segment in the middle of the sequence that contains no lysine or arginine. Thus, trypsin digestion gives an intact peptide piece that is not MS favorable. Several unmodified peptides spanning this region were detected, none of which were of sufficient abundance to warrant looking for their modified siblings. To achieve better

analytical coverage, Glu-C digestion was performed separately to complement trypsin digestion, as shown in Figure 4.2.

The detection of the substitution of H by OH can be done with high certainty. Typical extracted ion chromatograms for a peptide and its modified siblings with a mass increase of 16 (15.9949) show convincingly the incorporation of an oxygen (Figure 4.3 (a) and (b)). Modified peptides with the same mass are separated in the chromatogram, quantified by using the integrated EIC peak area in the high resolving power MS domain, and identified by MS/MS (Figure 4.3 (c)). We described a residue-level comparison of barstar's two states in paper that demonstrates FPOP data analysis; this work is led by Brian Gau. The conclusion is that 19 residues were detected as modified, of which ten residues are significantly more labeled and hence more solvent-accessible in the cold state. With better sequence coverage achieved by complementary digestion, more modified residues were identified, discussed later in the chapter. Some residues, however, are insensitive to FPOP; thus, they cannot report any changes of solvent accessibility even though they may be important in the formation of the native structure of protein.

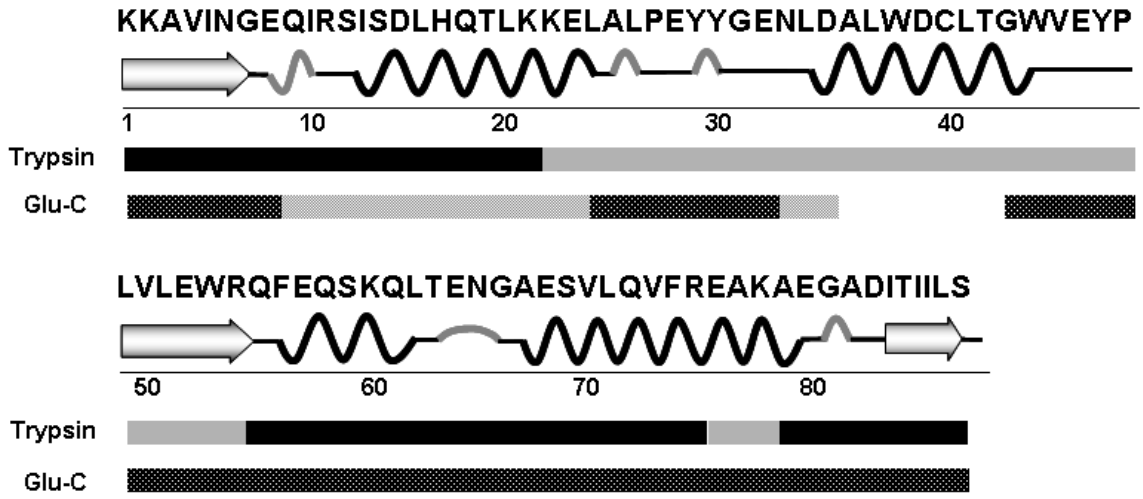


Figure 4.2: Sequence coverage map of barstar resulting from trypsin and Glu-C digestion. Black bars: sequence covered with peptides of adequate intensity for analysis; gray bars: sequence covered with peptides of poor intensity due to miscleavage or the lack of cleavable sites.

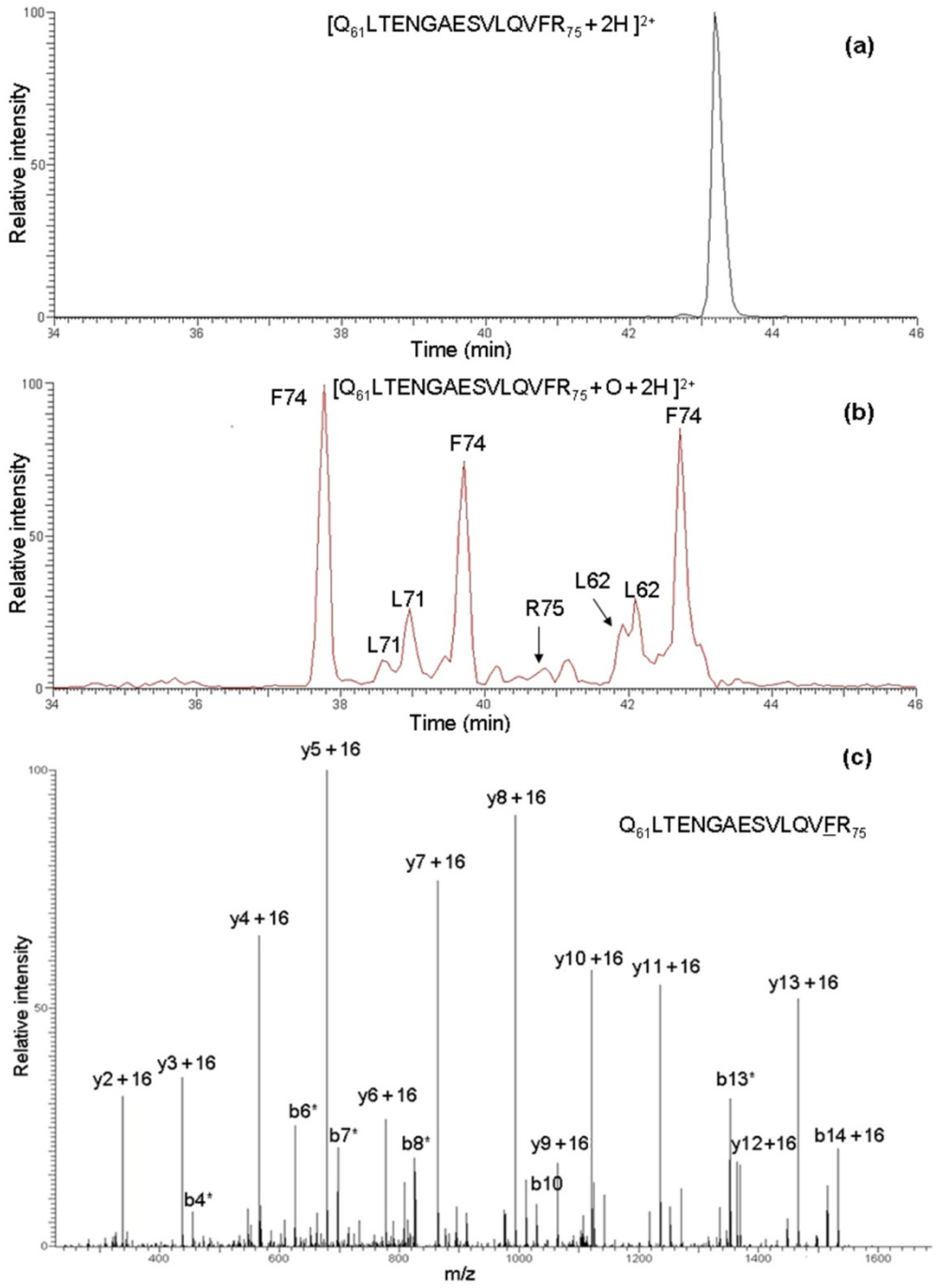


Figure 4.3: (a) EIC of an unmodified tryptic peptide of barstar. (b). EIC of oxygen-modified tryptic peptides. (c) Product-ion spectrum of a modified peptide eluting at 37.8 min in (b). The presence of y_2^{+16} and b_{14}^{+16} ions identifies F74 as the modified site.

4.4.3 Folding Dynamics at Amino-acid Level

The amino-acid residues having detectable modification can be grouped into four categories: (1) residues showing no/little difference of modification between native and denatured states, (2) residues modified to different extents between the native and the denatured states but show no/little change during early folding, (3) residues modified to a different extent for the native and the denatured states and showing significant changes during early folding, (4) residues partially protected in the intermediate state but showing no significant trend in oxidative modification during the early folding. Residues bearing relatively small percentages of modification are not included in the discussion because the error for the measurements exceeded the measured values.

We plot the percentage of modification of these four sets in Figures 4.4-7 and discuss each of them in the following paragraphs. Every plot has two end points at folding time 0 (cold, denatured) and infinity (room temperature, folded), colored green and red, respectively. Error bars are the standard deviations calculated from triplicate experiments.

Three example residues bearing either little or no difference of modification between native and denatured states (Figure 4.4) show the sensitivity of the method to regions of a protein that do not change significantly in folding. According to the NMR structure of barstar (PDB ID: 1BTA), E57 and L62 are located in $helix_3$, and R75 is located in $helix_4$, with their side chains exposed or partially exposed to the solvent. We conclude from the plots that these residues are not forming any hydrophobic interactions throughout folding, and are always exposed. In particular, L62, although a hydrophobic

residue and a candidate for interaction with other hydrophobic residues during folding, bears minimal change in solvent accessibility when the protein folds. The lack of change suggests that for small proteins like barstar, the folding process is straightforward. The residues exposed in the final folded stage do not participate in early folding.

Two example residues belonging to the second category are modified to different extent in native and denatured states but show no change of modification during early folding (Figure 4.5). Both W53 and L88 are buried in the hydrophobic core in the native state; thus, they are more protected when folded. In the process of forming the first intermediate state, however, these residues are not involved in hydrophobic interactions. They are likely to be incorporated in the hydrophobic core during the late folding stage when the secondary and tertiary structures consolidate.

Residues of the third category (Figure 4.6) are of particular interest because they are tightly buried in the hydrophobic core in the native state of protein, and show a significant trend of increasing protection during early folding. We propose that these residues play the key role in the formation of the folding intermediate by forming a fairly compact structure. As shown in the plots, the percentages of modification of these residues in the intermediate state are close to those for the final folded state, indicating that this region of protein is condensed to a high extent. It should be noted that these residues are all located in $helix_1$.

Residues in the fourth category (Figure 4.7) are also important for understanding the folding mechanism of barstar. These residues are away from $helix_1$, but they gain protection by the time the intermediate state is formed. These residues are probably

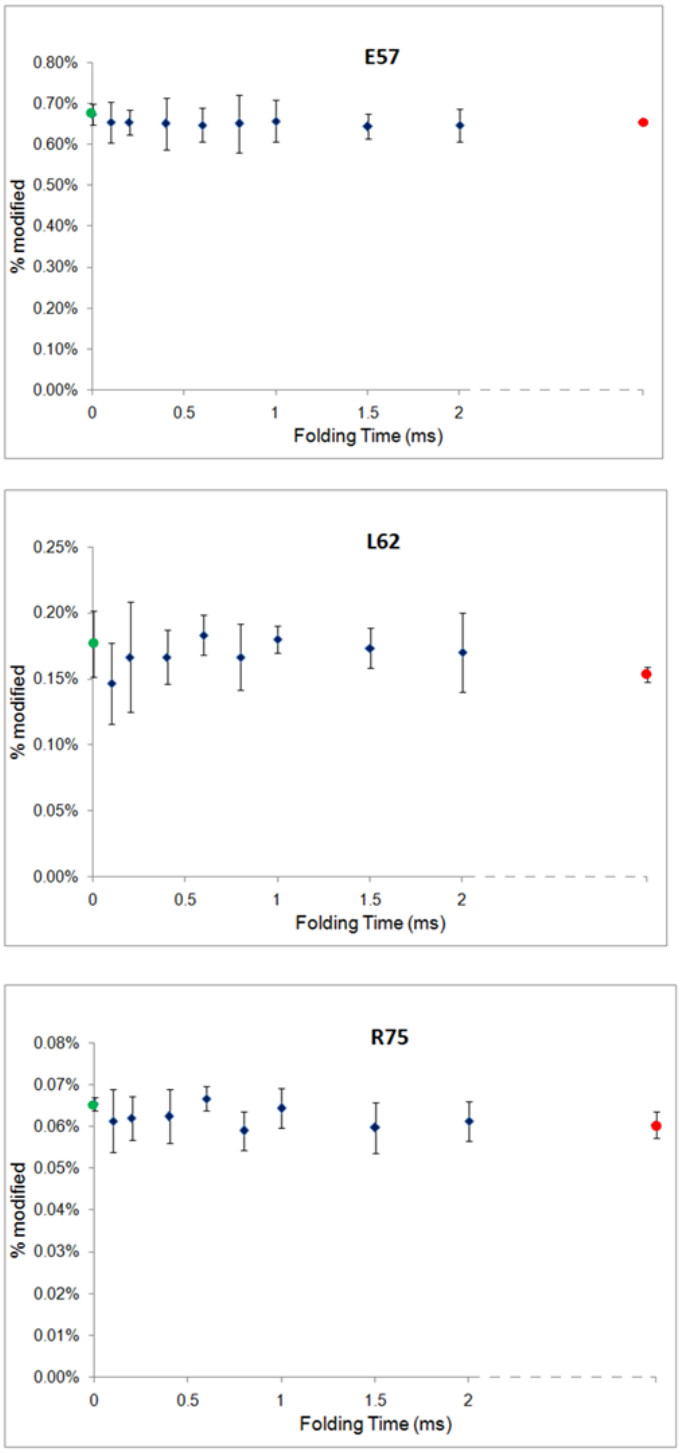


Figure 4.4: The time-dependent modification extents for three example residues that show no significant difference between the native and unfolded states. No trend of modification change throughout the early folding was observed.

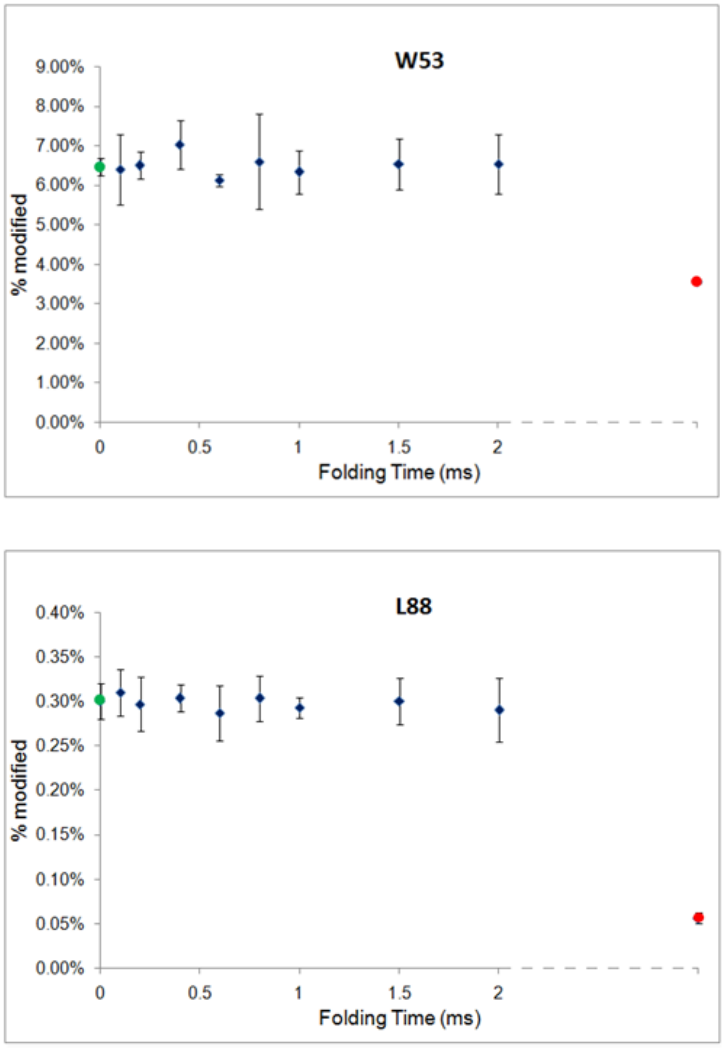


Figure 4.5: Examples of residues that show significant difference in oxidative modification between the native and unfolded states. No trend of modification change throughout the early folding, however, was observed.

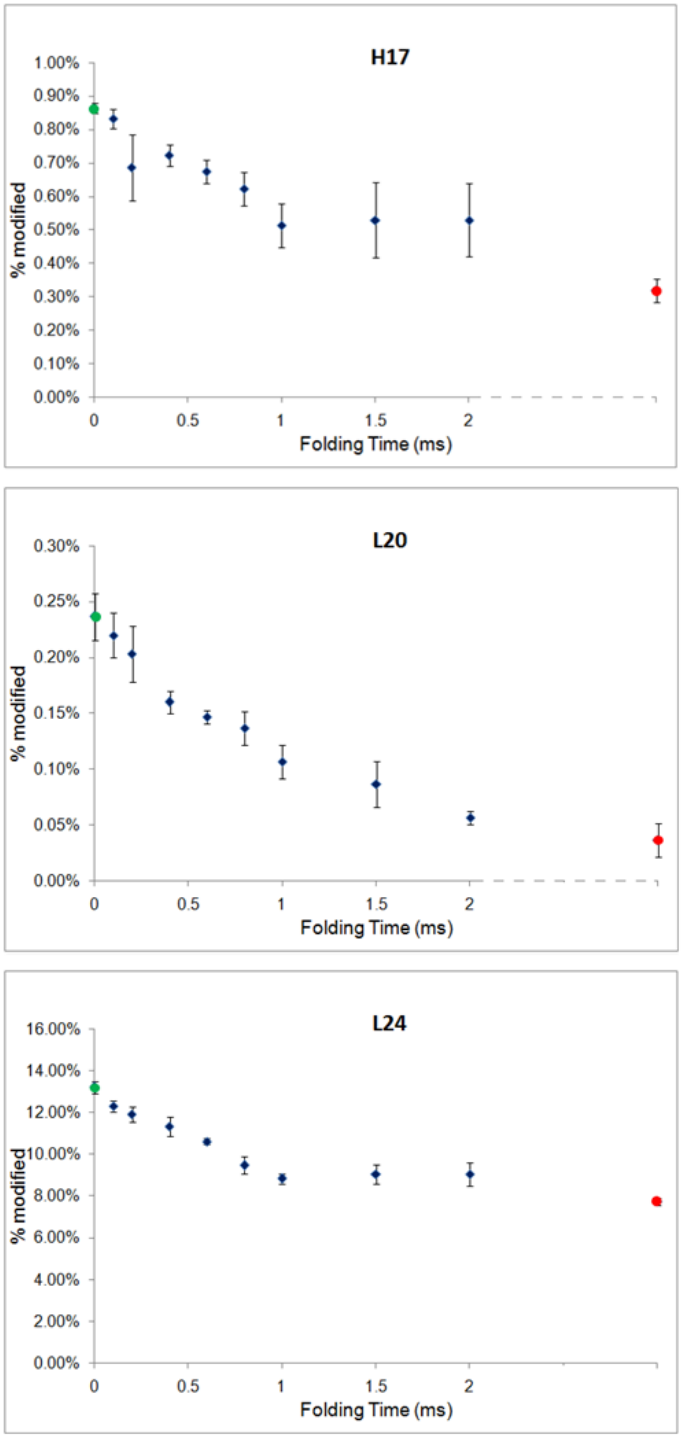


Figure 4.6: Example residues showing significant difference in modification extent between the native and unfolded states. A clear trend of modification decrease (increase in protection) is demonstrated for the early folding.

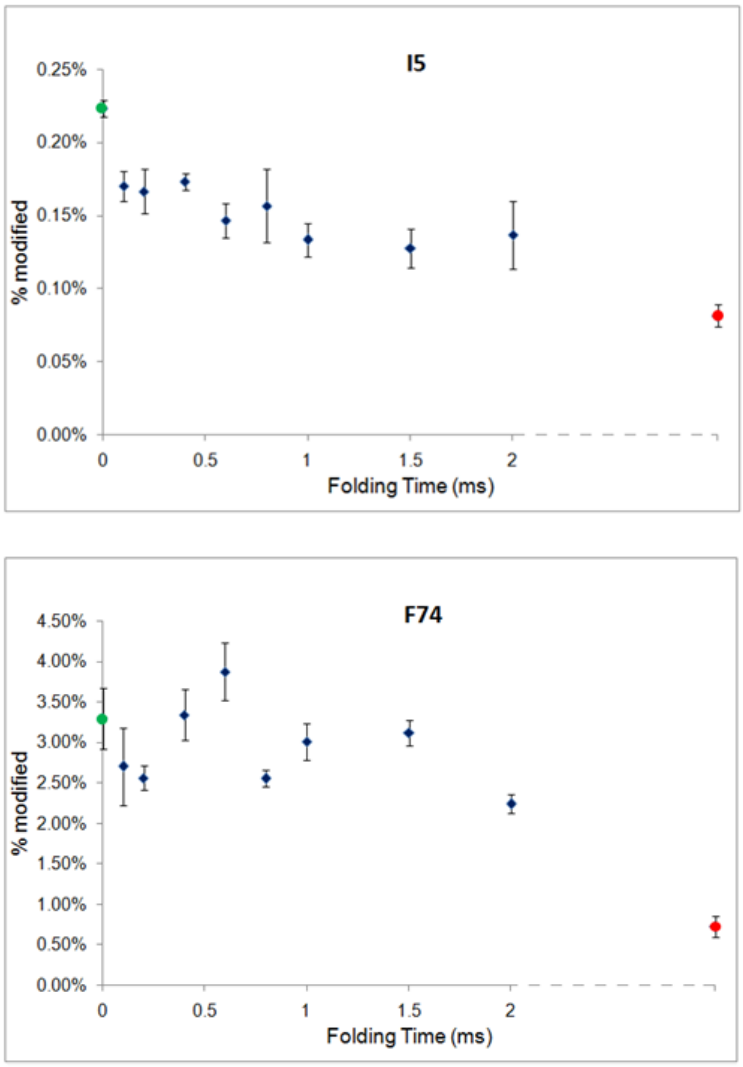


Figure 4.7: Example residues that show significant difference in oxidative modification extent between the native and unfolded states. Residues are partially protected in the intermediate state as shown by less modification than in the denatured state. The trend of the modification extent throughout the early folding was not obvious, especially for F74.

involved in the early folding; their involvement is via an interaction with the condensed structure around $helix_1$. The trend of percentage of modification is not obvious, especially for F74, suggesting that the interactions are rather weak and fluctuating in the first several hundred of μs.

It is worth noting that the error bars are generally larger for the kinetic measurements when the protein is folding than the steady-state measurements when the protein is equilibrated in either cold temperature or room temperature. We attribute the error to both experimental error and the nature of protein folding. In this experiment, we are observing an ensemble of conformations of protein, which is likely to be more diversified and uncertain while folding.

4.4.4 Characterization of the Intermediate State

The intermediate state of barstar folding can be represented by the structure formed 2 ms after initiation of folding. The percentage of modification of each residue, shown in Figures 4.4-7, cannot be used for cross comparison among all residues, owing to the fact that different residues have different reactivity towards hydroxyl radical and, thus, are modified to different extents even when they have the same solvent accessibility. Thus, we used a degree-of-folding value for each residue to evaluate and compare the folding extent. The degree-of-folding value of a residue in a particular state (*s*) is defined in the following equation:

$$degree\ of\ folding = \frac{(\%\ of\ modification\ in\ the\ denatured\ state) - (\%\ of\ modification\ in\ state\ s)}{(\%\ of\ modification\ in\ the\ denatured\ state) - (\%\ of\ modification\ in\ the\ native\ state)}$$

When this value equals 0, the residue is as exposed as in the denatured state. When the value equals 1, the residue is protected as in the native state. Intermediate values indicate

either that they are involved partially in formation of a folded structure or they constitute a mixture of structures with different degrees of folding.

Folding can be better presented in the degree-of-folding plot of three states (cold-denatured, intermediate and native states) for selected residues representing different segments of protein (Figure 4.8). All residues are valued 0 in the denatured state and 1 in the native state by definition. Among these residues, L20, a residue in $helix_1$, has the largest degree of folding (> 0.8) in the intermediate state, indicating the formation of a partially solvent-excluded core. I5 and F74, residues in β-$sheet_1$ and $helix_4$, have degrees of folding of 0.57 and 0.29, respectively. They are likely to be involved in weaker interactions with the hydrophobic core than L20. W53 and L88, residues in β-$sheet_2$ and β-$sheet_3$, have degrees of folding that are nearly 0, indicating that these residues in the intermediate state are as exposed as in the denatured state. All five residues (H17, L20, L24, I5, F74) are all involved in the hydrophobic core formation in the native state, as shown in Figure 4.9, but are folded to different extent in the intermediate state.

This result can be compared to the Φ values analyzed by Nolting et al. (*10*). The engineered proteins with the residues involved in the hydrophobic core that were substituted one-at-a-time in site-directed mutagenesis mutated are arranged in a decreasing order in of Φ values as follows: L16V, I5V, F56A, L51V, A67G and A77G. It was suggested that L16V with a high Φ probes mainly interactions in $helix_1$ and between $helix_1$ and $helix_4$. Our data are consistent with their results. We are not able to probe A67 and A77 by FPOP, however, owing to the low reactivity of alanine to hydroxyl radicals.

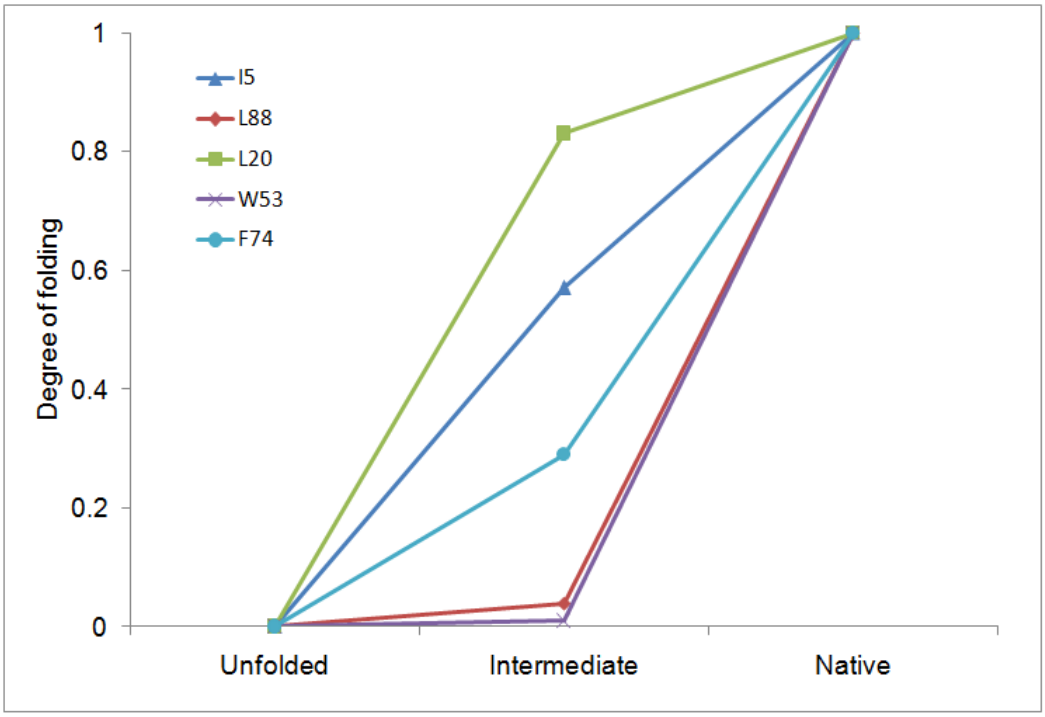


Figure 4.8: The degree-of-folding values for selected residues show the change from the cold-denatured, intermediate and native (final) states.

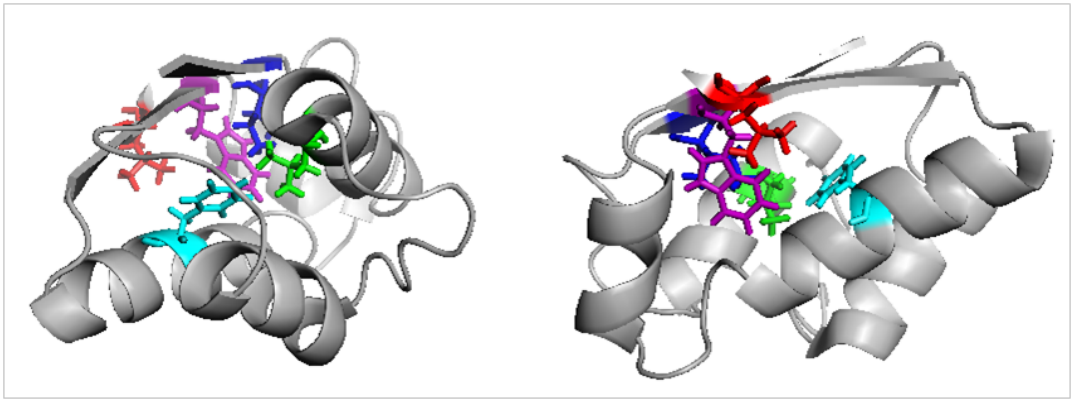


Figure 4.9: Two views of native barstar with the five residues identified as important in folding in our pump/probe experiment. The sidechains are colored coded: Blue: I5, Green: L20, Purple: W53, Cyan: F74, Red: L88.

4.4.5 Mechanism of Folding

The FPOP results on barstar folding are consistent with a general nucleation-condensation scheme that implicates a diffuse nucleus that consisting of some neighboring residues. According to our data, the nucleus is centered in $helix_1$, which contains several residues showing increasing protection during early folding. Some other residues, such as I5 and F74, may be involved in the formation of a folding intermediate but to a lesser extent. They are involved in long-range interactions with the nucleus, and they serve to stabilize the nucleus as it forms. There are also residues that are essential for the formation of the hydrophobic core in the native state but have no significant contact with the nucleus in the folding intermediate. The loose hydrophobic center with weak interactions is then further consolidated to become a well-established core during the second folding stage, which was not characterized in these experiments but can be predicted. The proposed model of barstar early folding is shown in Figure 4.10.

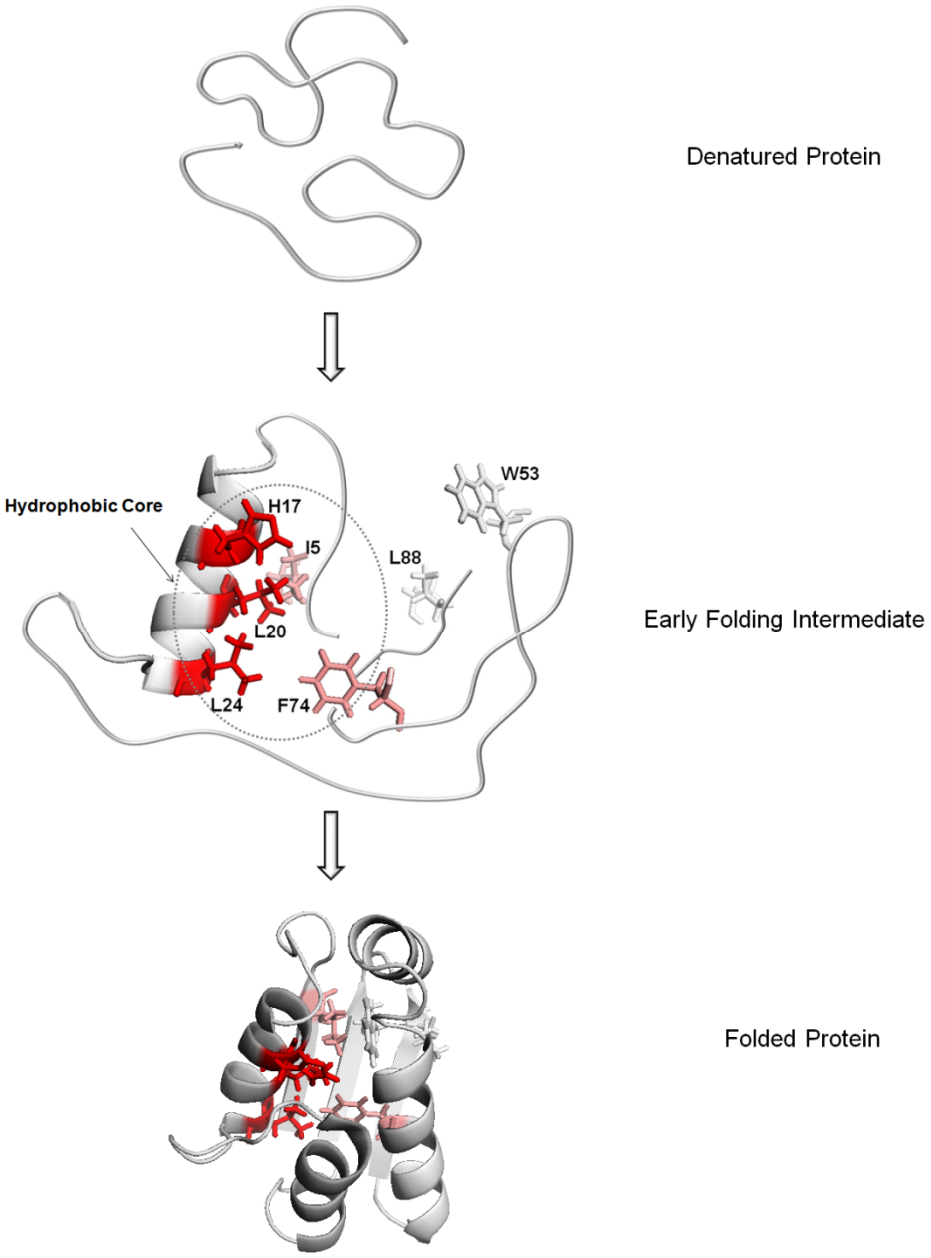


Figure 4.10: Proposed mechanism of barstar early folding. The folding starts with an ensemble of different conformations of the polypeptide chain, and the protein structure moves within 2 ms to an intermediate possessing moderate structure. Residues colored coded in red are closely associated with the hydrophobic center, whereas those color-coded pink weakly interact with the core. Residues colored gray do not participate in the formation of early folding intermediate.

4.5 Conclusion

The two-laser pump probe approach that we introduced allows us to study directly the mechanisms of protein folding. The approach provides insights into the conformational changes occuring at the amino-acid residue level during protein folding. In particular, a nucleus centered on the $helix_1$ region of barstar is formed to a significant extent in the intermediate state before further consolidation of structure occurs. Other residues are involved but via weaker interactions with the nucleus, and yet others are not involved at all in the formation of the nucleus. Owing to the nature of this experiment, we cannot conclude how much secondary structure is formed during the early folding. Although FPOP using a hydroxyl radical as the reactant is a non-selective labeling method, it does not modify all residues to a detectable extent. To obtain better footprinting coverage, other reagents might be used as complementary probes (*28, 29*).

The time frame of such experiment may be extended to faster processes using the current device or by using an ultra-fast mixing device, taking advantage of the short radical-exposure time (less than one µs) during FPOP. In such cases, this "pump-probe" combination could detect protein folding in early collapsed states that occur on the µs time scale (*30, 31*). This experiment, on the other hand, can also allow the study of slower process or the late stages of protein folding by displacing the probe laser from the pump. The extension of the time frame will be essential for obtaining a more complete picture of protein folding than was obtained here. This is the subject of ongoing research in our laboratory. The methodology presented here for structural analysis of transiently formed protein states is not restricted to folding intermediates but may also pertain to detecting transient states important for function.

4.6 References

1. Dill, K. A. (1985) Theory for the folding and stability of globular proteins, Biochemistry 24, 1501-1509.
2. Kim, P. S., and Baldwin, R. L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding, Annual review of biochemistry 51, 459-489.
3. Marianayagam, N. J., and Jackson, S. E. (2004) The folding pathway of ubiquitin from all-atom molecular dynamics simulations, Biophysical chemistry 111, 159-171.
4. Brylinski, M., Konieczny, L., and Roterman, I. (2006) Hydrophobic collapse in late-stage folding (in silico) of bovine pancreatic trypsin inhibitor, Biochimie 88, 1229-1239.
5. Yunger, J. (2007) A Study of Barstar Folding Events Using Boundary Value Simulations, Physica A 386, 791-798.
6. Gilmanshin, R., Dyer, R. B., and Callender, R. H. (1997) Structural heterogeneity of the various forms of apomyoglobin: implications for protein folding, Protein Sci 6, 2134-2142.
7. Vidugiris, G. J., Markley, J. L., and Royer, C. A. (1995) Evidence for a molten globule-like transition state in protein folding from determination of activation volumes, Biochemistry 34, 4909-4912.
8. Eigen, M. L. D. M. (1963) Relaxation Methods, in In Technique of Organic Chemistry (S. Freiss, E. L., A. Weissberger, Ed.), pp 895-1054, Interscience Publishing Inc., New York.

9. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) Direct observation of fast protein folding: the initial collapse of apomyoglobin, Proceedings of the National Academy of Sciences of the United States of America 93, 5759-5764.

10. Nolting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) The folding pathway of a protein at high resolution from microseconds to seconds, Proceedings of the National Academy of Sciences of the United States of America 94, 826-830.

11. Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) Mapping the transition state and pathway of protein folding by protein engineering, Nature 340, 122-126.

12. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, Journal of molecular biology 224, 771-782.

13. Fersht, A. R., and Sato, S. (2004) Phi-value analysis and the nature of protein-folding transition states, Proceedings of the National Academy of Sciences of the United States of America 101, 7976-7981.

14. Cho, J. H., O'Connell, N., Raleigh, D. P., and Palmer, A. G., 3rd. (2010) Phi-value analysis for ultrafast folding proteins by NMR relaxation dispersion, Journal of the American Chemical Society 132, 450-451.

15. Guillet, V., Lapthorn, A., Fourniat, J., Benoit, J. P., Hartley, R. W., and Mauguen, Y. (1993) Crystallization and preliminary X-ray investigation of barstar, the intracellular inhibitor of barnase, Proteins 17, 325-328.

16. Lubienski, M. J., Bycroft, M., Freund, S. M., and Fersht, A. R. (1994) Three-dimensional solution structure and 13C assignments of barstar using nuclear magnetic resonance spectroscopy, Biochemistry 33, 8866-8877.

17. Khurana, R., and Udgaonkar, J. B. (1994) Equilibrium unfolding studies of barstar: evidence for an alternative conformation which resembles a molten globule, Biochemistry 33, 106-115.

18. Schreiber, G., and Fersht, A. R. (1993) The refolding of cis- and trans-peptidylprolyl isomers of barstar, Biochemistry 32, 11195-11203.

19. Shastry, M. C., and Udgaonkar, J. B. (1995) The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates, Journal of molecular biology 247, 1013-1027.

20. Agashe, V. R., Shastry, M. C., and Udgaonkar, J. B. (1995) Initial hydrophobic collapse in the folding of barstar, Nature 377, 754-757.

21. Takamoto, K., and Chance, M. R. (2006) Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes, Annual review of biophysics and biomolecular structure 35, 251-276.

22. Stocks, B. B., and Konermann, L. (2010) Time-dependent changes in side-chain solvent accessibility during cytochrome c folding probed by pulsed oxidative labeling and mass spectrometry, Journal of molecular biology 398, 362-373.

23. Chen, J., Rempel, D. L., and Gross, M. L. (2010) Temperature jump and fast photochemical oxidation probe submillisecond protein folding, Journal of the American Chemical Society 132, 15502-15504.

24. Privalov, P. L., Griko Yu, V., Venyaminov, S., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin, Journal of molecular biology 190, 487-498.

25. Griko, Y. V., Privalov, P. L., Sturtevant, J. M., and Venyaminov, S. (1988) Cold denaturation of staphylococcal nuclease, Proceedings of the National Academy of Sciences of the United States of America 85, 3343-3347.

26. Privalov, P. L. (1990) Cold denaturation of proteins, Critical reviews in biochemistry and molecular biology 25, 281-305.

27. Pastore, A., Martin, S. R., Politou, A., Kondapalli, K. C., Stemmler, T., and Temussi, P. A. (2007) Unbiased cold denaturation: low- and high-temperature unfolding of yeast frataxin under physiological conditions, Journal of the American Chemical Society 129, 5374-5375.

28. Gau, B. C., Chen, H., Zhang, Y., and Gross, M. L. Sulfate radical anion as a new reagent for fast photochemical oxidation of proteins, Analytical chemistry 82, 7821-7827.

29. Gomez, G. E., Cauerhff, A., Craig, P. O., Goldbaum, F. A., and Delfino, J. M. (2006) Exploring protein interfaces with a general photochemical reagent, Protein Sci 15, 744-752.

30. Arai, M., Kondrashkina, E., Kayatekin, C., Matthews, C. R., Iwakura, M., and Bilsel, O. (2007) Microsecond hydrophobic collapse in the folding of Escherichia coli dihydrofolate reductase, an alpha/beta-type protein, Journal of molecular biology 368, 219-229.

31. Lapidus, L. J., Yao, S., McGarrity, K. S., Hertzog, D. E., Tubman, E., and Bakajin, O. (2007) Protein hydrophobic collapse and early folding steps observed in a microfluidic mixer, Biophysical journal 93, 218-224.

5. *Fast Photochemical Iodination Footprinting Combined with Top-down and Bottom-up Mass Spectrometric Analysis Reveal Protein Conformation

*This chapter is based on a recent manuscript:

Chen, J., Cui, W., Giblin, D. & Gross,M. L. Fast Photochemical Iodination Footprinting Combined with Top-down and Bottom-up Mass Spectrometric Analysis Reveal Protein Conformation.

5.1 Abstract

We report a new reagent for the FPOP (fast photochemical oxidation of proteins)-based protein footprinting method whereby iodine species are the modifying reagent. We generate the radicals by photolysis of iodobenzoic acid at 248 nm; the radical then rapidly modifies the target protein. This iodine-radical labeling is sensitive, tunable and site-specific, modifying only histidine and tyrosine residues. In this work, myoglobin (Mb) and apomyoglobin (aMb) were iodinated in their native states and analyzed by both top-down and bottom-up proteomic strategies. Top-down sequencing selects a certain level (addition of one I, two I's) of modification and determines the major components produced in the modification reaction, whereas bottom-up reveals details for each modification site. Tyr146 is found to be modified for aMb but less so for Mb. His82, His93 and His97 are at least 10 times more modified for aMb than for Mb. For carbonic anhydrase and its apo form, there are no significant differences of the modification extents, indicating their similarity in conformation. For lispro, insulin-EDTA and insulin-zinc, different iodination yields are sensitive to differences in insulin oligomerization state. The iodine radical labeling is a promising addition to protein footprinting methods, offering higher specificity than ·OH and $SO_4^{-\cdot}$, two other radicals already employed in FPOP.

5.2 Introduction

Mass spectrometry (MS) approaches using chemical probes appear to be useful for characterizing protein structure. Typical chemical probes include hydrogen-deuterium (H/D) exchange (*1, 2*), hydroxyl radical footprinting (*3-6*), specific chemical modifications of one or two residues (*7-9*) and cross-linking (*10*). As the most widely used probe for protein structure, H/D exchange interrogates the backbone amide hydrogen accessibility, whereas the more recently developed hydroxyl-radical oxidation probes the solvent accessibility of various but not all side chains.

Described in chapter 1, 2 and 3, we developed the FPOP method so that exposed amino-acid residues can be irreversibly and covalently labeled by oxidation with hydroxyl radicals produced by photolysis of hydrogen peroxide (*4, 11*). The modified residues can be detected and quantified by an approach using trypsin proteolysis followed by LC/MS/MS analysis (a standard approach in proteomics). Fourteen of 20 amino acids can be modified by hydroxyl radicals for the purpose of footprinting, thus potentially covering ~65% of the sequence of a typical protein (*12, 13*). Given that FPOP occurs on a time scale faster than most protein conformational changes (*14*), it is suitable for both steady-state studies and time-resolved investigations (*15, 16*). Recently, we reported that the sulfate radical anion can be used as a new FPOP reagent (*17*). This development indicates that the scope of FPOP can be increased by using other reactive species including radicals and possibly carbenes.

In this chapter, we describe a new FPOP-based footprinting method using the iodine species produced by laser photolysis to label proteins in a more specific way than OH radicals. Iodine atoms can be incorporated enzymatically or non-enzymatically

mainly into tyrosine residues of proteins, forming iodotyrosines (*18-20*). Some histidines can also be iodinated to an extent depending on the experimental conditions and the structure of the protein (*21*). The iodination strategy was previously used for labeling proteins in a wide variety of experiments, such as protein-metabolism studies, radioimmunoassay (*18*), and protein-structure determination (*22-25*). Recently, Ly and Julian (*26*) utilized the photodissociation of the carbon-iodine bond in iodo-tyrosine-containing protein to generate a residue-specific radical site for protein cleavage in gas phase. Although numerous studies focus on protein iodination, none have applied FPOP or any equivalent fast labeling via photo dissociation of iodobenzoic acid. Given that the iodine is the least reactive halogen, we anticipate that, unlike OH radicals, a more specific modification platform will result. In contrast, chlorine and bromine species result from photolysis react with water promptly to produce hydroxyl radicals.

One traditional MS strategy of locating modification sites and quantifying the fraction of modified sites is bottom-up sequencing (*27*). In this approach, both the modified and unmodified proteins are proteolyzed in separate experiments. A liquid-phase chromatographic separation followed by MS analysis locates the sites of modification. This approach provides reliable identification of the protein and its modified "siblings" but suffers the risk of peptide losses during digestion and analysis, thus possibly forfeiting vital information.

More recently, top-down sequencing has been developed to characterize the intact protein and determine its post-translational modifications (PTMs) (*28-30*). With recently developed MS/MS techniques of electron-transfer dissociation (ETD) (*31*) and electron-capture dissociation (ECD) (*32*), the top-down approach has become more efficient and

gives better sequence coverage and more tolerance towards labile PTMs. Moreover, top-down sequencing informs us whether a modified amino acid has been produced for various labeled states of the protein (e.g., for singly modified, doubly modified. . .) Given the more specific reactivity of iodine radicals and their higher mass compared to ·OH, proteins labeled with ·I are more suitable for top-down analysis than are ·OH-modified proteins.

In this work, we demonstrate that FPOP-based iodination footprints proteins by modifying solvent-accessible histidine and tyrosine residues. We located the modifications and quantified them by both a top-down approach, which isolates and characterizes the intact mono-iodinated protein, for example, and a bottom-up approach, which identifies the iodinated residues after sample footprinting and quantifies each of them. To our knowledge, this is the first example of iodination being used on the FPOP platform.

5.3 Material and Methods

5.3.1 Chemicals and Proteins

Equine skeletal myoglobin, apo-myoglobin, carbonic anhydrase II from bovine erythrocytes, human insulin, iodobenzoic acid, proteomics grade trypsin, ethylene diamine tetraacetic acid (EDTA), zinc chloride, phosphate-buffered saline (PBS, pH = 7.4), trifluoroacetic acid, 1,10-phenanthroline, and ammonium bicarbonate were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO). Lispro was donated by Dr. Stephen Bayne of NovoNordisk, Denmark. Acetonitrile was obtained from Honeywell Burdick and Jackson (Muskegon, WI). Purified water (18 MΩ) was obtained from an in-house Milli-Q Synthesis system (Millipore, Billerica, MA). All chemicals

were used without further purification. Apo-carbonic anhydrase was prepared by using dialysis of the native protein against 0.1 M sodium acetate, pH 5.0, containing 2×10^{-3} M 1,10-phenanthroline at 4 °C for 7 d followed by the removal of the chelating agent by buffer exchange against PBS buffer (*33*). The apo-state was confirmed by native MS analysis. Insulin-EDTA and insulin-zinc stock solutions were prepared by incubating 0.1 mM insulin with 0.1 mM EDTA and 0.05 mM zinc chloride, respectively. Depletion of zinc ions by EDTA likely caused dissociation of insulin hexamers to dimers, and the addition of zinc ions enhanced the formation of hexamers. The concentrations of all the protein stock solutions were determined by UV absorption (1.55×10^{4} $M^{-1}cm^{-1}$ for myoglobin and apo-myoglobin, 5.22×10^{4} $M^{-1}cm^{-1}$ for carbonic anhydrase, 5.7×10^{3} $M^{-1}cm^{-1}$ for insulin and lispro, all at 280 nm) (*34*).

5.3.2 Photochemical Protein Iodination

Each 50 µL sample was prepared in PBS (10 mM phosphate buffer, 138 mM NaCl, 2.7 mM KCl, pH 7.4 at 25 °C), with a final protein concentration of 10 µM. Histidine was chosen as the radical scavenger, and it was added to a final concentration of 0.5 mM. An iodobenzoic acid solution was added to a final concentration of 50 µM, just prior to FPOP. The experimental setup and procedure for photolysis were described previously (*4*). To ensure that the protein was not "double-shot" (*4*, *14*), laser frequency and sample flow rate was adjusted to exclude 15% of the flowing protein solution from irradiation. The sample solution following irradiation was collected in a centrifuge tube and purified by C18 Ziptip™ (Millipore Corporation, Billerica, MA) solid-phase extraction. Proteins were eluted from the Ziptip with 10 µL 50% acetonitrile, 0.1% trifluoroacetic acid. Each sample interrogation was conducted in triplicate. Control

samples of each set were treated following the same procedure except no laser irradiation was employed.

5.3.3 Electron-Capture Dissociation (ECD) of Iodinated Mb and aMb

Global analysis and ECD of iodinated Mb and aMb were performed in the positive-ion mode of a Bruker SolariX^{TM} 12 T FT-ICR mass spectrometer (Bruker Daltonics, Bremen, Germany). Direct infusion of protein samples after purification was carried out by pumping at flow rate of 100 nL/min with a Harvard Apparatus syringe pump (Instech Laboratories, Inc., Plymouth Meeting, PA). The samples, at an estimated final concentration of 5 µM, were sprayed at 800 V into the mass spectrometer via a nanospray emitter (360 µm o.d., 150 µm i.d., Polymicro Technologies, Phoenix, AZ), which was pulled in-house on Polymicro silicon tubing by a P2000 laser puller (Sutter Instrument Co., Novato, CA). Mass spectra were recorded at 60,000 mass resolving power (at m/z = 1000) with ion accumulation times of 0.001 s in the ion source, 0.5 s in the collision/accumulation cell for full spectra, and 2 s in the collision/accumulation cell for ECD of isolated ions. The argon pressure in the collision/accumulation cell was 5.8×10^{-6} mbar; the pressure of the ICR trap region held within the core of the superconducting magnet was 1.4×10^{-9} mbar. The trapping voltages of the infinity ICR cell were 0.7 V on the front plate and 0.8 V on the back plate. Excitation was 25% of full power. To perform ECD on monoiodinated species of both aMb and Mb, an ion of m/z 1068 corresponding to $[*M + 16H]^{+16}$, was isolated with a window of 7 Da by the front-end quadrupole. ECD parameters were 1.6 A to heat the hollow cathode dispenser, 10 V on the grid, - 0.6 V for electron extraction, and 60 ms of extraction length, with slight adjustments of the latter two to optimize the ECD fragmentation of each protein. Up to

200 scans were averaged with 1 M data points for each scan. Ubiquitin was used for calibration.

5.3.4 Proteolysis and LC/MS/MS Analysis

Triplicate protein samples were dried under vacuum in a SpeedVac and then dissolved in 100 mM ammonium bicarbonate buffer. Trypsin solution (0.4 µL) was prepared according to a protocol provided by the manufacturer and added to each sample. Samples were incubated at 37 °C for 8 h. An aliquot (5 µL) of sample was loaded onto a custom-built silica capillary column packed with C18 reverse-phase material (Magic, 5 µm, 300 Å, Michrom, Auburn, CA). The HPLC gradient was from 2% solvent B (97% acetonitrile, 3% water, 0.1% formic acid) and 98% solvent A (97% water, 3% acetonitrile, 0.1% formic acid) to 50% solvent B over 50 min, then to 85% solvent B for 5 min at a flow rate of 260 nL/min followed by a 5 min re-equilibration step. The solution was sprayed directly from the column into the LTQ-Orbitrap mass spectrometer (Thermo Fisher, Waltham, MA) by using a PicoView PV-500 nanospray source (New Objective, Woburn, MA). A full mass spectrum of eluting peptides was recorded at high mass resolving power (100,000 for ions of *m/z* 400) while MS/MS experiments on the six most abundant ions from the eluent were performed in the LTQ at a normalized collision energy of 35% of the maximum, with a 2 Da isolation width and wide-band activation. Ions submitted to MS/MS were placed in a dynamic exclusion list for 8 s.

5.3.5 Data Analysis

The protein mass spectra from top-down sequencing were analyzed with the Bruker DataAnalysis tool. ECD fragments were assigned with the help of Bruker

BioTool™ 3.2 software. Quantification was done by dividing the sum of intensities of the peaks corresponding to modified and unmodified protein, assuming comparable charging and ionization efficiencies. For the LC/MS/MS data analysis, a general Mascot search and a sequential error-tolerant search as the first step were necessary not only to evaluate sequence coverage but also to provide valuable information of the possible modified sites. Then a list was created manually for each protein; the list contained the *m/z* of all charge states of the identified peptides, both unmodified and modified. Peptides were validated by manual inspection of their accurate mass, retention time, and product-ion (MS/MS) spectra. Ion chromatograms were extracted according to the peptide list, and the peak areas were measured. Modification yields for all the associated peptides were expressed as the ratio of the modified peptide to the sum of both modified and unmodified peptide.

5.3.6 Circular Dichroism Experiments

CD spectra were obtained at 4 °C using a J-815 CD spectrometer (JASCO Analytical Instruments, Tokyo, Japan). All experiments were done after placing the protein in a PBS buffer. Spectra were measured with a sample cell having a 0.1 cm path length, at a scan resolution of 0.2 nm, and a scan rate of 50 nm/min.

5.3.7 Theoretical Calculations

Theoretical calculations were performed to characterize the potential-energy surface (PES) associated with halide radical attachment and fixation. As surrogate for tyrosine, we chose *p*-cresol which retains the same functionality and requires less computational overhead. Structures of precursors and intermediates were explored by

MMFF molecular molecular mechanics, and associated transition states were explored by using the PM3 semi-empirical (*35, 36*) algorithm (Spartan for Linux v. 02: Wavefunction, Inc.). Minima and transition states were optimized by DFT (Density Function Theory: B3LYP or PBE0 [designated PBE1PBE]) or MP2 (Moeller-Plesset), as in the Gaussian 98/03 suites (*37*)(*38*) (Gaussian Inc.) by using the LAND2DZ basis set and confirmed by vibrational frequency analysis. In addition, connection of transition states to minima was examined by path calculations. Calculated energies were scaled thermal-energy corrections were applied and reported in kJ/mol as enthalpies of formation relative to a selected, suitable precursor (*39*). Theoretical calculations were done by Dr. Daryl Giblin.

5.4 Results and Discussion

5.4.1 Iodination as a Footprinting Method

The photodissociation mechanism of iodobenzoic acid is expected to be similar to aryl halides such as iodobenzene (*40*). The cleavage of the C-I bond in aryl iodides is an indirect process consisting of an excitation in the range 250 ± 30 nm promoting the molecule to the (π, π^*) state and an internal conversion to the repulsive (n, σ^*) state before the C-I bond breaks rapidly (*40-42*). In the absence of other reactive species, the iodine radicals and benzoic-acid radicals then undergo recombination reactions (*43*).

Although the amino-acid residues of a protein react productively with the iodine species generated by the laser photolysis, this reaction has not yet been well studied. We are able to demonstrate that both histidine and tyrosine can be modified, depending on the solvent environment of these residues and the ratio of the protein concentration to that

of the iodine species that were formed by laser photolysis. Theoretical calculations using different method (see Experimental) all indicate that iodine radical itself does not attach covalently to the π-electron system of an aromatic ring. It has been established, however, that an iodine radical can form a loose complex with the π system of the aromatic ring with carbon-iodine distance greater than 3 angstroms. Oxygen attacking an adjacent carbon induces the attachment of the iodine radical to the ring. Subsequent elimination of the $HO_2\cdot$ radical then leaves the iodine substituted on the ring, in analogy to the proposed mechanism for hydroxyl radical substitution on an aromatic ring (*3*).

The main products of traditional iodination of histidine are 2-iodohistidine, 2,5-diiodohistidine and 1,2,5-triiodohistidine, whereas 3,5-diiodotyrosine is the major product for tyrosine iodination (*44*). Previous studies suggested that the rate of iodination of histidine is 30-100 times slower than that of tyrosine (*44*). To avoid overlabeling and inducing conformational changes in the protein, free histidine amino acid was added to the protein solution as a scavenger to control radical exposure time. The basis for this approach is the capability of histidine to react with iodine radicals at moderate rate and yet not completely quench the radicals before they react with the protein. We monitored the amount of scavenger *vs.* the level of modification and found an appropriate concentration of 0.5 mM for the experiments.

Like other radicals used in FPOP, the radical life time should be short in the presence of the scavenger such that the protein is modified before structural changes are induced by overlabeling. Under controlled experimental conditions, no modified histidine other than mono-iodohistidine, was obtained whereas both mono-iodotyrosine and di-iodotyrosine were detected by MS. CD spectra of proteins before and after

iodination are nearly identical (see Figure 5.1). Therefore, iodination under our experimental condition does not alter the proteins' secondary structure.

5.4.2 Global Mass Spectrometric Analysis

Upon modification by reaction with I species, the mass of the protein increases in integral multiples of 125.90 Da corresponding to [M + n(I - H)] (n = 1,2… where n = 0 corresponds to no modification); the single modification (n = 1) is the most prevalent in the case of Mb and aMb. Mb is a 153 amino-acid protein existing in a highly folded and compact structure with eight alpha helical secondary structures that wrap around a central pocket containing the heme group. AMb, lacking the heme group between helices E and F, is the intermediate in biosynthesis of Mb. Structural characterization by NMR suggests that although the majority of the aMb polypeptide chain adopts a well-defined structure, some of the structural elements including the EF loop, the F helix, the FG loop, the first few residues of the G helix, and the C-terminal end of the H helix fluctuate between a native conformation and less structured conformations (*45*). Limited proteolysis experiments also demonstrated that helix F is highly flexible or largely disrupted in aMb (*46*).

Global MS measurements of modified Mb and aMb that were iodinated under the same experimental conditions (Figure 5.2) clearly show a difference in the extent of iodination. Mono-iodinated species are present at 38% of the unmodified species for Mb (Figure 5.2 (a)), but they increase to 130% in aMb (Figure 5.2 (b)). Furthermore, there is no detectable tri-iodinated Mb, but tri-iodinated aMb is clearly visible, consistent with the premise that heme removal leads to a looser protein conformation that has an increased number of solvent-exposed histidine and tyrosine residues. It appears that iodination is a

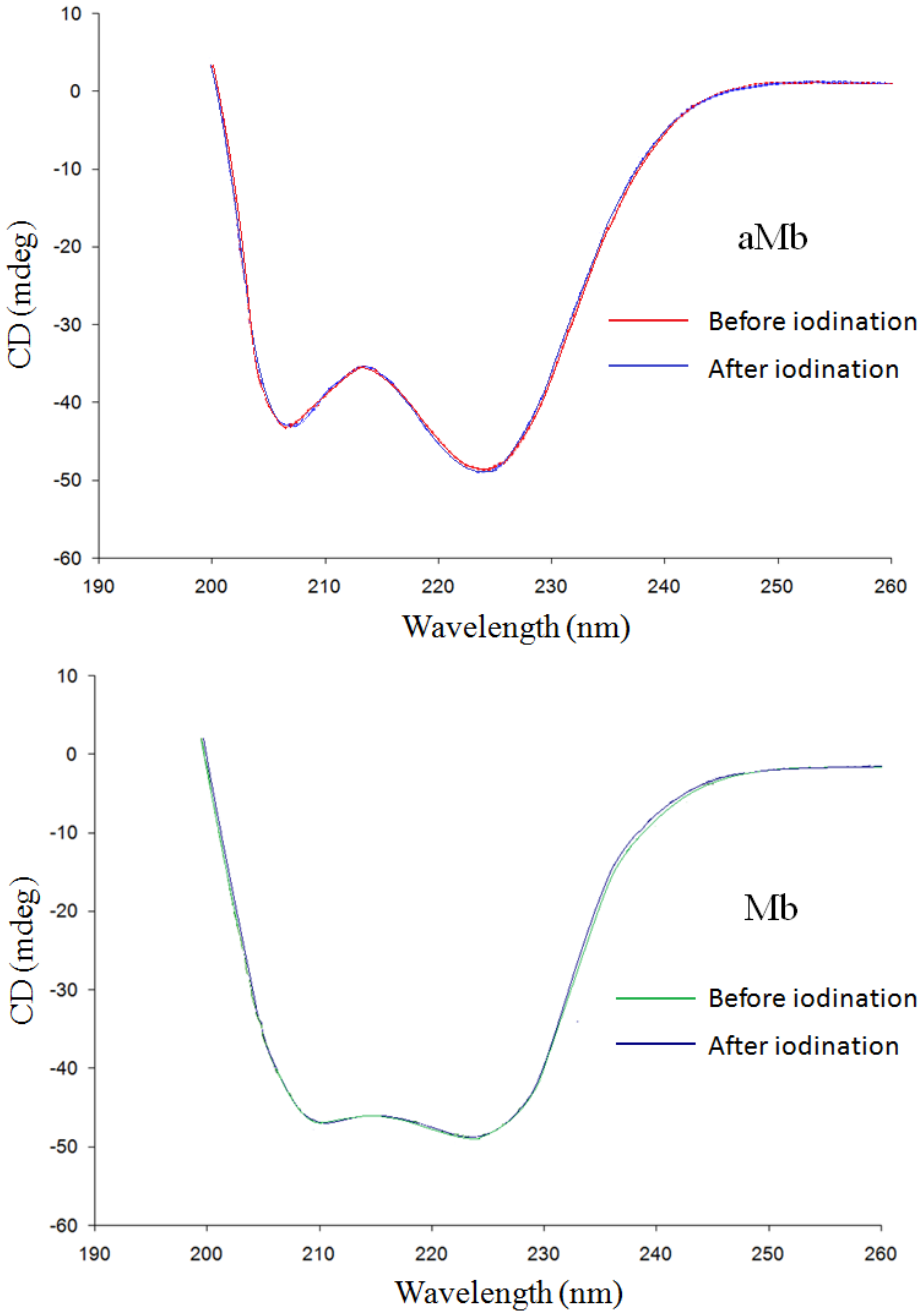


Figure 5.1: CD spectra of aMb and Mb before and after iodination. No significant secondary structure change occurs upon iodination of the protein.

very sensitive probe for structural change when combined with MS analysis; results for the whole protein and for individual residues should be informative.

5.4.3 Top-down Analysis

There are precedents for top-down analysis of iodinated proteins. Inieta et al. (*47*) reported the first top-down experiment to locate a single iodinated tyrosine of myoglobin; the modification resulted from electrosynthesis. Recently, Sun et al. (*25*, *48*) applied radical direct dissociation (RDD) and ECD to locate sites of iodination in several proteins.

Top-down sequencing may be conducted either by selecting all the states of protein, unmodified and modified, or by isolating one particular modification state (e.g., mono iodinated). To identify the location(s) of the site(s) of modification and to avoid interference of the unmodified protein, we chose to isolate the protein possessing one modified site. In this experiment, $[*\text{Mb} + 16\text{H}]^{+16}$ of *m/z* 1068, representing a singly-modified protein, when isolated and subjected to ECD, gave a series of *c* and *z* ions for identification and quantification (Figure 5.3). Given that the proteins either without modification or with more than one modification were excluded prior to ECD, we should be able to determine the distribution of iodine in the mono-iodinated species. Consider fragmentation at a site that can yield a *c* (*z*) ion containing only one modifiable site (i.e., a histidine or tyrosine). If the histidine or tyrosine is fractionally modified, two *c* ions will form, one carrying the modifications and one not. The fractional modification at that site is the yield of the modified *c* (*z)* ion divided by the sum of the modified and unmodified.

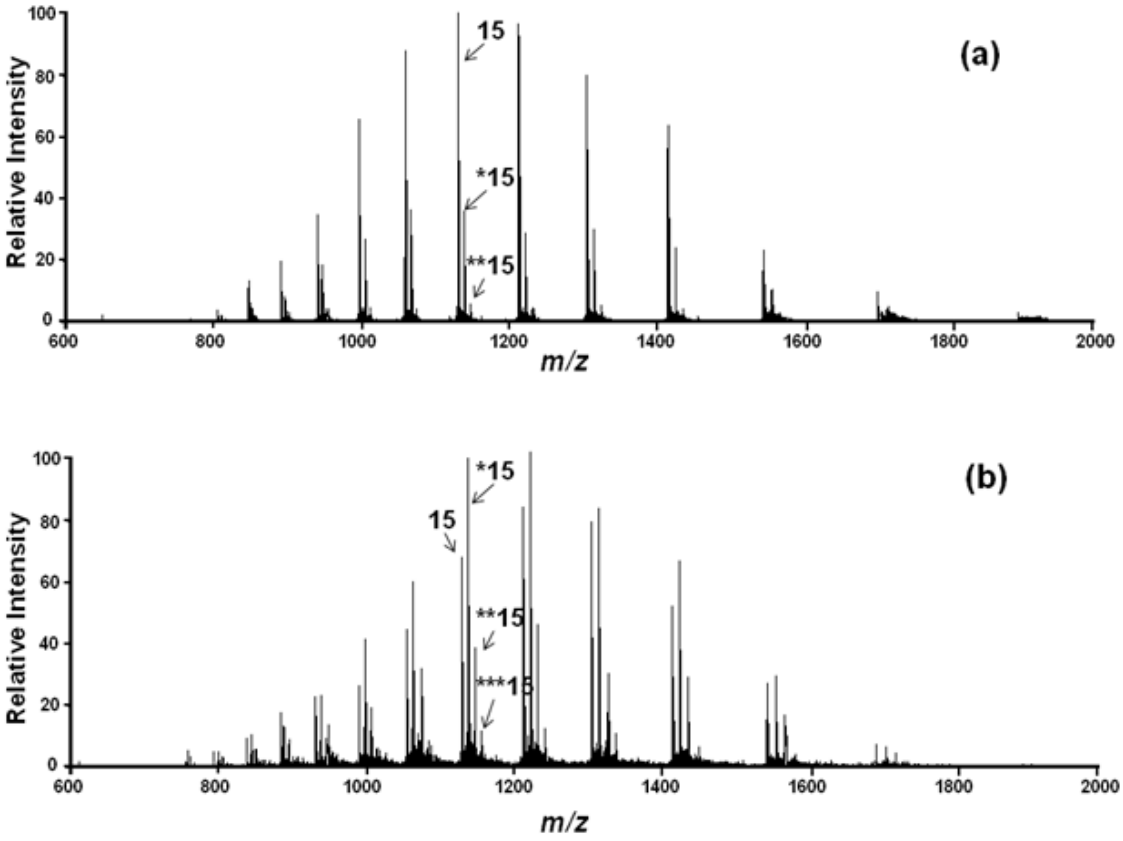


Figure 5.2: Full ESI mass spectra of (a) of iodinated myoglobin (Mb), and (b) of iodinated apomyoglobin (aMb). The number of iodines incorporated in the protein to give mono-, di- and tri-iodinated species are indicated for the 15^{th} charge state by the number of stars.

For aMb, we measured the yield of iodinated z_8^+ ion to be 26% of the total of modified and unmodified z_8^+, indicating that among all the mono-iodinated species, 26% of them are iodinated at Y146 because this amino-acid residue is the only one that can be modified in the region represented by the fragment. There is no modified z_8^+ for Mb, however, indicating that Y146 is now protected from iodination. Determining that no modification occurred in a region of the protein is even simpler. For example, there is no observable modified c_{59}^{4+} ion for both aMb and Mb, indicating that H24, H36 and H48, the three modifiable sites in this fragment, are highly protected and do not react with I species.

Conclusions become more difficult for larger ions that contain more than one modifiable site. Consider two *z* ions, the smaller containing x modifiable sites and the larger containing x + 1 modifiable sites. This is the case for z_{44}^{5+} and z_{55}^{5+}. The former contains Y146 and the latter Y146 and Y103. The percent modified z_{44}^{5+} and z_{55}^{5+} are 28% and 52%, respectively, for aMb, allowing us to surmise that 24% (52 – 28%) of the mono-iodinated proteins are iodinated at Y103 because it is the only modifiable site in the sequence between z_{44}^{5+} and z_{55}^{5+}. For Mb, the percentages are lower at 12% and 40%, but the difference in Y103 iodination is 28%, slightly larger than that for aMb. Thus, Y103 has comparable solvent accessibility for both aMb and Mb, whereas Y146 is exposed in aMb but protected in Mb.

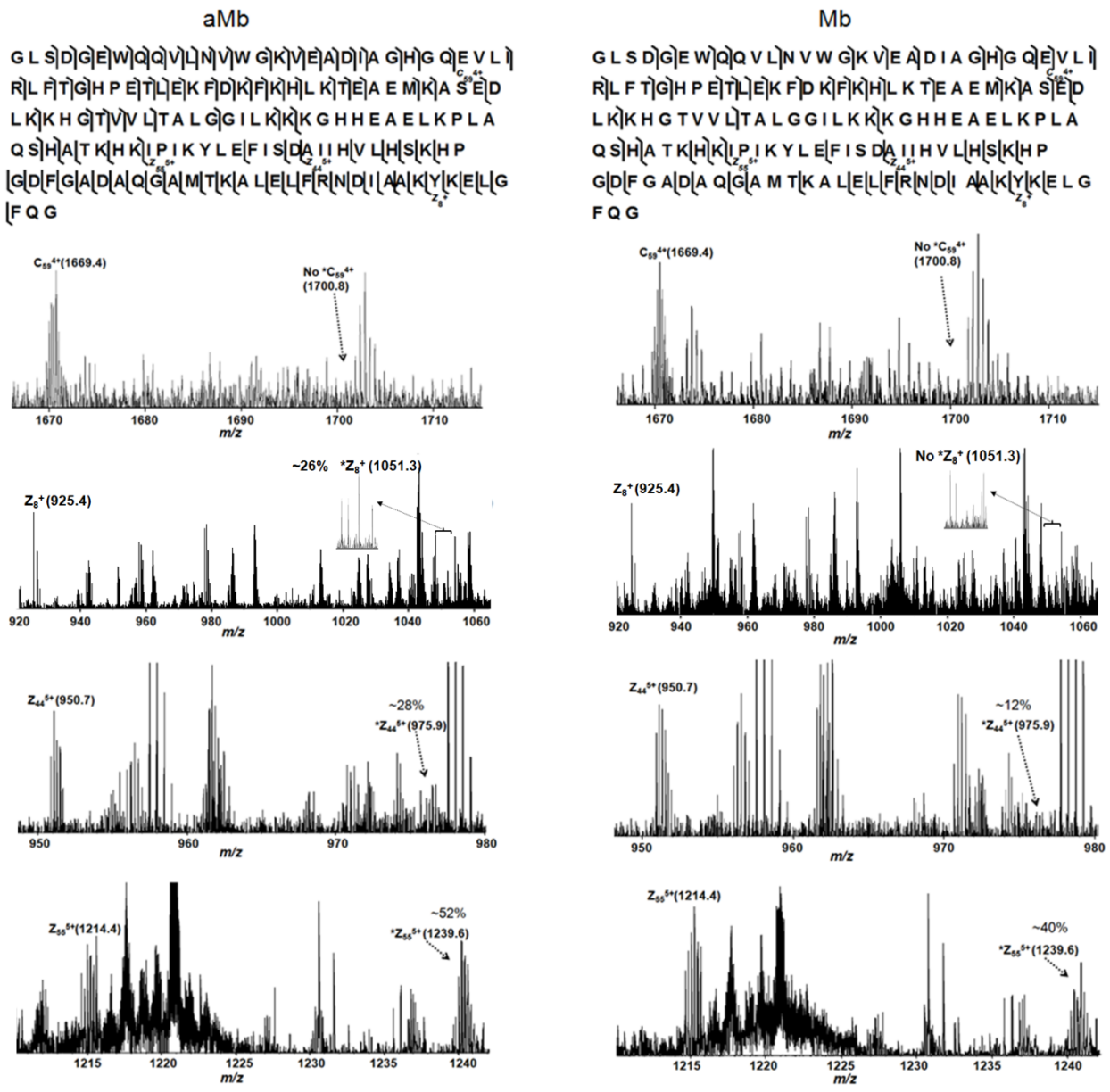


Figure 5.3: Sequence with the observed fragmentation pattern with ECD for aMb and Mb (top). The zoomed-in ECD spectra of $[*aMb+16H]^{16+}$ and $[*Mb+16H]^{16+}$ (bottom).

The remainder of the iodination occurs mostly in the middle of the sequence of Mb. Although we could obtain reasonable sequence coverage in the top-down sequencing, the modification yields at the residue level were difficult to determine for those amino-acid residues that incurred minor amounts of modification or were located in the middle of an unfragmented protein sequence. For these cases, top-down sequencing does not allow assignment of 11 histidine residues, some of which are located close to each other in the middle of the protein sequence. We calculated the percentage of mono-iodinated z_{55}^{5+} in aMb and Mb to be 52% and 40%, respectively, as described above, leaving the remaining 48% and 60% not assigned to any particular sites. Thus, we look to bottom-up sequencing for confirmation.

5.4.4 Bottom-up Sequencing

To complete the assignments and aid the evaluation of the top-down data, we applied bottom-up sequencing to both Mb and aMb. We relied primarily on the accurate masses of each tryptic peptide and those of its modified "siblings" to obtain extracted-ion chromatograms. For peptides containing more than one histidine/tyrosine site, high chromatographic efficiency and informative MS/MS data can locate the modification. To calculate the fraction-modified, we summed in the numerator the signal intensities corresponding to all peptides having a modification at a given residue and summed in the denominator all the intensities of detected peptides that contain the residue; this summation includes all charge states of a peptide. As an example, the extracted ion chromatograms of peptide 80-96, which contains three histidines (Figure 5.4), show that three peptides, extricated by digestion of the mono-iodinated aMb, elute sequentially at

22.8, 23.0, and 23.5 min. On the basis of product-ion spectra (MS/MS), we conclude that the comparably abundant eluents are iodinated H93, H81 and H82, respectively. Turning to Mb, we find that the abundance ratios for these three components change, and now H93 and H82 were modified to a small extent.

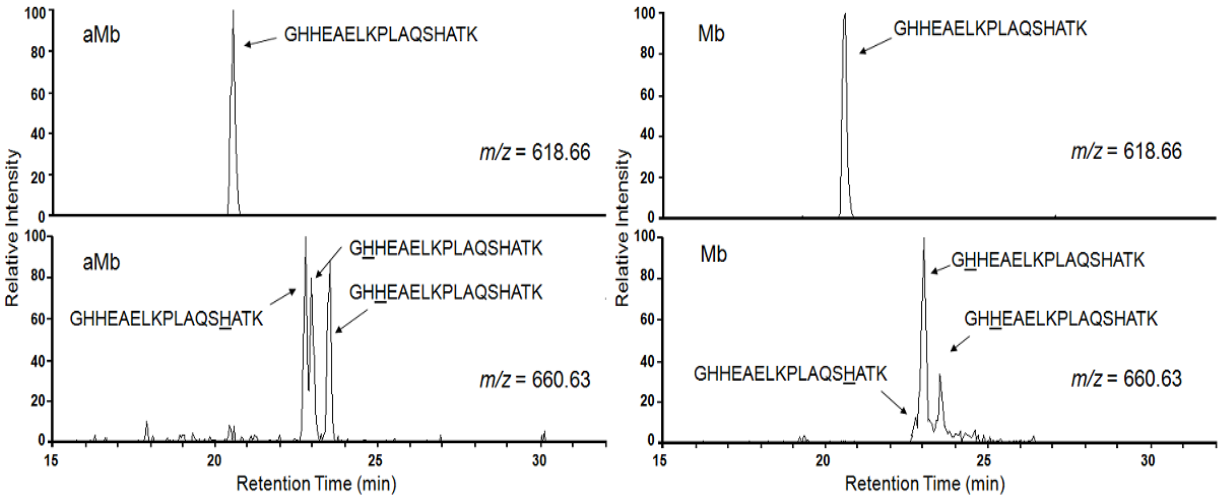


Figure 5.4: High MS resolution extracted ion chromatograms of doubly charged peptide 80-96 of aMb (top left) and Mb (top right), monoiodinated peptide 80-96 of aMb (bottom left) and aMb (bottom right). The assigned modification sites are underlined.

A comparison of overall calculated residue-level modifications of aMb and Mb (Figure 5.5) shows that residues H82, H93 and H97 are modified more than 2% in aMb but less than 0.2% in Mb, indicating a significant difference between the apo- and holo-proteins at the sites containing these residues. Residues H24, H36, H48, H64, H113, H116 and H119 are modified at less than 1.0% in both the apo and holo proteins. Among those histidines, H48 and H113 are iodinated to a greater extent in Mb than in aMb. Residues H82, H93, and H97 are modified in excess of 1.0% for aMb, significantly greater than for Mb. Residue H81 is modified most extensively among all histidines in Mb; the yield is 1.5%, slightly smaller than for aMb (2.0%). Although it is likely that the modified and unmodified peptides have different ionization efficiencies that might compromise the quantification, any bias is suffered by all the peptides, and, thus, the bias does not affect the quantitative trends we observe and the conclusions we draw. Indeed, the trends reflect structural differences between apo and holo states.

It is relatively common to produce di-iodinated tyrosines in this experiment. In bottom-up data analysis, we took into account every iodinated species, no matter how many modifications are on one site, summing all their signal areas to obtain the total abundance of modified peptides even though our focus is singly modified species of all types. For this reason and larger dynamic range of bottom-up sequencing, we find larger numbers of tyrosine modifications in the bottom-up rather than the top-down approach.

5.4.5 Comparison of Bottom-up and Top-down

Bottom-up and top-down strategies are complementary and have different advantages in analyzing differences in protein structures. Bottom-up experiments,

despite possible biases in enzymatic cleavage, can determine the modification fractions for each amino acid, down to less than 0.2%. This low level of detection in bottom-up sequencing is an advantage over the top-down approach when studying a protein that is rich in histidine and tyrosine. Top-down sequencing (as shown for Mb and aMb) provides estimates of where modifications happen most frequently, but it lacks the sensitivity and dynamic range that are available in the bottom-up approach. The top-down approach selects only a given modified state of a protein (we chose only the singly modified) and, thus, excludes those proteins that have undergone two or more modifications. Bottom-up analysis integrates over all modification states of the protein (e.g., mono-iodination, di-iodination). Top-down sequencing can be used to determine the distribution of modifiable sites within a protein state or among different states, and the results can roughly evaluate the modification level in regions where there are several reactive residues. Top-down sequencing also eliminates sample-preparation steps and avoids peptide loss.

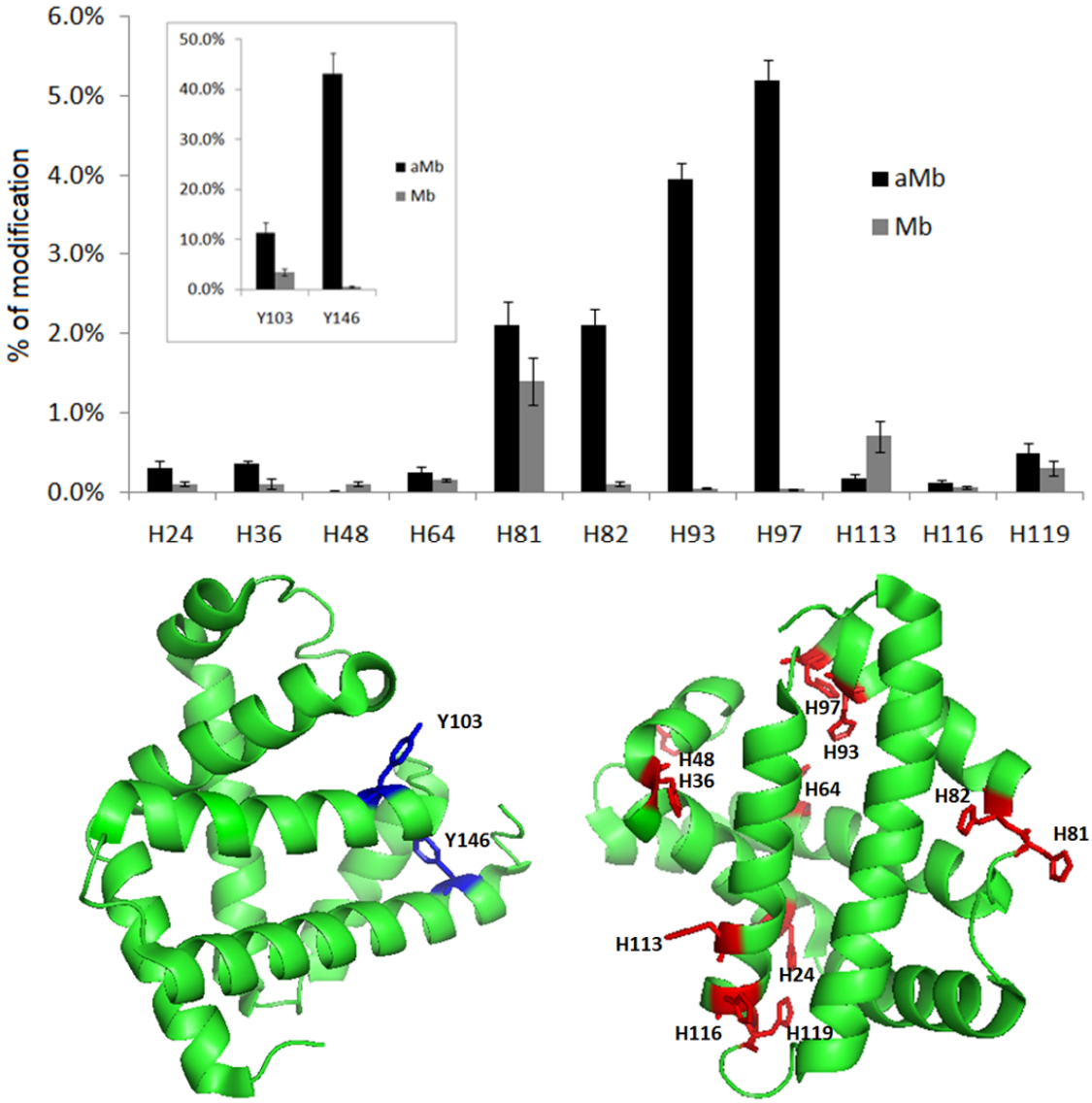


Figure 5.5: Plot of the yield of iodinated histidine and tyrosine residues in aMb and Mb, analyzed by the bottom-up strategy (top). Tyrosine and histidine sites mapped in a crystal structure of myoglobin (bottom, PDB ID: 1WLA).

5.4.6 Application 1: Iodination of Apo- and Holo- Carbonic Anhydrase II

We applied this FPOP-based iodination method to apo- and holo-carbonic anhydrase. Carbonic anhydrase is a zinc-containing metalloenzyme that catalyzes the hydration of CO_2 in solution. We argue that the extents of modification at the residue level are a function of the solvent accessibility of the particular residue, as shown in the protein structure onto which the tyrosine and histidine residues are mapped (Figure 5.6). The comparison of the modification extent of the same residue between two states shows no significant difference of the solvent accessibility (Figure 5.6), examined by T-test. This result supports previous conclusions showing that the removal of the zinc causes very little change to the tertiary structure of the protein (*49, 50*). This experiment serves as a negative control and proves that iodination extents are a measure of conformation and its change between two states.

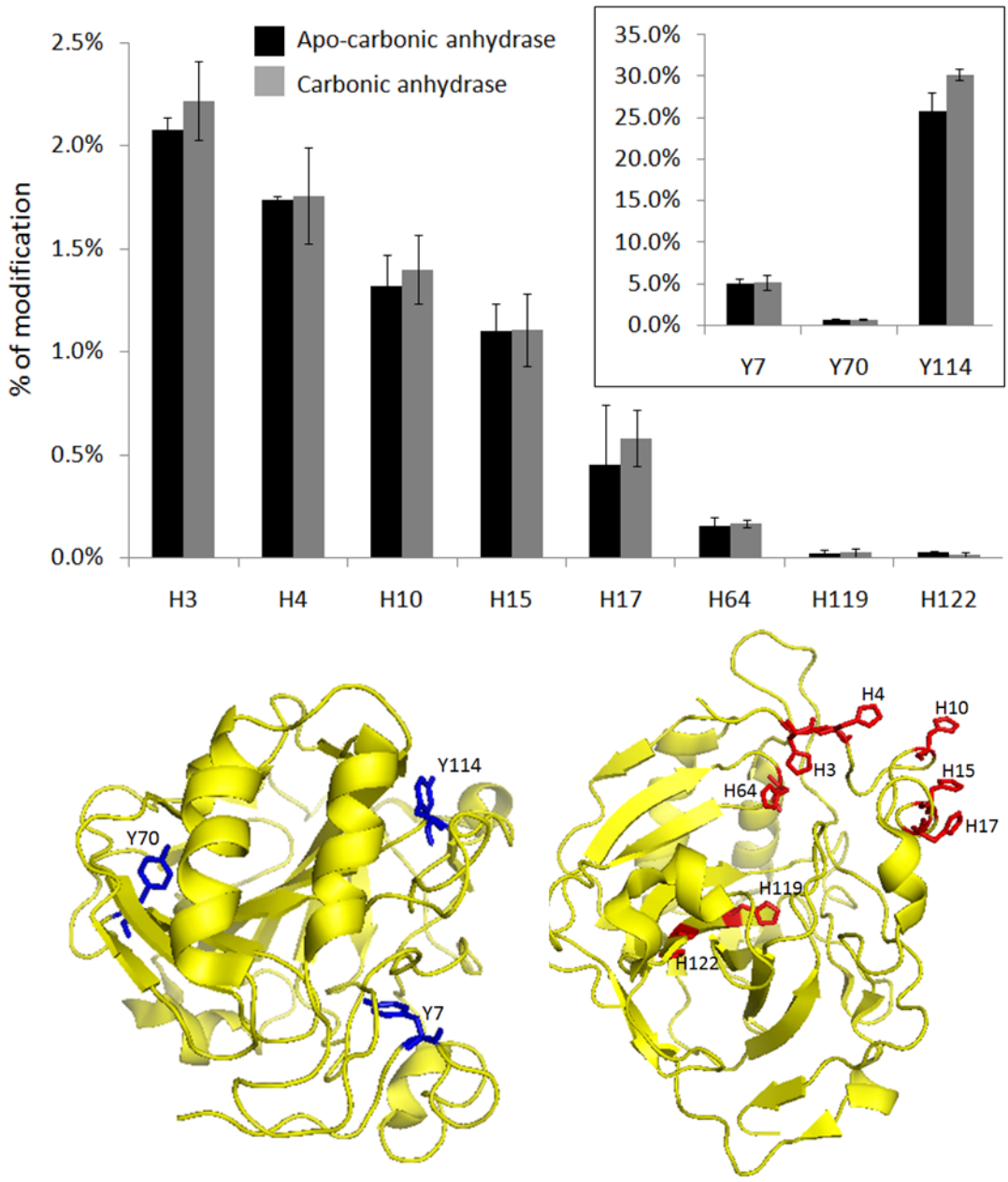


Figure 5.6: Plot of the yield of iodinated histidine and tyrosine residues in apo- and holo-carbonic anhydrase, analyzed by bottom-up strategy (top). Tyrosine and histidine sites mapped in a crystal structure of carbonic anhydrase (bottom, PDB ID: 2CBA).

5.4.7 Application 2: Iodination of lispro, insulin-EDTA and the zinc form of insulin

We also submitted lispro, insulin-EDTA, and the zinc form of insulin to iodination to test the applicability of this footprinting strategy for mapping the interaction sites in protein complexes. We chose these systems because human insulin exists as dimers in the absence of zinc and hexamers in the presence of zinc, whereas the lispro isoform is monomeric (*51*). Furthermore, there are four tyrosines in each insulin (or lispro) monomer: Y14 and Y19 in A chain, Y16 and Y26 in B chain, providing sufficient targets for the footprinting. Under the experimental conditions we chose for FPOP, an insulin-EDTA solution is likely comprised of monomers and dimers, whereas the zinc form of insulin is mainly an equilibrated mixture of dimers and hexamers. Lispro, a protein in which proline 28 and lysine 29 of the B chain in human insulin are interchanged, exists mainly as a monomer in solution and, therefore, has a more rapid response in diabetes therapy (*52*).

The iodination yields for lispro, insulin-EDTA and zinc form of insulin (Figure 5.7) show that the two tyrosines in A chain are iodinated to a comparable extent for lispro and insulin-EDTA. The other two tyrosines (Y16 and Y26) of the B chain, however, are more modified in lispro, indicating that they are involved in the dimerization interface of the human form. All tyrosines in the zinc form of insulin are modified to a lower extent than in the other two forms, owing to the increased protection that occurs upon the formation of hexamers. Although an ensemble of different oligomeric states is likely sampled in these experiments, the modification differences among them reveals some of the interaction sites involved in oligomerization.

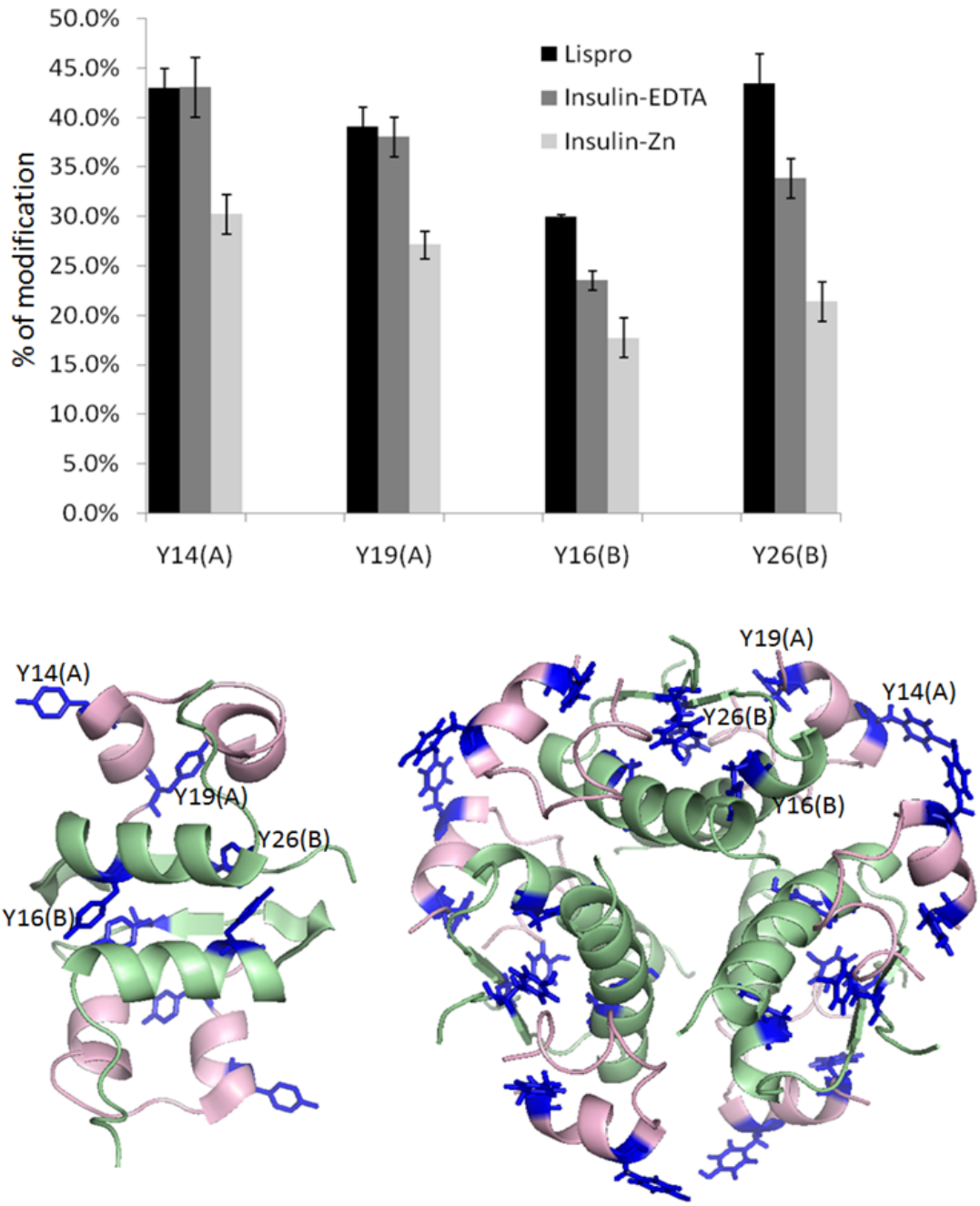


Figure 5.7: Plot of the yield of iodinated tyrosine residues in lispro, insulin with EDTA and the zinc form of insulin, analyzed by the bottom-up strategy (top). Tyrosine sites mapped in one insulin molecule present in a native dimer structure (bottom left, PDB ID: 1GUJ) and a hexamer structure (bottom right, PDB ID: 1AI0).

5.5 Conclusion

One outcome of this study is a demonstration of the versatility of the FPOP strategy in its ability to accommodate various reagents. The iodine species produced from photolysis of iodobenzoic acid selectively modifies histidine and tyrosine residues in proteins, and, more importantly, serves as a specific probe for changes in protein conformation between apo *vs.* holo states and as an indicator of protein oligomerization. The difference in modification extents can be seen in either the global mass spectral data or the peptide and amino-acid residue-level comparisons. The data processing is relatively easy compared to ·OH FPOP owing to the modification specificity and larger mass displacement than for oxidative modifications by ·OH. Unlike hydrogen peroxide, which gradually oxidizes protein even without photolysis, the precursor for ·I does not react with a protein in the absence of laser irradiation; thus, modification is only due to reactions with an iodine species. For this reason, removing the radical-generating reagent following protein modification is not as critical as in ·OH FPOP experiments (*53*). Unlike other specific footprinting reactions (e.g., acetylation of Lys or addition of *N*-ethylmaleimide to Cys), radical reactions occur rapidly, most likely on a time scale that preempts protein unfolding, providing an advantage for this specific footprinting reagent. Characterizing iodination modification by bottom-up and/or top-down depends largely on the protein sequence and the residues that are of interest. To our knowledge, this is the first example of combining protein footprinting and top-down MS analysis.

In the future, we intend to develop for FPOP more reagents that generate different radicals with different selectivity towards amino acids, affording a protein footprinting

tool box containing a variety of radicals appropriate for answering different and difficult questions.

5.6 References

1. Englander, J. J., Del Mar, C., Li, W., Englander, S. W., Kim, J. S., Stranz, D. D., Hamuro, Y., and Woods, V. L., Jr. (2003) Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry, *Proc Natl Acad Sci U S A 100*, 7057-7062.
2. Kaltashov, I. A., Bobst, C. E., and Abzalimov, R. R. (2009) H/D exchange and mass spectrometry in the studies of protein conformation and dynamics: is there a need for a top-down approach?, *Anal Chem 81*, 7892-7899.
3. Xu, G., and Chance, M. R. (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics, *Chem Rev 107*, 3514-3543.
4. Hambly, D. M., and Gross, M. L. (2005) Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale, *J Am Soc Mass Spectrom 16*, 2057-2063.
5. Konermann, L., Stocks, B. B., Pan, Y., and Tong, X. (2009) Mass spectrometry combined with oxidative labeling for exploring protein structure and folding, *Mass Spectrom Rev*.
6. Kiselar, J. G., and Chance, M. R. Future directions of structural mass spectrometry using hydroxyl radical footprinting, *J Mass Spectrom*.

7. Jha, S. K., and Udgaonkar, J. B. (2007) Exploring the cooperativity of the fast folding reaction of a small protein using pulsed thiol labeling and mass spectrometry, *J Biol Chem 282*, 37479-37491.

8. Guan, L., and Kaback, H. R. (2007) Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins, *Nat Protoc 2*, 2012-2017.

9. Wen, J., Zhang, H., Gross, M. L., and Blankenship, R. E. (2009) Membrane orientation of the FMO antenna protein from Chlorobaculum tepidum as determined by mass spectrometry-based footprinting, *Proc Natl Acad Sci U S A 106*, 6134-6139.

10. Back, J. W., de Jong, L., Muijsers, A. O., and de Koster, C. G. (2003) Chemical cross-linking and mass spectrometry for protein structural modeling, *J Mol Biol 331*, 303-313.

11. Hambly, D. M., and Gross, M. L. (2007) Laser Flash Photochemical Oxidation to Locate Heme Binding and Conformational Change in Myoglobin, *Int. J. Mass Spectrom. 259*, 124-129.

12. Guozhong Xu, M. R. C. (2005) Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting, *Analytical Chemistry 77*, 4549-4555.

13. Xu, G., and Chance, M. R. (2004) Radiolytic modification of acidic amino acid residues in peptides: probes for examining protein-protein interactions, *Anal Chem 76*, 1213-1221.

14. Gau, B. C., Sharp, J. S., Rempel, D. L., and Gross, M. L. (2009) Fast photochemical oxidation of protein footprints faster than protein unfolding, *Anal Chem 81*, 6563-6571.

15. Chen, J., Rempel, D. L., and Gross, M. L. (2010) Temperature jump and fast photochemical oxidation probe submillisecond protein folding, *J Am Chem Soc 132*, 15502-15504.

16. Stocks, B. B., and Konermann, L. (2009) Structural characterization of short-lived protein unfolding intermediates by laser-induced oxidative labeling and mass spectrometry, *Anal Chem 81*, 20-27.

17. Gau, B. C., Chen, H., Zhang, Y., and Gross, M. L. Sulfate radical anion as a new reagent for fast photochemical oxidation of proteins, *Anal Chem 82*, 7821-7827.

18. Morrison, M., and Schonbaum, G. R. (1976) Peroxidase-catalyzed halogenation, *Annu Rev Biochem 45*, 861-888.

19. Huwiler, M., Burgi, U., and Kohler, H. (1985) Mechanism of enzymatic and non-enzymatic tyrosine iodination. Inhibition by excess hydrogen peroxide and/or iodide, *Eur J Biochem 147*, 469-476.

20. Taurog, A., Dorris, M. L., and Doerge, D. R. (1996) Mechanism of simultaneous iodination and coupling catalyzed by thyroid peroxidase, *Arch Biochem Biophys 330*, 24-32.

21. Wolff, J., and Covelli, I. (1969) Factors in the iodination of histidine in proteins, *Eur J Biochem 9*, 371-377.

22. Kretsinger, R. H. (1968) A crystallographic study of iodinated sperm whale metmyoglobin, *J Mol Biol 31*, 315-318.

23. Sigler, P. B. (1970) Iodination of a single tyrosine in crystals of alpha-chymotrypsin, *Biochemistry 9*, 3609-3617.

24. Ghosh, D., Erman, M., Sawicki, M., Lala, P., Weeks, D. R., Li, N., Pangborn, W., Thiel, D. J., Jornvall, H., Gutierrez, R., and Eyzaguirre, J. (1999) Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase, *Acta Crystallogr D Biol Crystallogr 55*, 779-784.

25. Ly, T., and Julian, R. R. Elucidating the tertiary structure of protein ions in vacuo with site specific photoinitiated radical reactions, *J Am Chem Soc 132*, 8602-8609.

26. Ly, T., and Julian, R. R. (2008) Residue-specific radical-directed dissociation of whole proteins in the gas phase, *J Am Chem Soc 130*, 351-358.

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

28. Meng, F., Forbes, A. J., Miller, L. M., and Kelleher, N. L. (2004) Detection and localization of protein modifications by high resolution tandem mass spectrometry, *Mass Spectrom Rev 24*, 126-134.

29. Pesavento, J. J., Mizzen, C. A., and Kelleher, N. L. (2006) Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4, *Anal Chem 78*, 4271-4280.

30. Ge, Y., Rybakova, I. N., Xu, Q., and Moss, R. L. (2009) Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that

truncation alters protein phosphorylation state, *Proc Natl Acad Sci U S A 106*, 12658-12663.

31. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proc Natl Acad Sci U S A 101*, 9528-9533.

32. McLafferty, F. W., Horn, D. M., Breuker, K., Ge, Y., Lewis, M. A., Cerda, B., Zubarev, R. A., and Carpenter, B. K. (2001) Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance, *J Am Soc Mass Spectrom 12*, 245-249.

33. Lindskog, S., and Malmstrom, B. G. (1962) Metal binding and catalytic activity in bovine carbonic anhydrase, *J Biol Chem 237*, 1129-1137.

34. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci 4*, 2411-2423.

35. Stewart, J. J. P. (1989) Optimization of parameters for semiempirical methods. I. Method., *J Comp. Chem. 10*.

36. Stewart, J. J. P. (1989) Optimization of parameters for semiempirical methods. II. Applications., *J Comp. Chem. 10*.

37. 3. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Vakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A.. Petersson, P. Y Ayala, Q. Cui, K.

Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D. J. Fox, T. Keith, M. A. Al- Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M W. Gill,. B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople. (1998) Gaussian 98, Revision A.6, Gaussian, Inc., Pittsburgh PA.

38. 4. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A .Pople. (2004) Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT.

39. Scott, A. P., and Radom, L. (1996) Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, MØller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors, *J. Phys. Chem. 100*, 16502-16513.

40. Dzvonik, M., Yang, S., and Bersohn, R. (1974) Photodissociation of molecular beams of aryl halides, *J Chem Phys 61*, 4408-4421.

41. Kawasaki, M., Lee, S. J., and Bersohn, R. (1977) Photodissociation of molecular beams of aryl halides: Translational energy distribution of the fragments, *J Chem Phys 66*, 2647-2655.

42. Cheng, P. Y., Zhong, D., and Zewail, A. H. (1995) Kinetic-energy, femtosecond resolved reaction dynamics. Modes of dissociation (in iodobenzene) from time-velocity correlations, *Chem Phys Lett 237*, 399-405.

43. Furlan, A. (1999) Photodissociation of a Surface-Active Species at a Liquid Surface: A Study by Time-of-Flight Spectroscopy, *J Phys Chem B 103*, 1550-1557.

44. Li, C. H. (1944) Kinetics of Reactions between Iodine and Histidine, *Journal of American Chemical Society 66*, 225-227.

45. Eliezer, D., and Wright, P. E. (1996) Is apomyoglobin a molten globule? Structural characterization by NMR, *J Mol Biol 263*, 531-538.

46. Picotti, P., Marabotti, A., Negro, A., Musi, V., Spolaore, B., Zambonin, M., and Fontana, A. (2004) Modulation of the structural integrity of helix F in apomyoglobin by single amino acid replacements, *Protein Sci 13*, 1572-1585.

47. Iniesta, J., Cooper, H. J., Marshall, A. G., Heptinstall, J., Walton, D. J., and Peterson, I. R. (2008) Specific electrochemical iodination of horse heart myoglobin at tyrosine 103 as determined by Fourier transform ion cyclotron resonance mass spectrometry, *Arch Biochem Biophys 474*, 1-7.

48. Sun, Q., Yin, S., Loo, J. A., and Julian, R. R. Radical directed dissociation for facile identification of iodotyrosine residues using electrospray ionization mass spectrometry, *Anal Chem 82*, 3826-3833.

49. Brewer, J. M., Spencer, T. E., and Ashworth, R. B. (1968) Zinc effect on physical properties of bovine carbonic anhydrase, *Biochem Biophys Acta 168*, 359-361.

50. Hakansson, K., Carlsson, M., Svensson, L. A., and Liljas, A. (1992) Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes, *J Mol Biol 227*, 1192-1204.

51. Jeffrey, P. D., and Coates, J. H. (1966) An equilibrium ultracentrifuge study of the self-association of bovine insulin, *Biochemistry 5*, 489-498.

52. Holleman, F., and Hoekstra, J. B. (1997) Insulin lispro, *The New England journal of medicine 337*, 176-183.

53. Hambly, D. M., and Gross, M. L. (2009) Cold chemical oxidation of proteins, *Anal Chem 81*, 7235-7242.

6. Mass Spectrometric Characterization of *Enterococcus faecium* Peptidoglycan

This work was done collaboratively with Gary Patti. Some of it was published in: Patti, G. J., Chen, J., Schaefer, J. & Gross, M. L. Charaterization of Structural Variations in the Peptidoglycan of Vancomycin-Susceptible *Enterococcus faecium*: Understanding Glycopeptide-Antibiotic Binding Sites Using Mass Spectrometry. *J Am Soc Mass Spectrom 2008, 19, 1467-1475*. This chapter is a recapitulation of the work with emphasis on the method development.

6.1 Abstract

Enterococcus faecium (*E. faecium*) is a human pathogen that causes nosocomial bacteremia, surgical-wound infection, endocarditis, and urinary tract infections. It presents a clinical challenge due to the increasing number of infections that are highly drug resistant, with vancomycin-resistant enterococcus (VRE) infection being one of them. The drugs vancomycin and other glycopeptides are large, rigid molecules that inhibit a late stage in bacterial cell wall peptidoglycan synthesis. To characterize the fine structure of peptidoglycan and elucidate the mode of action of glycopeptides, we developed a bottom-up mass spectrometry approach as a complementary method to solid-state NMR. We identified approximately 50 muropeptide structures and their variations as the main components of the peptidoglycan by using accurate mass measurement and MS/MS. We also measured in a semi-quantitative way the amount of each by extracting specific ion chromatograms and integrating their peak areas. Although the organism here is susceptible to vancomycin, only 3% of the digested peptidoglycan has the well-known *D*-Ala-*D*-Ala vancomycin-binding site. The data agree with the NMR results and supports a proposed template model of cell-wall biosynthsis. Our method provides a fast and reliable way of characterizing digested peptidoglycan and detecting structural variations with high sensitivity. We applied this method to confirm the *D*-Ala-*D*-Ala/*D*-Ala-*D*-Lac existing in the clinical *vanA*-containing vancomycin-susceptible enterococcus (VSE) samples.

6.2 Introduction

We describe in this chapter the development and application of the modern LC/MS technique for determining the structure of bacterial peptidoglycan fragments (*1*). We chose *Enterococcus faecium* as a model for the development because *E. faecium* is a clinically important organism.

The enterococci are a diverse and important group of bacteria that have important implications in human health. They cause many serious clinical infections such as urinary tract infections, bacteremia, bacterial endocarditis, etc. The enterococci have the capacity to acquire a wide variety of antimicrobial resistance factors that present serious problems in the treatment of patients (*2, 3*). In the last two decades, VRE have emerged in nosocomial infections of hospitalized patients especially in the united states (*4*).

The enterococcal PG is organized as a "fisherman's net", made of glycan chains held together by peptide cross-links (*5, 6*). The glycan is composed of alternating β-1,4-linked units of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc). Peptide stems are connected to the glycan chains through amide linkages between the carboxyl groups on the muramyl residues and the terminal amino groups of the peptide. In enterococci, the stem is *L*-Ala-*D*-Glu-*L*-Lys-*D*-Ala (figure 6-1). Enterococcal resistance is generally thought to involve the conversion of *D*-Ala-*D*-Ala carboxyl termini of peptidoglycan, the vancomycin binding site, to *D*-Ala-*D*-Lac (*2*). In vitro studies show a 1000-fold decrease of the binding constant of the modified vancomycin-peptidoglycan complex than the unmodified one (*7*). Recently, researchers suggested that vancomycin-like glycopeptides recognize other peptidoglycan structural motifs in addition to *D*-Ala-*D*-Ala (*8, 9*). For the purpose of better understanding the modes of

action of both the widely-used glycopeptides as well as their derivatives, some fundamental studies giving atomic-level detail of the peptidoglycan structure are needed.

Here we describe a method to use LC/MS coupled with accurate-mass measurements of peptidoglycan fragments to identify individual muropeptide species and, by integrating selected ion-current chromatographic peaks, estimate the relative amount of each. The identifications were also verified by MS/MS experiments within the same runs, as well as by stable-isotope labeling experiments. With this method, over 90% of the peaks eluting in the expected LC retention times were assigned to various muropeptides and their derivatives. This LC/MS/MS method provides complementary information to that from solid-state NMR.

The MS method was applied to investigate clinical *vanA*-containing VSE samples. Resistance to vancomycin among enterococci is commonly mediated by the horizontal transfer of a mobile genetic element containing the *vanA* operon (*10*), a mobile genetic element that alters cell-wall structure. Here an isolate from a new vancomycin-susceptible *E. faecium* that harbors biochemically active *vanA* operons was studied by MS.

6.3 Experimental

6.3.1 Muropeptide sample preparation

Starting cultures of *E. faecium* (ATTC 49624) were prepared by inoculating brain-heart infusion media with a single colony. Cultures were incubated overnight at 37 °C, but not aerated. Samples in this work were prepared by inoculating either brain-heart infusion or sterile enterococcal standard media (ESM) with the overnight starter cultures (1% final volume).

Cells were harvested at the end of log phase, when the absorbance at 660 nm reached ~1.0, by centrifugation at 10,000×*g* for 25 min at 4 °C. Pellets were rinsed with 40 mM triethanolamine hydrochloride (pH 7.0) three times followed by three rinses with H_2O, centrifuging after each rinse. Cells were then resuspended in H_2O, frozen, and lyophilized.

Mass spectrometry samples were prepared from the lyophilized whole cell by physical disruption, as was detailed before (*11*). The isolated cell walls were digested into muropeptides with lysozyme and mutanolysin following the protocol described by others (*12*) with slight modifications. All experiments were performed at neutral pH.

Clinical *vanA*-containing VSE are provided by Dr. Kyu Y. Rhee at Weill Cornell Medical College. The cell walls were isolated and digested using the same protocol described above.

6.3.2 LC/MS/MS

Liquid chromatography/MS and MS/MS were performed in the electrospray ionization (*13*) mode by using a PicoView PV-500 (New Objective, Woburn, MA) nanospray stage attached to either an LTQ-FT mass spectrometer or an LTQ-Orbitrap mass spectrometer (ThermoFisher, San Jose, CA).

Muropeptide samples were loaded into an uncoated 75 µm i.d. fused-silica capillary column with a 15 µm picofrit tip (New Objectives, Woburn, MA), packed with C18 reverse-phase material (3 µm, 100 Å; Phenomenex, Torrance, CA) for 15 cm. The column was eluted at a flow rate of 250 nL/min for 10 min with 0.1% (vol/vol) formic acid in water and subsequently with a 60-min linear acetonitrile gradient (0%-40%) with

0.1% formic acid. The samples, as they emerged from the column, were sprayed into an LTQ-FT mass spectrometer. Full mass spectra were recorded in the FT component of the instrument at 100,000 resolving power (at m/z = 400).

Accurate-mass product-ion spectra of muropeptides were acquired by introducing the samples by nanospray as they eluted from the LC to an LTQ-Orbitrap mass spectrometer. To obtain major components' product-ion spectra, cycles consisting of one full FT-scan mass spectrum and five ensuing data-dependent MS/MS scans acquired by the Orbitrap (with a normalized collision energy setting of 35% of the maximum available energy) were repeated continuously throughout the elution with the following dynamic exclusion settings: repeat count, 3; repeat duration, 15 s; exclusion duration, 30 s. Sometimes, for minor components, an inclusion list containing the precursor ions' masses was applied.

6.4 Results and Discussion

6.4.1 Optimization of the Chromatograms

Enterococcal peptidoglycan units, shown in Figure 6.1, are cross-linked via *D*-aspartic acid or *D*-asparagine (*D*-Asx) bridges. By using *N*-acetylmuramidase enzymes, we selectively hydrolyzed the β-1,4 linkage between *N*-acetylmuramic acid and *N*-acetylglucosamine but left the cross-links intact. We call those muropeptides having only one stem a monomer, and two stems connected by one bridge a dimer, three stems with two bridges a trimer, etc. Chromatographic separation is critical and challenging here, owing to the structural similarity among the muropeptides.

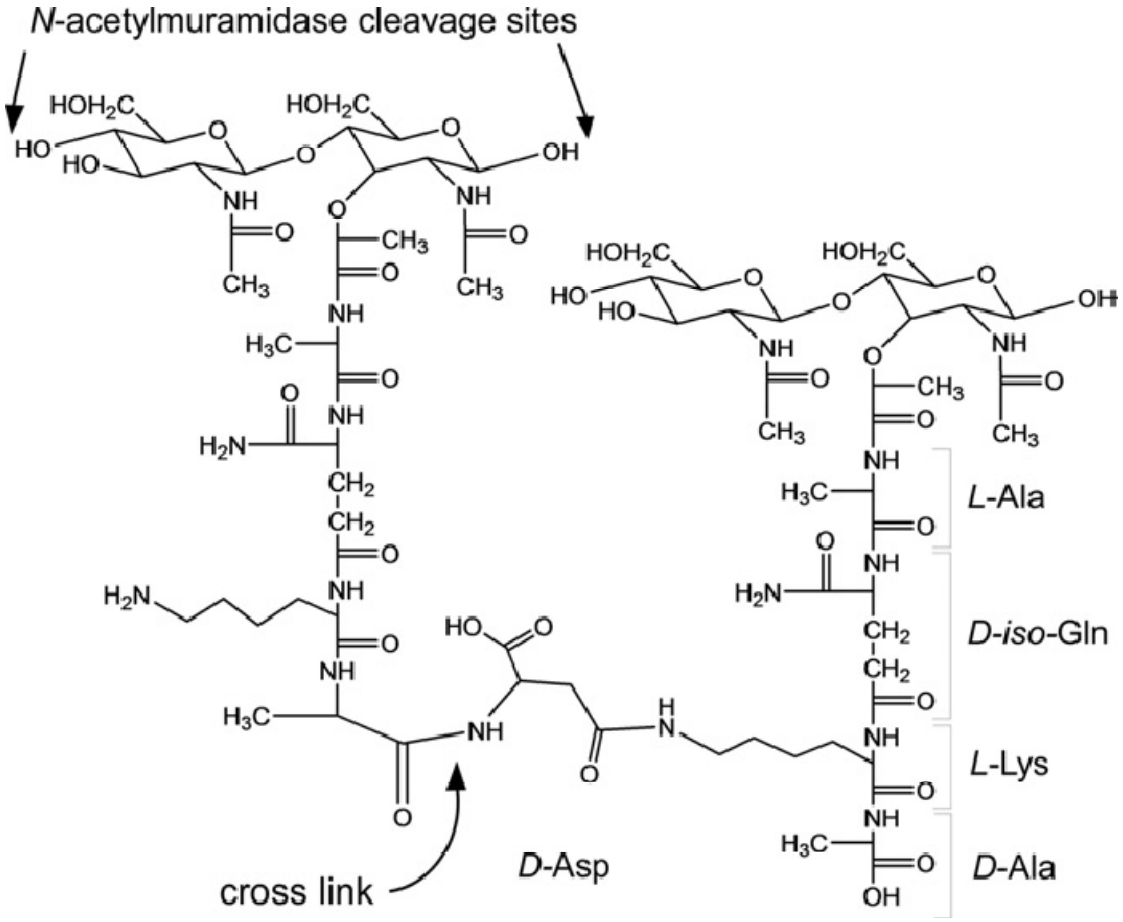


Figure 6.1: Representative chemical structure of *E. faecium* peptidoglycan. The glycan is composed of alternating β-1,4-linked units of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc). Peptide stems are connected to the glycan chains through an amide linkage between the carboxyl groups on muramyl residues and the terminal amino groups of the peptide. Cross-links are formed between *D*-Ala on one peptide and the *D*-Asp bridge on another. Figure is reproduced from reference (1).

The behavior of muropeptides on reversed phase supports is sensitive to pH and ionic strength of the eluent, to the gradient, and also to temperature. Researchers optimized the HPLC condition for muropeptide samples (*14*) and achieved good separation. Such conditions, however, are not compatible with mass spectrometry, owing to the high salt content of the gradient and also to an inability to change the HPLC solvents for proteomic experiments; the solvents were water with 0.1% formic acid as aqueous phase (A) and acetonitrile with 0.1% formic acid as organic phase (B). We were able to optimize the chromatogram by adjusting the loading method, gradient profile, the column size and the type of the packing material in the column. One variation was made at a time, followed by the test and evaluation based on both peak separation and width. The optimized conditions are stated in the experimental section, and the total ion chromatogram is shown in Figure 6.2.

Although we did not achieve base-line separation for all species, we are still able to select a particular ion based on its *m/z,* allowing the differentiation of peaks that are co-eluting with each other. This is an advantage of coupling chromatography with mass spectrometry as the detector over using the traditional off-line separation followed by mass spectrometric analysis.

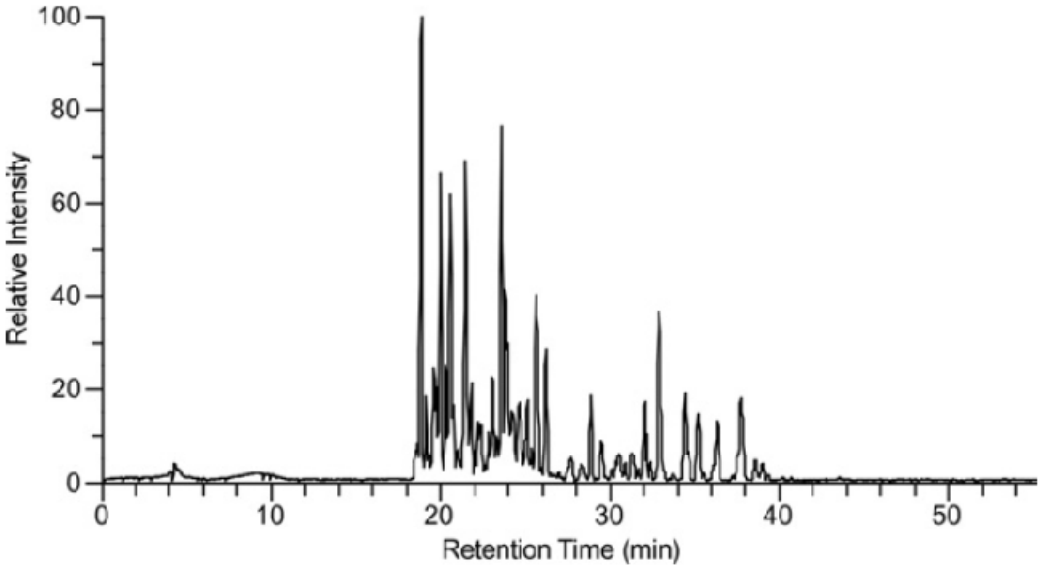


Figure 6.2: Total ion chromatogram for digested *E. faecium* muropeptides. The relative intensity is represented by total ion current as a function of retention time.

6.4.2 Muropeptide Identification

The identification of muropeptides relies on the accurate mass measurement, MS/MS assignments, and stable-isotope labeling experiments. Three sets of exemplary spectra (Figure 6.3) show the ions of *m/z* 696.2936, 824.3881 and 1010.4525, and these were assigned to the species with a dipeptide stem, tripeptide stem and tetrapeptide with a bridge (see structures 1,2 and 6 in Scheme 6.1), all having an accurate mass measurement within 1 ppm of the theoretical *m/z* value. The product-ion (MS/MS) spectra (Figure 6.3 (b), (e) and (h)) show the major fragmentations of the parent ions. In the isotopically labeled sample where the lysine and alanine were replaced by *L*-[1-^{13}C]lysine and *D*-

[^{15}N]alanine, the spectrum (c) of the dipeptide stem shows no mass shift compared to the unlabeled stem, whereas the mass spectrum of the tripeptide stem shows +1 shift due to the incorporation of one heavy lysine whereas the mass spectrum of the tetrapeptide stem with a bridge shows a +2 shift due to the incorporation of one heavy lysine and one heavy *D*-alanine. The labeled sample was originally made for solid-state NMR measurements but can be utilized here for the purpose of confirming the composition of muropeptide. Additionally, MS measurements can help determine the labeling efficiency, which is an important factor for the NMR analysis.

Most of the masses for eluents within 15-35 min retention time could be assigned as either the traditional muropeptides or their variations. Assisted by accurate mass measurement, MS/MS and the stable-isotope labeling experiment, we were able to assign with confidence 90% of the peaks in the mass spectra as muropeptide structures.

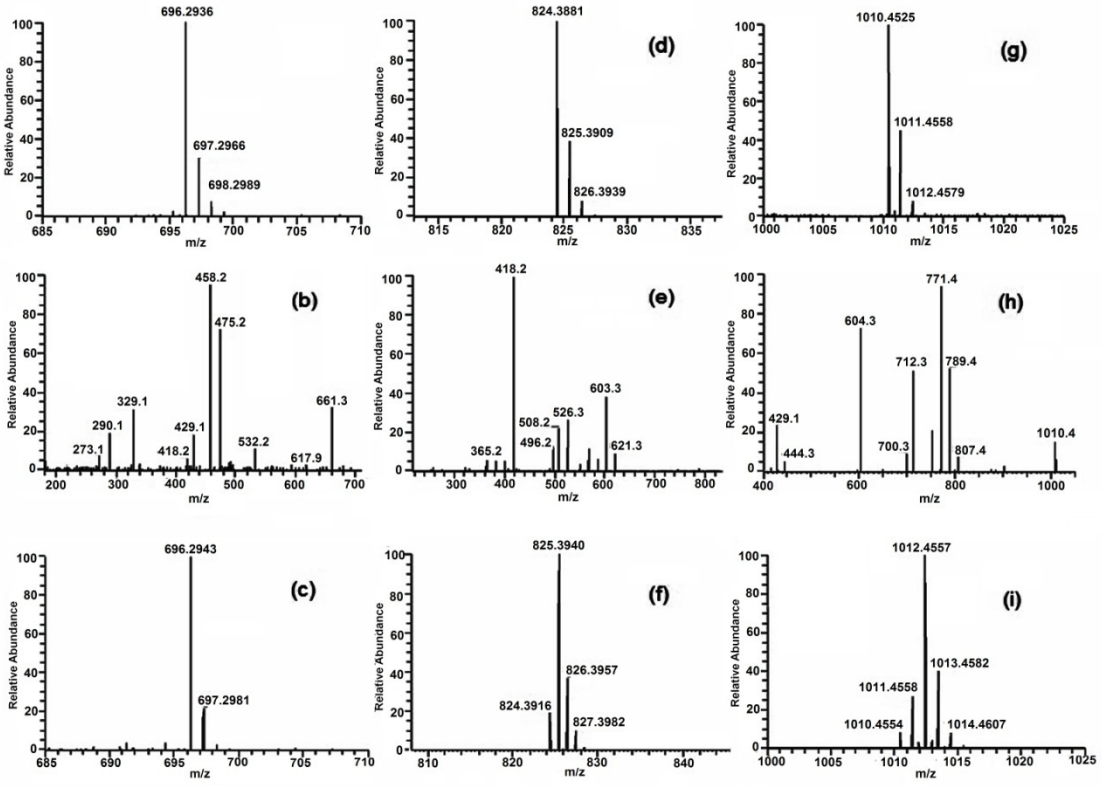
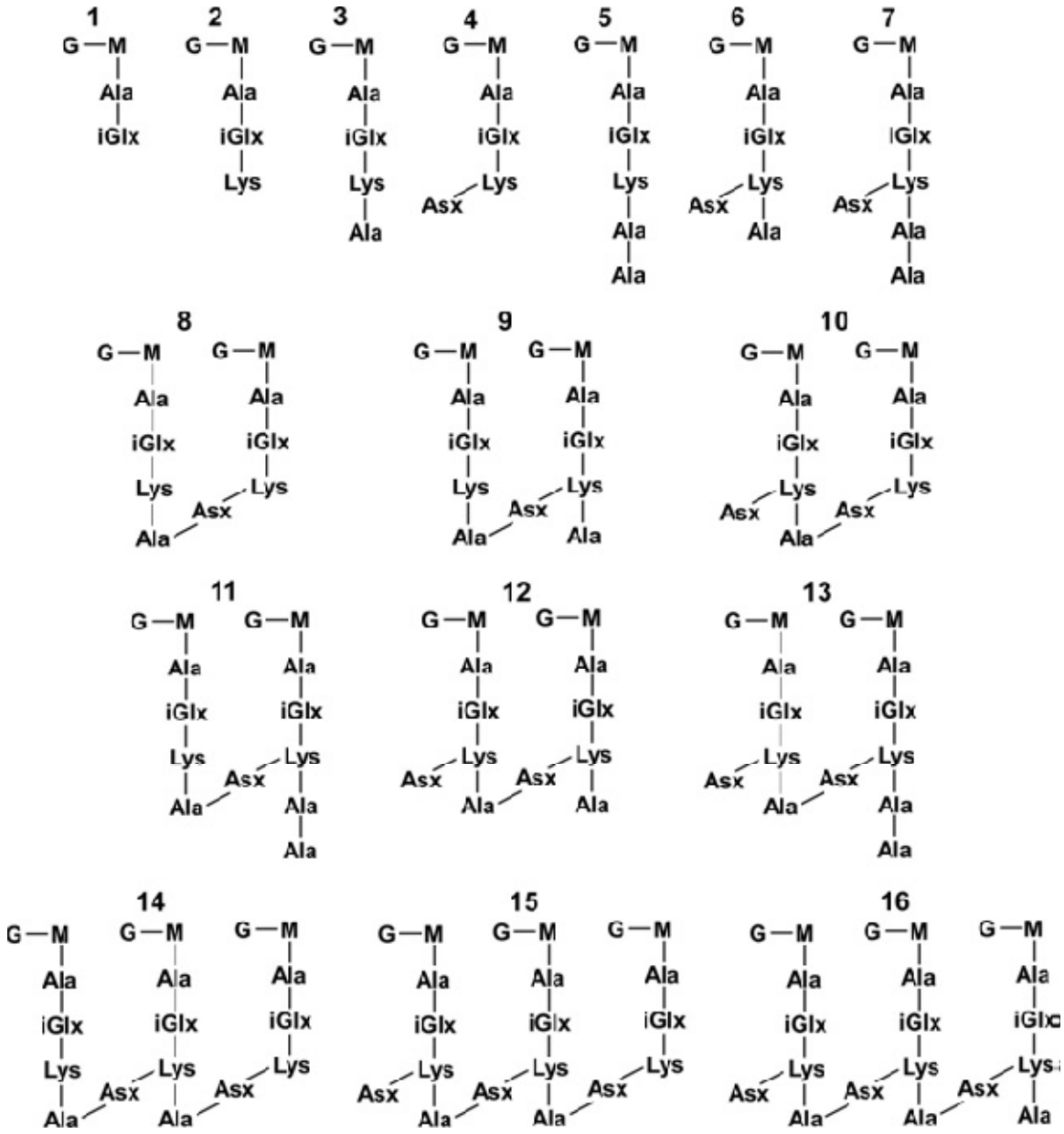


Figure 6.3: Exemplary spectra of three muropeptides. (a) Mass spectrum of muropeptide 1 (see Scheme 6-1); (b) Product-ion spectrum of muropeptide 1; (c) mass spectrum of isotopically enriched muropeptide 1; (d) mass spectrum of muropeptide 2; (e) Product-ion spectrum of muropeptide 2; (f) mass spectrum of isotopically enriched muropeptide 2; (g) mass spectrum of muropeptide 6; (h) product-ion spectrum of muropeptide 6; (i) mass spectrum of isotopically enriched muropeptide 6.



Scheme 6.1. Structures of assigned muropeptides. They include monomers (1-7), dimers (8-13) and trimers (14-16). Variations based on these primary structures are not shown in this figure. Figure is reproduced from reference (1).

6.4.3 Quantification

The quantification of muropeptide species was based on using extracted ion chromatograms. First, a list of *m/z* values associated with all muropeptides including structural variations was created. Different charge states of the same muropeptide were also added to the list. Then, the chromatograms of these ions were extracted from the total ion chromatogram, and the quantity was estimated by integrating the peak areas. Finally, the quantities of structural isomers and multiply charged ions of the same species were summed to give a total value, which was then used to calculate the percentage of this species among all the muropeptides. The accurate mass measurement and the baseline separation of the chromatogram ensure the unambiguous selection of ions.

The quantification of muropeptides, listed in Table 6.1, was presented as two sets of values. The percentages in the left column were calculated with respect to monomers (47%), dimers (43%) and trimers (8%). The percentages in the right column, however, were calculated with respect to stems. For example, the percentage of trimer stems is calculated according to the following equation:

$$\%\ of\ trimer\ stems = \frac{[trimer] \times 3}{[monomer] + [dimer] \times 2 + [trimer] \times 3 + [tetramer] \times 4}$$

This way of calculating percentages allows the comparison of our data with the results obtained by solid-state NMR. Surprisingly, given that *D*-Ala-*D*-Ala is the target of glycopeptides antibiotics, the number of muropeptide stems with *D*-Ala-*D*-Ala in the terminus (structures 5, 7, 11, 13 in scheme 6.1) is only 3% of the total stems. This number agrees with that determined by solid-state NMR, which was employed to measure these species to be 7% in whole cells (*15*). The difference between the two

values may be attributed to the peptidoglycan precursor and *N*-acetylglucosamine-*N*-acetyl-muramyl-pentapeptide-pyrophosphoryl-undecaprenol (Lipid II), which contain *D*-Ala-*D*-Ala stems and are not present in the samples for mass spectrometry analysis.

Species	Theoretical mass[a]	Measured mass	Percentage [%][b,c]	Percentage [%][c,d]
Monomers			47	30
1	695.2856	695.2855	<1	<1
2	823.3805	823.3802	22	13
3	894.4177	894.4174	5	3
4	938.4075	938.4071	13	8
5	965.4548	965.4555	<1	<1
6	1009.4446	1009.4446	6	4
7	1080.4817	1080.4825	<1	<1
Dimers			43	51
8	1814.8151	1814.8150	8	10
9	1885.8522	1885.8510	3	4
10	1929.8421	1929.8420	23	27
11	1956.8894	1956.8899	2	2
12	2000.8792	2000.8784	7	8
13	2071.9163	2071.9158	<1	<1
Trimers			8	15
14	2806.2497	2806.2450	2	3
15	2921.2767	2921.2767	5	10
16	2992.3138	2992.3140	1	2
Tetramers			2	4

[a]Theoretical masses were calculated with Xcalibur 2.0.
[b]Percentages with respect to monomers, dimers, trimers, and tetramers.
[c]Both percentages include structural variations of the species (see Table 2). Under optimized digestion conditions for data from triplicate runs of four different sample preparations, each species contribution was within ± 3% of the values listed here.
[d]Percentages with respect to total number of peptidoglycan stems.

Table 6.1. Distribution of structures in *E. faecium* peptidoglycan. Table is reproduced from reference (1).

6.4.4 *O*-Acetylated Muropeptides

One important structural variation of muropeptides is *O*-acetylation of *N*-acetylmuramic acid, identified as a mass increase of 42.011 (Figure 6.4, bottom left). The product-ion spectra of structures with and without *O*-acetylation show the same

fragmentation masses of *N*-acetylglucosamine and the peptide stem but different masses for fragments of the *N*-acetylmuramic acid stem, indicating that *O*-acetylation occurs on at *N*-acetylmuramic acid. This is consistent with previous findings (*16*).

Different *O*-acetylation extents were observed for those muropeptide structures, shown in Table 6.2. The extents suggest that either the enzyme performing *O*-acetylation has preference for certain target structures over others, or the exposure time of different muropeptide structures to the enzyme is different (a correlation of *O*-acetylation with muropeptide structures will be discussed in the next chapter). One average, 5-15% of the total peptidoglycan is *O*-acetylated. This is lower than the reported value (46-57%) for *E. faecium* (*17*). The difference may be within reproducibility of the experiment; a source of variation is the extent of *O*-acetylation may vary in different strains of the same bacterial species (*17-20*).

Species	Acetylation [%]	Cyclic Imide [%]	Gln/Asn [%]	Gln/Asp (Glu/Asn) [%]	Glu/Asp [%]
2	19				
3	2				
4	11	8	41	58	1
5	4	3	16	83	1
6	3	2	23	75	2

Table 6.2. Distribution of selected monomers. Table is reproduced from reference (1).

As indicated above, we used two *N*-acetylmuramidase enzymes, mutanolysin and lysozyme (*20*). They serve as complementary enzymes because the *O*-acetylated peptidoglycan is resistant to the activity of lysozyme but susceptible to mutanolysin isolated from *Streptomyces globisporus* (*21*). In a separate experiment, we digested *E.*

faecium peptidoglycan by using only lysozyme, and detected no *O*-acetylated species. The lack of detection is consistent with the earlier observation that cell walls of different enterococcal strains respond differently to lysozyme digestion owing to the different *O*-acetylation extents (*22*).

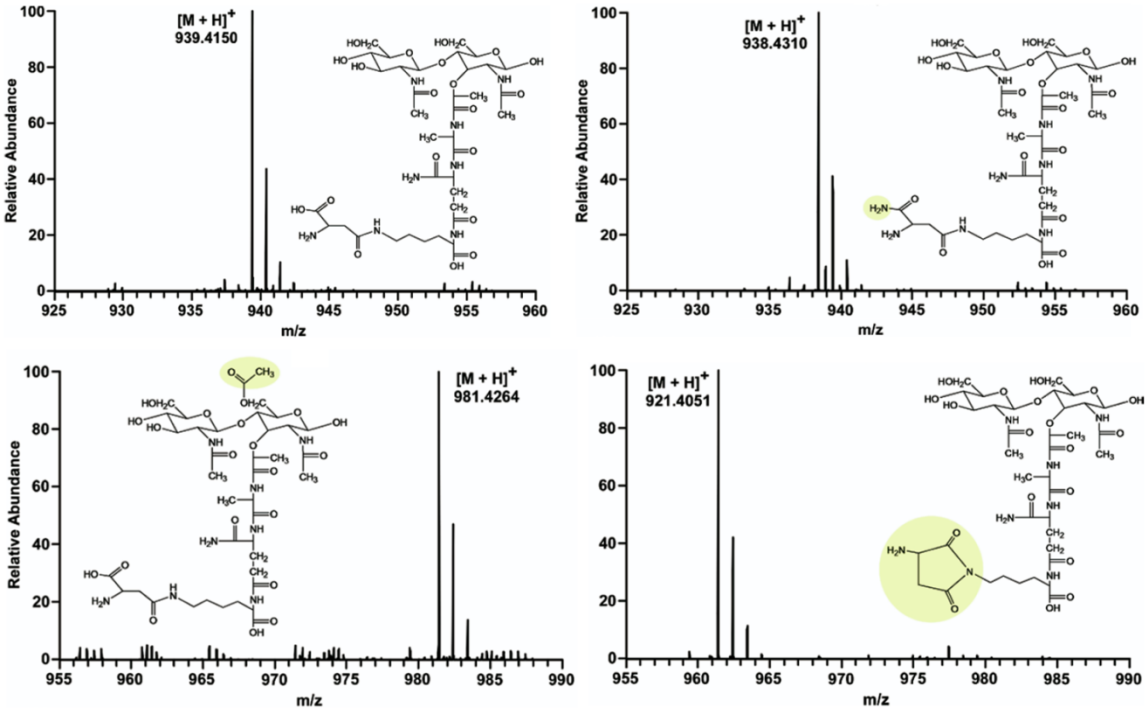


Figure 6.4: Mass spectra and assigned structures for variations of the muropeptide of *m/z* = 939.4150. The most likely structure for an m/z of 939.4150 (top left) is that with a *D*-aspartic acid bridge. An ion of I of 938.4310 (top right) is from the muropeptide with double amidation, *D*-asparagine and *D*-iso-glutamine. A mass increase of 42.011 is consistent with *O*-acetylation of *N*-acetylmuramic acid (bottom left). A mass decrease of 18.010 suggests the loss of water, and could be indicative of two types of cyclic imides (bottom right). The structure shown is consistent with NMR data (see Figure 6.6).

6.4.5 Amidation of *D*-iso-Glu and *D*-Asp

It was suggested that the presence of *D*-*iso*-glutamic acid residues in muropeptides affects cross-linking (*23*) and vancomycin resistance in *Staphylococcus aureus* (*24*). Amidation of the *D*-*iso*-glutamic acid may change both the efficiency of transpeptidation (*25*) and the affinity for glycopeptides antibiotics. Thus, characterization and quantification of such species may be important for understanding antibiotic resistance. There are two sites, *D*-*iso*-Glx (*D*-*iso*-glutamine or *D*-*iso*-glutamic acid) and *D*-Asx (*D*-asparagine or *D*-aspartic acid), that can be amidated (see Scheme 6.1), resulting in four possible species containing stems with *D*-*iso*-Gln /*D*-Asn, *D*-*iso*-Gln/*D*-Asp, *D*-*iso*-Glu/*D*-Asn and *D*-*iso*-Glu/*D*-Asp. Because species with *D*-*iso*-Gln/*D*-Asp and *D*-*iso*-Glu/*D*-Asn have the same accurate mass, MS identification of each is not possible when product-ion spectra do not provide much information. Alternatively, we used the hydroxylation profile of the muropeptide without a bridge as a reference to help identify these species. The singly charged muropeptide of *m/z* of 824.4 (see Scheme 6.1, Structure 2) contains only one site, *D*-*iso*-glutamine, that can be hydroxylated. It is shown in Figure 6.5.a that the peak area corresponding to the hydroxylated muropeptide of *m/z* of 825.4 is relatively small compared with that corresponding to the nonhydroxylated one. Furthermore, the species with double hydroxylation occur infrequently, indicating that there are few muropeptide stems with *D*-*iso*-Glu. This is consistent with previous finding that few *D*-*iso*-glutamic acid-containing peptidoglycan stems was found for other enterococcal organisms (*26*). We conclude, based on the above evidence, that most muropeptide species with a single hydroxylation represent *D*-*iso*-Gln/*D*-Asp stems.

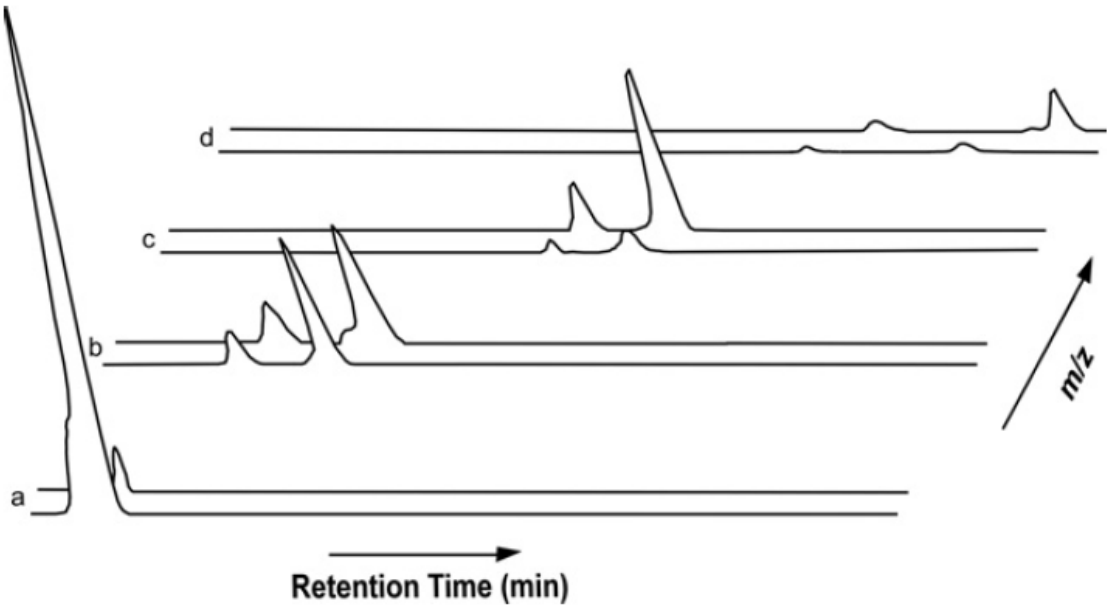


Figure 6.5: Extracted ion chromatograms for singly charged species with *m/z* of (a) 824.4 and 825.4; (b) 938.4 and 939.4; (c) 1009.5 and 1010.5; and (d) 1080.5 and 1081.5. The relative intensities of the signals for the *m/z* 1080.5 and 1081.5 ions are small, consistent with there being few *D*-Ala-*D*-Ala stems. Muropeptides with bridges elute as two distinct peaks with different retention times in the chromatogram. The selected ion chromatogram for the *m/z* 824.4 ion shows only one peak.

6.4.6 Structural Bridge Isomers

Interestingly, there are two chromatographic peaks corresponding to each ion having a monomer with a bridge, suggesting the presence of structural isomers associated with different bridge conformations (Figure 6.5). The product-ion mass spectra show no difference between the two isomers. It is known that the deamidation of *D*-Asn produces a mixture of α- and β-*D*-Asp residues (*27, 28*). Recently, however, it was suggested that all *D*-Asn bridges in *E. faecium* are from the amidation of *D*-Asp after its incorporation into the peptidoglycan (*29*). Additionally, it was proposed that the incorporation proceeds specifically by way of a β-aspartylphosphate intermediate, ruling out the possibility of α- and β- bridge isomers. Given that the β-*D*-Asp is the original species incorporated into peptidoglycan precursors and the 1- and 4-aspartyl carbons are similar, it is possible that either carbon could form a peptide bond with the 6-*N*-lysine and create unique bridge conformations.

6.4.7 Cyclic Imides

An accurate mass decrease of 18.010 associated with the loss of water was observed for the muropeptide monomers having aspartic acid bridges (Figure 6-4, bottom right). The water loss is not from the disaccharides or the first amino acid of the peptide stem, *L*-alanine, according to the product-ion mass spectrum. This modification is likely to be in the middle of the peptide stem, supported by the evidence that tripeptide monomers with bridges (Scheme 6-1, Structure 4) undergo water loss. We suggest two possibilities for the water loss: the formation of a succinimide structure between *D*-Asp

and 6-*N*-lysine (Figure 6-4, bottom), and the cyclization between *D*-Asp and *N*-*D*-iso-Gln.

We are unable to differentiate between the two proposed varieties of cyclic imides using the current mass spectrometric data. Monomers having *D*-*iso*-Gln and *D*-Asn do not show water loss because there is no free hydroxyl to be protonated and become a leaving group. The infrequency of a cyclic imide species makes detection and structural assignment difficult. The fact that there were two chromatographic peaks produced by some structures that are missing a water indicates the existence of both types of cyclic imide structures as a consequence of there being two bridge orientations, as discussed above.

The finding of the succinimide is consistent with solid-state NMR results (*15*). To our knowledge, this is the first detection of such structures in *E. faecium* peptidoglycan. Although of small quantity, this post-insertional modification may play an important role in the peptidoglycan structure and function.

6.4.8 Validity of MS Quantification: Comparison with Solid-State NMR Results

As indicated above, the MS quantification here suffers from the different ionization efficiencies that different muropeptide structures will have. To judge the quantitative integrity of the MS results, the percentage of total peptidoglycan stem cross-linked was calculated and compared with that directly measured by solid-state NMR. Cross-linking is equal to: (dimers + 2 × trimers +3 × tetramers) / (monomers + 2 × dimers + 3 × trimers + 4 × tetramers). From the MS data, ~40% of *E. faecium* peptidoglycan is cross-linked. This result is in good agreement with the 47% obtained from solid-state

NMR (*15*). MS is likely to discriminate against the larger fragments, thus underestimating the cross-linking value.

6.4.9 Peptidoglycan Structure and Cell-Wall Biosynthesis

Peptidoglycan stems terminating in *D*-Ala-*D*-Ala are crucial to cell-wall biosynthesis. The small quantity measured by MS is consistent with the template model proposed for *S. aureus* (*9*) and *E. faecium* (*15*). Those stems, which present only 3% of the total muropeptide, are incorporated into the nascent peptidoglycan to allow the cross-linking to occur (Figure 6-6). They are then post-insertionally modified to other species by certain enzymes. Such modification could prevent the binding of vancomycin and may be crucial for cell-wall assembly. We suggest that there must be an active *L*,*D*-carboxypeptidase that cleaves *D*-Ala-*D*-Ala stems that are not cross-linked in mature peptidoglycan. We also propose that other enzymatic modifications, including *O*-acetylation, amidation of aspartic acid, and formation of cyclic imides, occur in mature peptidoglycan rather than the nascent and template peptidoglycan (Figure 6-6). Although the modified peptidoglycan precursors may produce inefficient substrates for transglycosylase and transpeptidase, modified mature peptidoglycan may play an important role in biosynthetic regulation and prove important in virulence, allowing bacterial infections to evade host-immune system response.

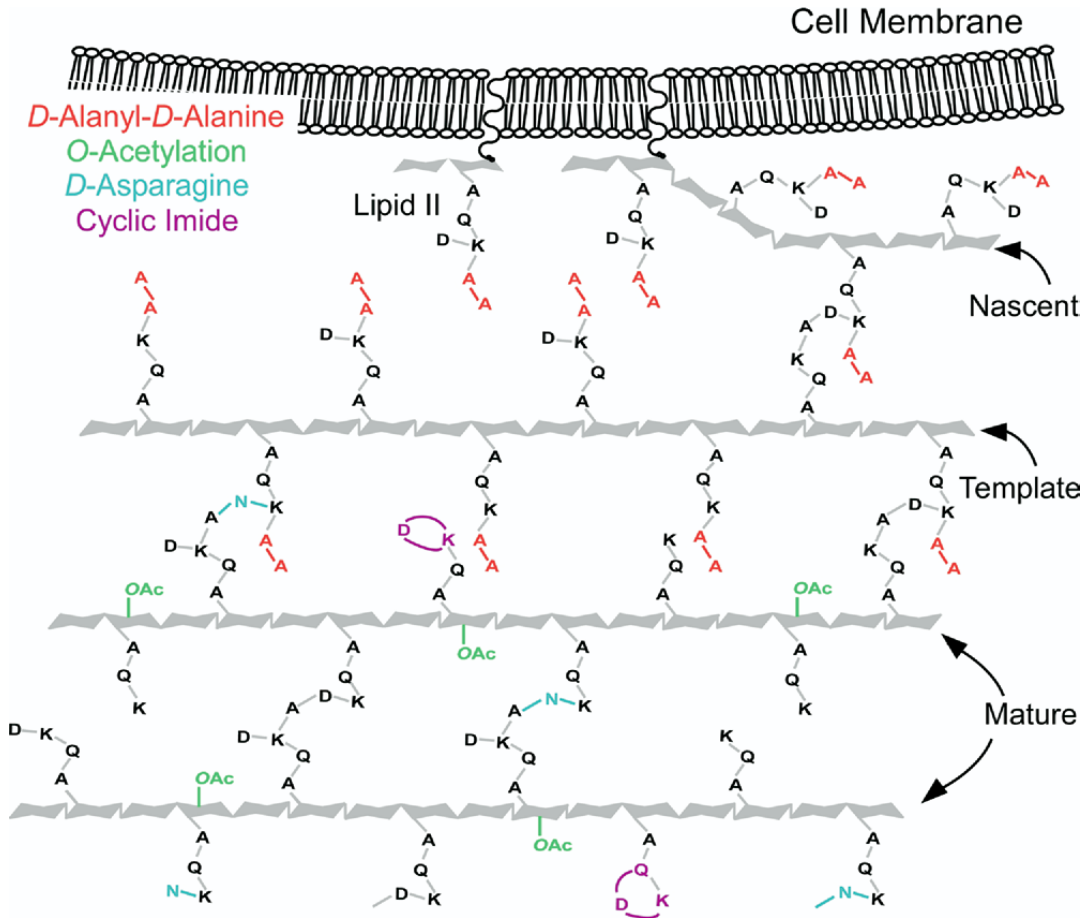


Figure 6.6: Schematic representation of the peptidoglycan biosynthesis for *E. faecium*. The few peptidoglycan stems terminating in *D*-Ala-*D*-Ala occur in nascent and template peptidoglycan where they are cross-linked. The post-insertional modifications are shown in the mature peptidoglycan. Figure is reproduced from reference (1).

6.4.10 Characterization of *VanA*-containing VSE

This is a collaborative work with Dr. Kyu Y. Rhee at Weill Cornell Medical College using the bottom-up MS method to characterize the peptidoglycan structure of clinical *vanA*-deficient and *vanA*-containing VSE samples obtained by our collaborator from the hospital. *VanA* is a type of gene cluster that causes acquired resistance of bacteria to glycopeptides antibiotics by redirecting cell-wall biosynthesis to produce *D*-Ala-*D*-Lac stems instead of *D*-Ala-*D*-Ala (*10*). The sample was obtained from a patient with VSE mitral valve endocarditis who failed vancomycin chemotherapy owing to a breakthrough infection with an isogenic vancomycin-resistant strain. Analysis of the peptidoglycan of the *vanA*-containing strain showed that peptidoglycan from both vancomycin-susceptible and –resistant isolates contained muropeptide with *D*-Ala-*D*-Lac (Figure 6.7) in the absence of vancomycin. In *vanA*-containing VSE, the ratio between *D*-Ala-*D*-Lac and *D*-Ala-*D*-Ala is approximately 1:4. This ratio suggests that the vanA operons of both *vanA*-containing VSE and VRE are active. Out MS results, together with other data (unpublished), reveal a previously unrecognized reservoir of *vanA*-type resistance (*30, 31*).

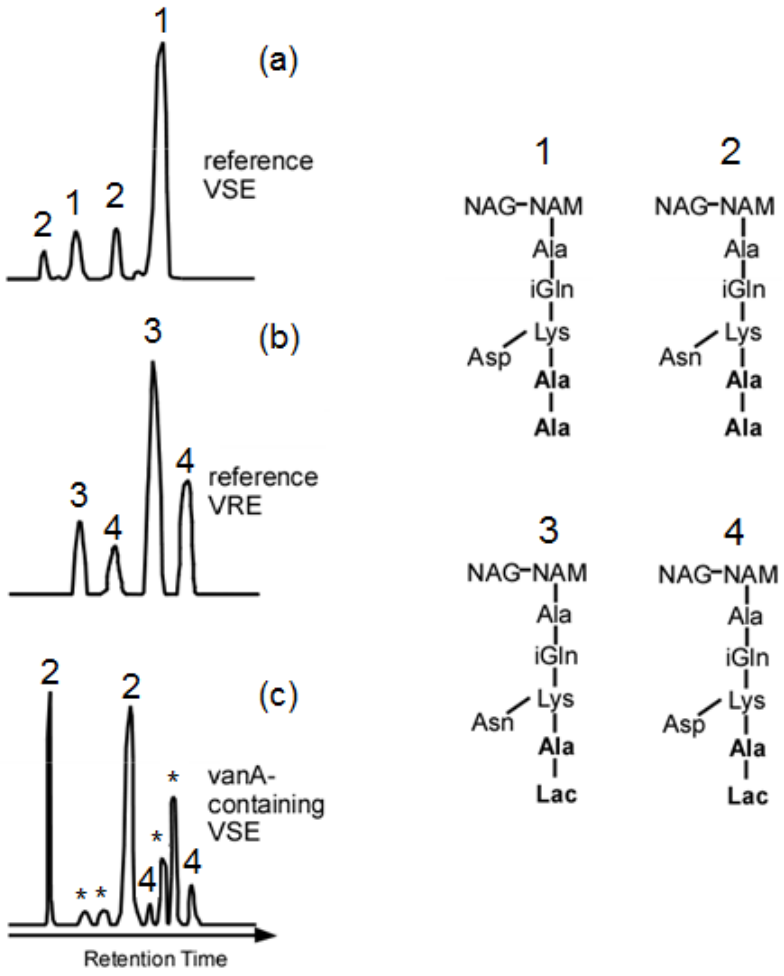


Figure 6.7. Extracted ion chromatograms of a pentapeptide containing *D*-Asp or *D*-Asn bridges for *vanA*-deficient reference VSE (a), *vanA*-containing reference VRE (b) and vanA-containing clinical VSE (c). The reference VSE chromatographic peaks correspond to stems with *D*-Ala-*D*-Ala (structures 1 and 2 on the right), and the reference VRE chromatographic peaks correspond to stems with *D*-Ala-*D*-Lac (structures 3 and 4 on the right). *VanA*-containing VSE peptidoglycan have both *D*-Ala-*D*-Ala and *D*-Ala-*D*-Lac, as shown by the chromatogram. Four peaks labeled * correspond to structures 1 or 3, and could not be distinguished further because they have identical masses.

6.5 Conclusion

We implemented an on-line LC/MS analysis for the identification and quantitation of *Enterococcus faecium* peptidoglycan. The method is sensitive and reliable, providing new structural detail pertaining to variations of peptidoglycan stems and allowing new understanding of glycopeptides antibiotic binding. The finding that vancomycin-susceptible *Enterococcus faecium* have very few *D*-Ala-*D*-Ala vancomycin binding sites is consistent with an NMR result(*15*). The importance of these infrequent *D*-Ala-*D*-Ala linkages as binding sites for one of the most powerful antibiotics clinically available underscores the need for characterizing peptidoglycan structural details such as were accomplished by this research. Using this MS method, we showed the presence of *D*-Ala-*D*-Lac species in a *vanA*-containing VSE sample. We suggest that the method described in this chapter is widely applicable for characterization of peptidoglycan in a variety of bacteria and for the investigation of effect of antibiotics on peptidoglycan.

6.6 References

1. Patti, G. J., Chen, J., Schaefer, J., and Gross, M. L. (2008) Characterization of structural variations in the peptidoglycan of vancomycin-susceptible Enterococcus faecium: understanding glycopeptide-antibiotic binding sites using mass spectrometry, *Journal of the American Society for Mass Spectrometry 19*, 1467-1475.
2. Courvalin, P. (2006) Vancomycin resistance in gram-positive cocci, *Clin Infect Dis 42 Suppl 1*, S25-34.

3. Aumeran, C., Baud, O., Lesens, O., Delmas, J., Souweine, B., and Traore, O. (2008) Successful control of a hospital-wide vancomycin-resistant Enterococcus faecium outbreak in France, *Eur J Clin Microbiol Infect Dis 27*, 1061-1064.

4. Fisher, K., and Phillips, C. (2009) The ecology, epidemiology and virulence of Enterococcus, *Microbiology (Reading, England) 155*, 1749-1757.

5. Archibald, A. R., I. C. Hancock, and C. R. Harwood. (1993) Cell wall structure, synthesis, and turnover, p. 381-410, In A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), Bacillus subtilis and Other Gram-Positive Bacteria, *ASM Press, Washington, DC*.

6. Schleifer, K. H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications, *Bacteriological reviews 36*, 407-477.

7. Bugg, T. D., Wright, G. D., Dutka-Malen, S., Arthur, M., Courvalin, P., and Walsh, C. T. (1991) Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA, *Biochemistry 30*, 10408-10415.

8. Kim, S. J., Cegelski, L., Stueber, D., Singh, M., Dietrich, E., Tanaka, K. S., Parr, T. R., Jr., Far, A. R., and Schaefer, J. (2008) Oritavancin exhibits dual mode of action to inhibit cell-wall biosynthesis in Staphylococcus aureus, *Journal of molecular biology 377*, 281-293.

9. Kim, S. J., Matsuoka, S., Patti, G. J., and Schaefer, J. (2008) Vancomycin derivative with damaged D-Ala-D-Ala binding cleft binds to cross-linked peptidoglycan in the cell wall of Staphylococcus aureus, *Biochemistry 47*, 3822-3831.

10. Werner, G., Klare, I., Fleige, C., and Witte, W. (2008) Increasing rates of vancomycin resistance among Enterococcus faecium isolated from German hospitals between 2004 and 2006 are due to wide clonal dissemination of vancomycin-resistant enterococci and horizontal spread of vanA clusters, *Int J Med Microbiol 298*, 515-527.

11. Tong, G., Pan, Y., Dong, H., Pryor, R., Wilson, G. E., and Schaefer, J. (1997) Structure and dynamics of pentaglycyl bridges in the cell walls of Staphylococcus aureus by 13C-15N REDOR NMR, *Biochemistry 36*, 9859-9866.

12. Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., van Heijenoort, J., and Gutmann, L. (1996) Peptidoglycan structure of Enterococcus faecium expressing vancomycin resistance of the VanB type, *The Biochemical journal 313 (Pt 3)*, 711-715.

13. Pascher, T., Chesick, J. P., Winkler, J. R., and Gray, H. B. (1996) Protein folding triggered by electron transfer, *Science (New York, N.Y 271*, 1558-1560.

14. Glauner, B. (1988) Separation and quantification of muropeptides with high-performance liquid chromatography, *Analytical biochemistry 172*, 451-464.

15. Patti, G. J., Kim, S. J., and Schaefer, J. (2008) Characterization of the peptidoglycan of vancomycin-susceptible Enterococcus faecium, *Biochemistry 47*, 8378-8385.

16. Vollmer, W. (2008) Structural variation in the glycan strands of bacterial peptidoglycan, *FEMS microbiology reviews 32*, 287-306.

17. Pfeffer, J. M., Strating, H., Weadge, J. T., and Clarke, A. J. (2006) Peptidoglycan O acetylation and autolysin profile of Enterococcus faecalis in the viable but nonculturable state, *Journal of bacteriology 188*, 902-908.

18. Rosenthal, R. S., Folkening, W. J., Miller, D. R., and Swim, S. C. (1983) Resistance of O-acetylated gonococcal peptidoglycan to human peptidoglycan-degrading enzymes, *Infection and immunity 40*, 903-911.

19. Swim, S. C., Gfell, M. A., Wilde, C. E., 3rd, and Rosenthal, R. S. (1983) Strain distribution in extents of lysozyme resistance and O-acetylation of gonococcal peptidoglycan determined by high-performance liquid chromatography, *Infection and immunity 42*, 446-452.

20. Clarke, A. J., and Dupont, C. (1992) O-acetylated peptidoglycan: its occurrence, pathobiological significance, and biosynthesis, *Canadian journal of microbiology 38*, 85-91.

21. Hamada, S., Torii, M., Kotani, S., Masuda, N., Ooshima, T., Yokogawa, K., and Kawata, S. (1978) Lysis of Streptococcus mutans cells with mutanolysin, a lytic enzyme prepared from a culture liquor of Streptomyces globisporus 1829, *Archives of oral biology 23*, 543-549.

22. Metcalf, R. H., and Deibel, R. H. (1972) Growth of Streptococcus faecium in the presence of lysozyme, *Infection and immunity 6*, 178-183.

23. Stranden, A. M., Roos, M., and Berger-Bachi, B. (1996) Glutamine synthetase and heteroresistance in methicillin-resistant Staphylococcus aureus, *Microbial drug resistance (Larchmont, N.Y 2*, 201-207.

24. Hanaki, H., Labischinski, H., Inaba, Y., Kondo, N., Murakami, H., and Hiramatsu, K. (1998) Increase in glutamine-non-amidated muropeptides in the peptidoglycan of vancomycin-resistant Staphylococcus aureus strain Mu50, *The Journal of antimicrobial chemotherapy 42*, 315-320.

25. Nakel, M., Ghuysen, J. M., and Kandler, O. (1971) Wall peptidoglycan in Aerococcus viridans strains 201 Evans and ATCC 11563 and in Gaffkya homari strain ATCC 10400, *Biochemistry 10*, 2170-2175.

26. de Jonge, B. L., Gage, D., and Handwerger, S. (1996) Peptidoglycan composition of vancomycin-resistant Enterococcus faecium, *Microbial drug resistance (Larchmont, N.Y 2*, 225-229.

27. Ghuysen, J. M. (1968) Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism, *Bacteriological reviews 32*, 425-464.

28. Ikawa, M. (1964) The Configuration of Aspartic Acid in Cell Walls of Lactic Acid Bacteria and Factors Affecting the Racemization of Aspartic Acid, *Biochemistry 3*, 594-597.

29. Bellais, S., Arthur, M., Dubost, L., Hugonnet, J. E., Gutmann, L., van Heijenoort, J., Legrand, R., Brouard, J. P., Rice, L., and Mainardi, J. L. (2006) Aslfm, the D-aspartate ligase responsible for the addition of D-aspartic acid onto the peptidoglycan precursor of Enterococcus faecium, *The Journal of biological chemistry 281*, 11586-11594.

30. Arias, C. A., Torres, H. A., Singh, K. V., Panesso, D., Moore, J., Wanger, A., and Murray, B. E. (2007) Failure of daptomycin monotherapy for endocarditis caused by an Enterococcus faecium strain with vancomycin-resistant and vancomycin-

susceptible subpopulations and evidence of in vivo loss of the vanA gene cluster, *Clin Infect Dis 45*, 1343-1346.

31. Khan, S. A., Sung, K., Layton, S., and Nawaz, M. S. (2008) Heteroresistance to vancomycin and novel point mutations in Tn1546 of Enterococcus faecium ATCC 51559, *International journal of antimicrobial agents 31*, 27-36.

7. Time-Dependent Isotopic Labeling and Quantitative Liquid-Chromatography/Mass Spectrometry Reveal Bacterial Cell-Wall Architecture

This work was done collaboratively with Gary Patti. Some of it was published in: Patti, G. J., Chen, J. & Gross, M. L. Method Revealing Bacterial Cell-Wall Architecture by Time-Dependent Isotope Labeling and Quantitative Liquid Chromatography/Mass Spectrometry. *Anal. Chem. 2009, 81, 2437-2445(1).* This chapter is a recapitulation of the work with emphasis on the method development.

7.1 Abstract

The molecular details of bacterial cell-wall biosynthesis are essential to understanding cell-wall growth and division as well as the action of glycopeptides antibiotics, but they still remain unclear. Here we report a method combining time-dependent isotopic labeling and liquid-chromatography/mass spectrometry (LC/MS) analysis to track the synthesis and fate of bacterial cell-wall precursors. By comparing isotopic enrichments of post-insertionally modified cell-wall precursor along the growth curve, we can propose the sequence of cell-wall maturation throughout a cell cycle. In *E. faecium* peptidoglycan, the pentapeptides terminating in *D*-Ala-*D*-Ala are the first muropeptides to be isotopically labeled. As these species undergo transpeptidation, the peptide stems not cross-linked and are cleaved into tri and tetrapeptides. Subsequently, some muropeptides become *O*-acetylated, and ultimately the cross-bridges are cleaved on stems not cross-linked. This method will advance the understanding the tertiary structure of bacteria cell wall and assist with the design of novel drugs for antibiotic-resistant pathogens.

7.2 Introduction

We developed an LC/MS based method to characterize the fine structure of bacterial peptidoglycan (*2*), described in the previous chapter. The molecular details regarding the organization and post-insertional modification of the antibiotic-binding sites in the peptidoglycan, although important to understanding the architecture of the cell wall, remain unclear. Here, we present a method based on mass spectrometric quantitation of pulse-labeled cells to track the fate of bacterial peptidoglycan, thus, providing important information that allows elucidation of cell-wall morphogenesis.

Isotopic labeling combined with mass spectrometry analysis has been widely used in proteomics studies. One example is the stable isotope labeling with amino acids in cell culture (SILAC) (*3, 4*). SILAC is a straightforward approach for *in vivo* incorporation of a label into cellular proteins for MS-based quantitative proteomics. In such experiment, two cell populations are grown in culture media that are identical except that one of them contains a "light" and the other a "heavy" form of a particular amino acid. Then, the two cell populations are combined and analyzed by mass spectrometry, taking advantage of the mass difference between pairs of chemically identical peptides with different stable-isotope composition. For kinetics studies of biological systems, pulse-chase experiments based on stable isotope labeling in culture and monitoring by mass spectrometric measurement were reported previously (*5*). In this experiment, cells were grown in the heavy media for a certain period of time, and then the media is changed to light media, where the growth continues. The labeling is followed by sample collection throughout the chase period.

In this work, we applied mass spectrometry-based quantitation of pulse-labeled cells by using an online LC/MS analysis that we described in the previous chapter, to investigate the sequence in which *E. faecium* peptidoglycan subunits containing vancomycin-binding sites are post-insertionally modified in cell-wall growth and maturation. Muropeptide profiles taken with respect to the length of the heavy-isotope pulse provide snapshots of the progression of peptidoglycan species throughout the cell cycle.

Mass-spectrometry-based quantitation of pulse-labeled cells is particularly effective for investigating the *E. faecium* cell wall given the high sensitivity and low sample consumption of the nano-LC/MS method. The method can accommodate the significant number of peptidoglycan structural variations and the low frequency of cross-linked species larger than dimers. Furthermore, the biosynthesis of *E. faecium* peptidoglycan is of great clinical interest as its understanding is important in the development and understanding of novel antibiotics (*6-8*). Using post-insertional peptidoglycan modifications as a probe to track murein growth and assembly, we describe the extension of our understanding of cell-wall biosynthesis and architecture.

7.3 Experimental

7.3.1 Time-dependent Isotopic Labeling

Starting cultures of *E. faecium* (ATTC 49624) were prepared by inoculating brain-heart infusion media with a single colony. Cultures were incubated overnight at 37 °C, but not aerated. The growth was started by inoculating brain-heart infusion (BHI) media with the overnight starter culture (1% final volume).

Enterococcal standard media (ESM) was prepared as described before (*9*). With the pH adjusted to 7.0 prior to sterile filtration, the natural abundance *L*-lysine was replaced with uniformly labeled *L*-[$^{13}C_6$,$^{15}N_2$]lysine. The cells were extracted from BHI media when the absorbance at 660 nm had become approximately 0.3 by centrifugation at 10,000 × g. Cells were then re-suspended in the pre-warmed ESM containing *L*-[$^{13}C_6$,$^{15}N_2$]lysine and incubated (Figure 7.1, pulse). *L*-Lysine was a good choice because it has high enrichment efficiency in *E. faecium* and was not readily metabolized (*9*). The expression "pulse labeling" in this text to be consistent with established convention, but the term "step" is a more appropriate description of the labeling strategy. Cells were harvested at various time points (see Figure 7.1, harvest points) by centrifugation at 10,000g and boiled for 5 min in 40 mM triethanolamine hydrochloride (pH 7.0) to quench biological activity before being frozen and lyophilized. Isotopic enrichment is time-independent after cells are harvested and boiled. Enrichment measurements were made immediately after sample preparation and were within ±0.5% of the values determined from the same sample 20 days after harvest. Ten time points were acquired to investigate cell-wall development over approximately one enterococcal cell cycle (estimated as 3 h after the pulse for these conditions). For pulse-chase samples, the cells were removed from the *L*-[$^{13}C_6$,$^{15}N_2$]lysine enriched ESM after a defined time duration (Figure 7.1, *i*) by centrifugation at 10,000g, resuspended in prewarmed BHI, incubated, harvested, and boiled at a later time (Figure 7.1, *f*). The initial and final time points were chosen such that the cells were chased over the latter half of exponential growth. Cell-wall isolates were prepared from lyophilized whole cells as described previously (*10*).

The isolated cell walls were digested with lysozyme and mutanolysin to give muropeptides (*11, 12*). All manipulations were performed at neutral pH.

7.3.2 LC/MS Analysis

Liquid chromatography/MS was performed by using a Pico-View PV-500 (New Objective, Woburn, MA) nanospray stage attached to an LTQ-Orbitrap mass spectrometer (ThermoFisher, San Jose, CA). The high-mass accuracy of the Orbitrap mass spectrometer was important for identifying structures of chemically modified species with low ambiguity. With the use of accurate-mass measurements, the formulae of approximately 90% of the components seen in the mass spectra occurring within the range of expected muropeptide retention times were determined, structures proposed, and isotopic enrichments calculated for each.

Muropeptide samples were loaded into an uncoated 75 μm i.d. fused-silica capillary column with a 15 μm picofrit tip (New Objective, Woburn, MA), packed with C18 reverse-phase material (3 μm, 100 Å; Phenomenex, Torrance, CA). The column was eluted with a flow rate of 250 nL/min for 10 min with 0.1% (vol/vol) formic acid in water and subsequently with a 60 min linear acetonitrile gradient (0%-40%) with 0.1% formic acid. The components, as they emerged from the column, were sprayed into an LTQ-Orbitrap mass spectrometer. The spray voltage was 2.0 kV. The capillary voltage and temperature were 27 V and 200 °C, respectively. Full mass spectra were recorded in the FT orbitrap component of the instrument at 60,000 resolving power (at m/z = 400).

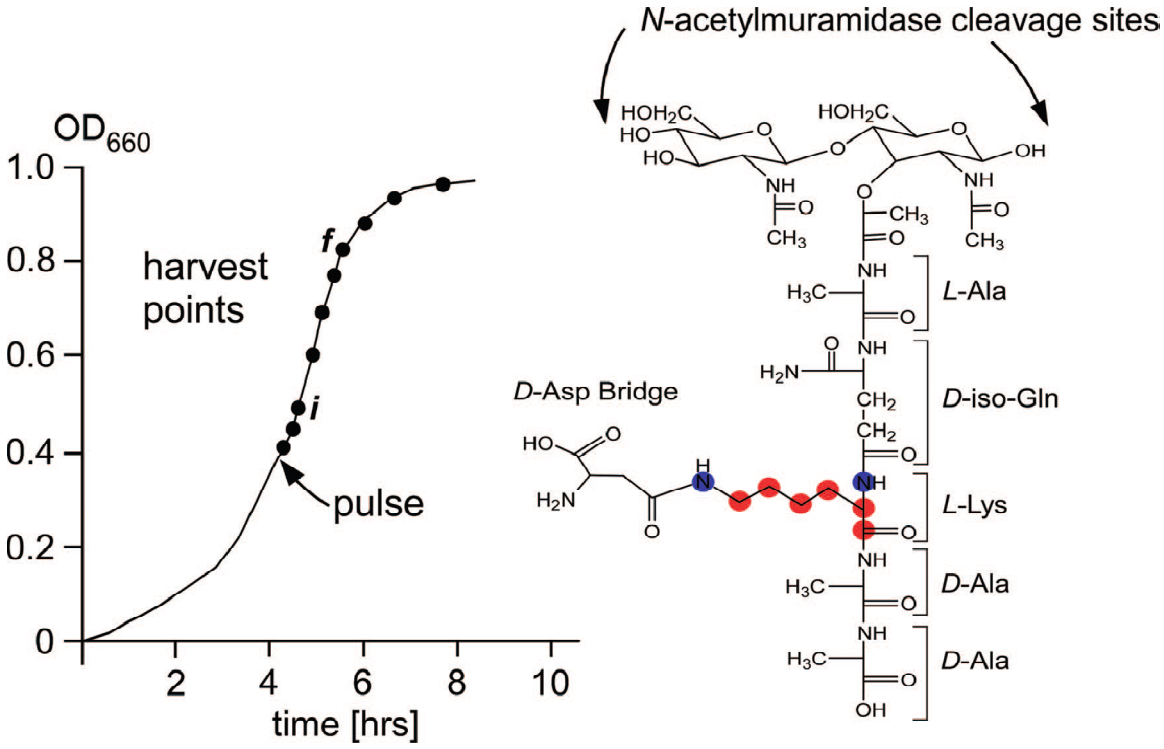


Figure 7.1: Labeling strategy: (left) growth of *E. faecium* as measured by optical density (660 nm) as a function of time. The cells were pulsed and harvested at the points indicated. For the pulse-chase experiment, cells were resuspended in media containing natural-abundance isotopes at point *i* and harvested at point *f*. (Right) Chemical structure of *E. faecium* peptidoglycan before modification, highlighting *L*-[$^{13}C_6$,$^{15}N_2$]lysine. The unmodified subunit contains a pentapeptide stem and a *D*-Asp bridge. Figure is reproduced from reference (1).

7.4 Results and Discussion

7.4.1 Semi-quantitative Analysis of Muropeptides

We identified the majority of resulting muropeptides with LC/MS accurate-mass measurements as variations of monomers, dimers, and a few trimers. The assignments of all muropeptides are consistent with the proposed structures presented in Scheme 6.1 in the previous chapter.

The total-ion chromatograms of muropeptides digested from cells that were harvested at three different times (Figure 7.2) are the basis for quantification of each muropeptide species. Quantification was determined by integrating extracted-ion chromatograms for singly and multiply charged ions. Consistent with the observation by others for a different organism (*13*), we detected no substantive differences in *E. faecium* cell-wall composition as cultures progressed from the exponential- to stationary-growth phase (Figure 7.3), although the amount of muropeptides varies among samples. Thus, the cross-linking remained relatively constant. This approach of estimating cross-linking yields results in good agreement with those obtained from solid-state NMR, as established previously (*9*).

7.4.2 Isotopic Labeling and MS Quantitation

We incorporated stable-isotopic labels into the bacterial growth cycles to investigate the dynamics of peptidoglycan processing and to track the development of specific muropeptides. By removing the cells growing in the natural-abundance media and re-suspending them in defined media enriched with *L*-[$^{13}C_6$,$^{15}N_2$]lysine, we had "pulse-labeled" the cells before quenching and analyzing. Muropeptides incorporating

n *L*-[$^{13}C_6$,$^{15}N_2$]lysines result in a mass increase of 8*n*, which reflects in the mass spectrum as a peak pattern 8*n*/*z* larger than the unlabeled one. Although quantitative comparison of peaks in ESI mass spectra are generally unreliable between different species because there is variability in efficiencies for ionization, the approach developed here avoids this problem given that molecular species carrying isotopes do not have inherently different ionization efficiencies (*3, 14*). Although the proportion of monomers, dimers and trimers did not significantly change throughout the cell cycle, the percentage of these species incorporating one or more *L*-[$^{13}C_6$,$^{15}N_2$]lysine isotopes was time-dependent. We calculated the enrichment percentages by considering the proportion of muropeptide structures incorporating *L*-[$^{13}C_6$,$^{15}N_2$]lysine with respect to the sum of those same structures with and without *L*-[$^{13}C_6$,$^{15}N_2$]lysine.

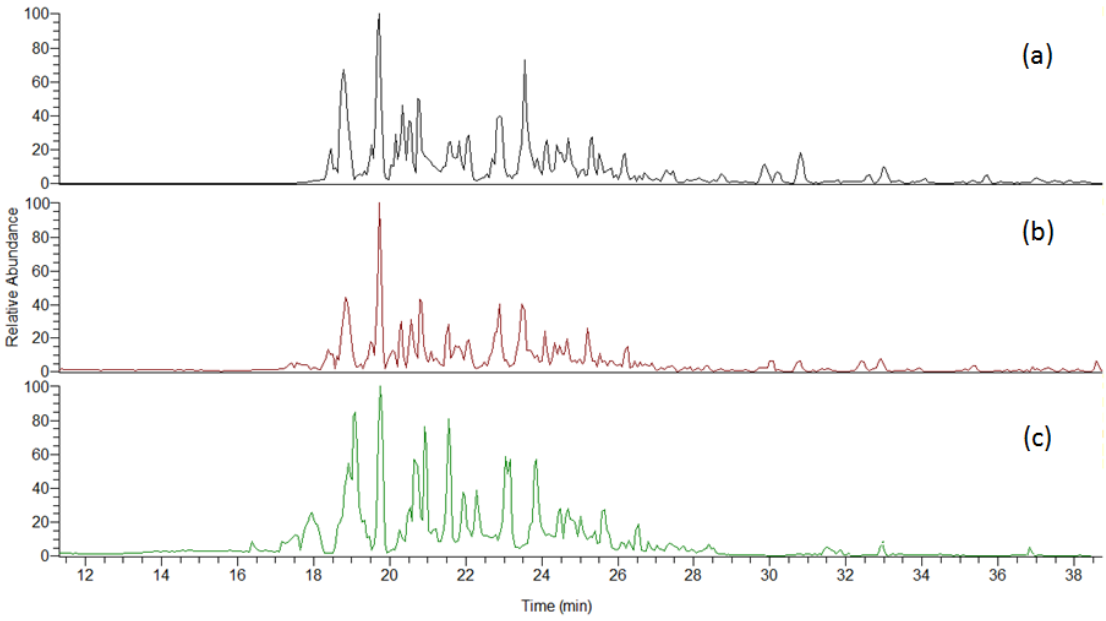


Figure 7.2: Total-ion chromatograms of peptidoglycan digests after (a) 11 min, (b) 70 min and (c) 202 min growing in heavy lysine media.

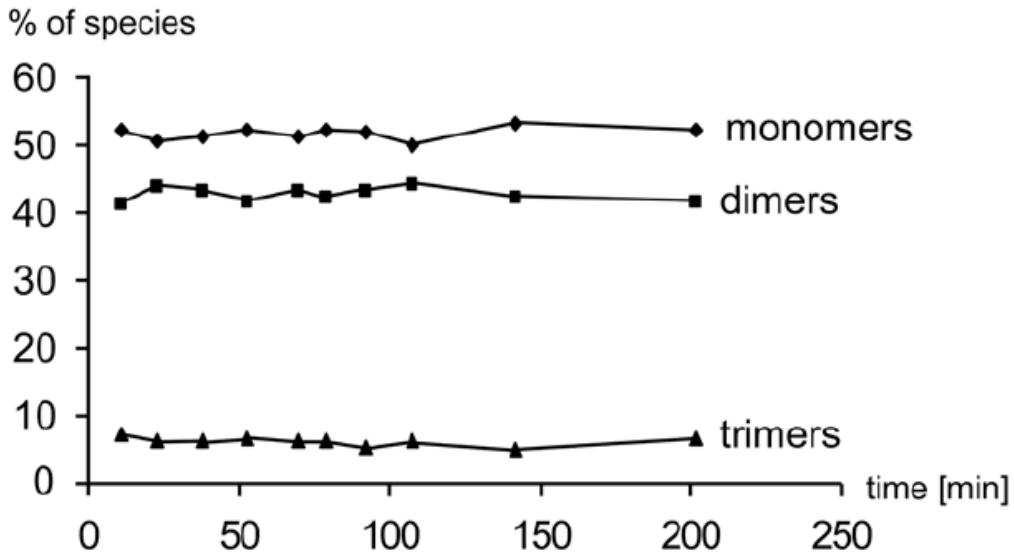


Figure 7.3: Distribution of monomers, dimers, and trimers with respect to total peptidoglycan.

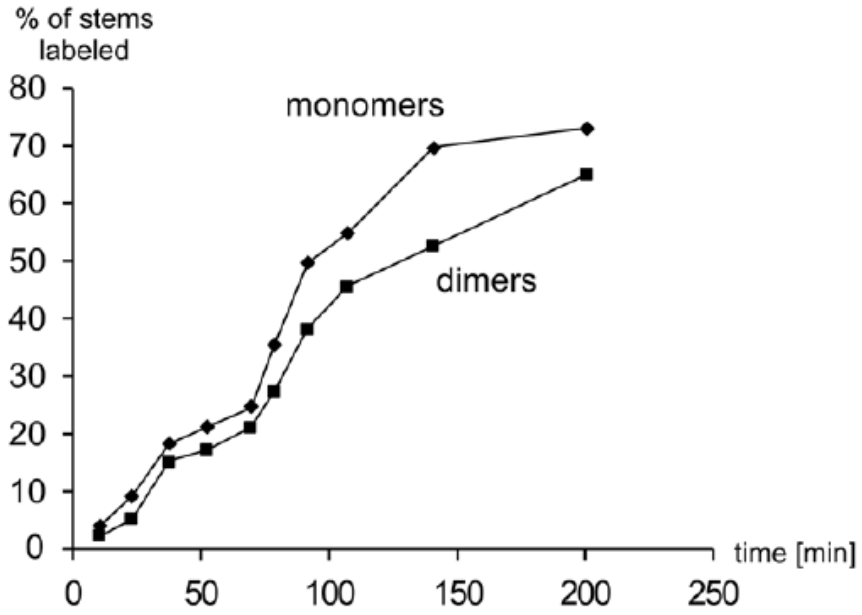


Figure 7.4: Plot of the average percentage of isotopically enriched monomers and dimers with respect to time after pulse. Trimers are excluded given their low frequency.

We suggest that monomers are the first muropeptides to incorporate *L*-[$^{13}C_6$,$^{15}N_2$]lysine, followed by dimers and trimers (Figure 7.4). At any given time point throughout the growth curve, the average enrichment for monomers is higher than that for dimers. For example, after 38 min in labeled media, monomers and dimers are enriched 19% and 14%, and after 143 min, 70% and 52%, respectively. The difference of enrichment between monomers and dimers is small in the early stage of growth because transpeptidation occurs early in the peptidoglycan synthesis. The difference becomes larger in the late stages of growth, likely owing to the slow-down of transpeptidation and the accumulation of isotopically labeled monomers. Trimer species are not discussed here, owing to their low frequency of occurrence.

m/z	species	final % enriched[a]
1081.491	G-M-Ala-iGln-Lys(Asp)-Ala-Ala	90
1010.454	G-M-Ala-iGln-Lys(Asp)-Ala	83
939.416	G-M-Ala-iGln-Lys(Asp)	81
1080.507	G-M-Ala-iGln-Lys(Asn)-Ala-Ala	75
938.432	G-M-Ala-iGln-Lys(Asn)	73
1009.470	G-M-Ala-iGln-Lys(Asn)-Ala	61
895.426	G-M-Ala-iGln-Lys-Ala	55
824.388	G-M-Ala-iGln-Lys	52

[a] For three different trials, the isotopic enrichments for all species described in this text were determined as ±0.5% of the values provided (e.g., 90 ± 0.5%, 83 ± 0.5%, etc.).

Table 7.1. Percentages of isotopic enrichment after 202 minutes growth in *L*-[$^{13}C_6$,$^{15}N_2$]lysine enriched media. Table is reproduced from reference (1).

7.4.3 Penta, Tetra, and Tripeptides

Modifications occurring at the peptide stem of the peptidoglycan muropeptides are of particular interest because the three amino acids are involved in transpeptidation. In forming a cross-link, the penultimate *D*-Ala of one stem forms an amide bond with a *D*-Asx bridge in another stem, and the terminal *D*-Ala is removed by *D,D*-carboxypeptidase. Here we want to study the modification removing the terminal *D*-Ala, the penultimate *D*-Ala, or the *D*-Asx bridge that occur independent of transpeptidation. We choose to present results for monomer species to simplify the discussion, although dimers and trimers not cross-linked follow the similar trend.

Among all the monomers, the species with *D*-Ala-*D*-Ala terminus and a *D*-Asp bridge (Figure 7.5, of *m/z* = 1081.491) is isotopically enriched to the highest extent, indicating that it is the first species incorporated into the peptidoglycan, consistent with the observations made by others (*15*). At the first time point after the media switch, however, this species is isotopically enriched but lower than the species with a *D*-Asn bridge (Figure 7.5, of *m/z* = 1080.507). This is also observed for other species with a bridge. We suggest, therefore, the amidation of *D*-Asp occurs rapidly in the beginning of cell division.

The tripeptide species without a bridge is the slowest to be enriched after the introduction of the *L*-[$^{13}C_6$,$^{15}N_2$]lysine pulse and is enriched to the lowest extent (Figure 7.5, of *m*/*z* = 824.388). The tetrapeptide without a bridge is the second lowest enriched muropeptide (Figure 7.6, bottom, of *m/z* = 895.426). These species show minimal isotopic enrichment at the first time point, indicating that they are processed from other

muropeptide species instead of being incorporated to the peptidoglycan as is. It also suggests the presence of an enzyme, which removes the bridge that is not cross-linked to other stems, in the cell wall.

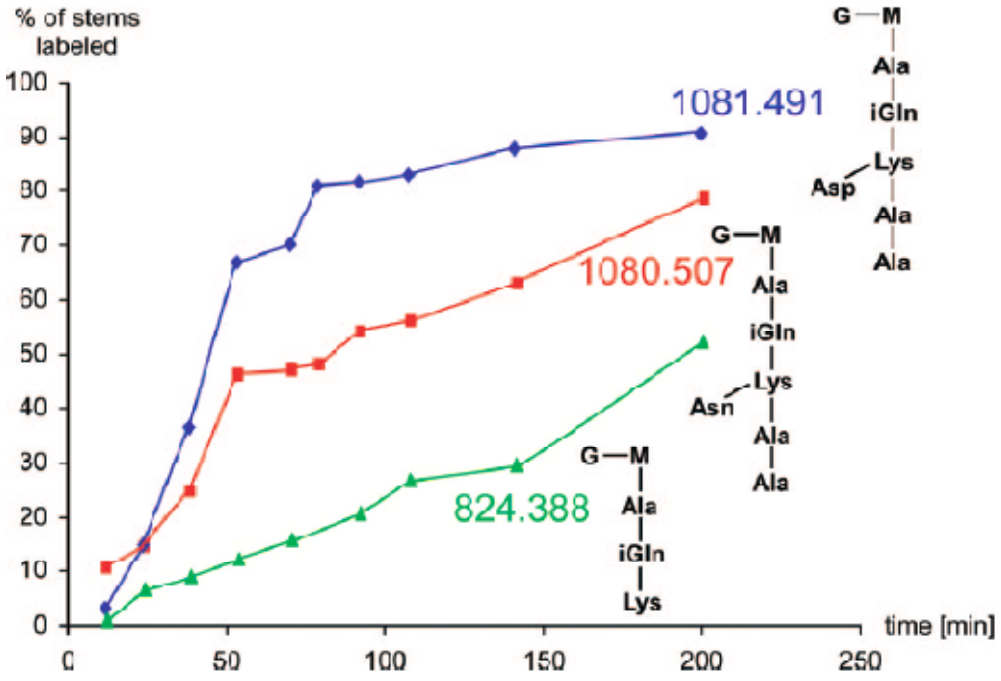


Figure 7.5: Plot of the percentage of isotopically enriched ions with respect to time after pulse. (Diamond) Ions with an *m*/*z* of 1081.491. (Square) Ions with an *m*/*z* of 1080.507. (Triangle) Ions with an m/z of 824.388.

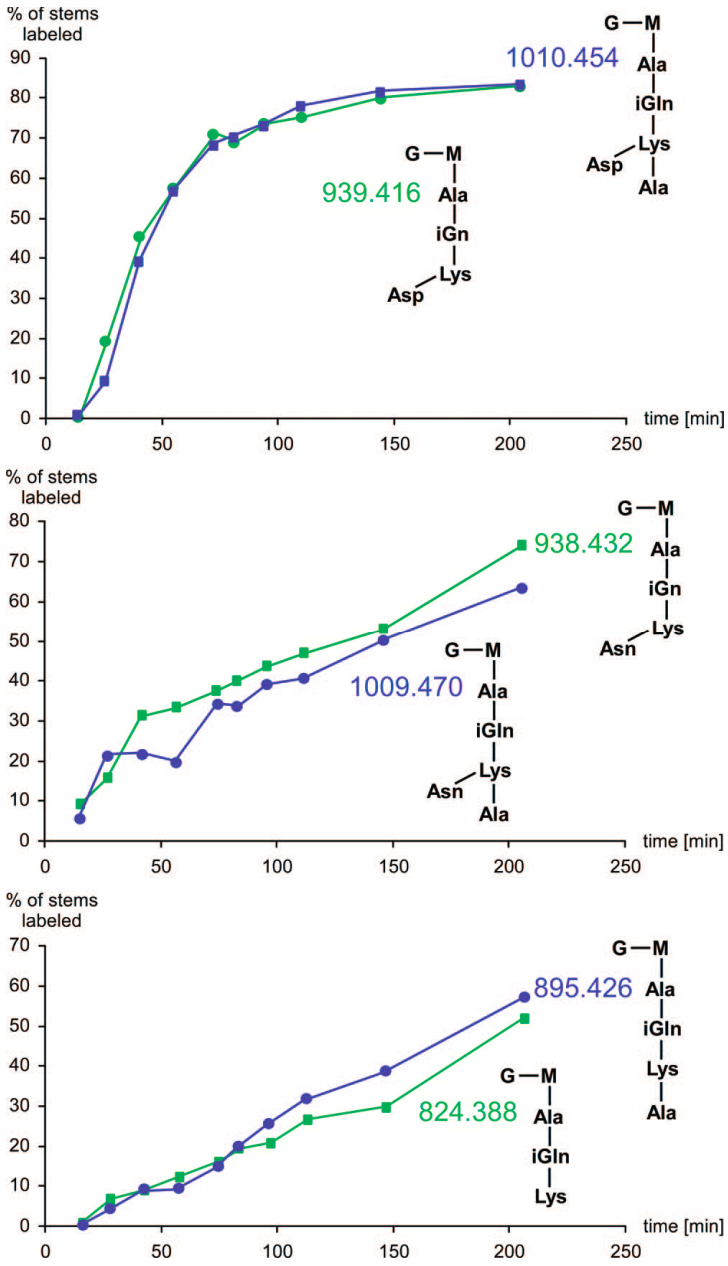


Figure 7.6: Plots of the percentage of isotopically enriched tri (square) and tetrapeptides (circle) with respect to time after pulse. (Top) Ions with an *m*/*z* of 1010.454 and 939.416, corresponding to muropeptides with an aspartic-acid bridge. (Middle) Ions with an *m*/*z* of 938.432 and 1009.470, corresponding to muropeptides with an asparagine bridge. (Bottom) Ions with an *m*/*z* of 895.426 and 824.388, corresponding to muropeptides without bridges.

The presence of tetra and tripeptides suggests there is an active *D*,*D*-carboxypeptidase and *L*,*D*-carboxypeptidase in the cell wall of *E. faecium*. *D*,*D*-carboxypeptidase hydrolyzes the peptide bond between the two *D*-Ala in the peptide terminus and *L*,*D*-carboxypeptidase cleaves the peptide bond between *L*-Lys and *D*-Ala. Not much is known about whether these two enzymatic processes are related to each other. The data show that the tetrapeptide (of m/z = 1009.470) is enriched less than the tripeptide (of m/z = 938.432) (Figure 7.6, middle), both of which are less enriched than the pentapeptide. The differences between the tri- and tetrapeptides, although not large, are statistically significant considering the small error in this experiment. Thus, it is unlikely that *D*,*D*-carboxypeptidase creates the substrate for *L*,*D*-carboxypeptidase. Instead, we propose that the reactions producing tripeptides and tetrapeptides are independent. The tripeptide (of m/z = 939.416) and tetrapeptide (of m/z = 1010.454) are enriched nearly equally (Figure 7.6, top), indicating that both *D*,*D*-carboxypeptidase and *L*,*D*-carboxypeptidase are targeting the same pentapeptide substrate, a conclusion that also agrees with the above hypothesis. For muropeptides without bridges (Figure 7.6, bottom), however, the tetrapeptide has a larger percentage of enrichment than the tripeptide, suggesting a different sequence of processing mature cell wall or greater activity of *D*,*D*-carboxypeptidase towards murein without bridges.

7.4.4 *O*-Acetylation

Postinsertional modifications of *E. faecium* peptidoglycan result in a number of structural variations, described in the previous chapter (*2*). The order of such modifications, however, could not be revealed by bottom-up MS analysis of the peptidoglycan digests. Thus, we analyzed the peptidoglycan of bacterial cells harvested

from enriched media over 10 different time intervals to investigate the trends in muropeptide variations and to study the dynamics thrughout an entire cell cycle.

It was suggested in the previous study that *O*-acetylation, a post-insertional modification giving rise to a mass increase of 42.011, occurs at C6 of *N*-acetylmuramic acid in mature cell walls (*16*). We determined the isotopic enrichments of *O*-acetylated muropeptides throughout the growth curve by averaging all the percentages of labeling in all species at 10 time points. Some muropeptides were rarely *O*-acetylated, consistent with the results obtained from cells in late exponential growth, discussed in the previous chapter (*2*). As shown in Figure 7.7, there is a lag in the beginning of the cell growth where the isotopic enrichment of *O*-acetylated muropeptides is low. The enrichment starts to increase more rapidly after cells have sufficient time to produce mature cell wall from the *L*-[$^{13}C_6$,$^{15}N_2$]lysine. The trend of enrichment for *O*-acetylation is similar to that for tri- and tetrapeptide without a bridge, indicating that these species are located in the mature cell wall.

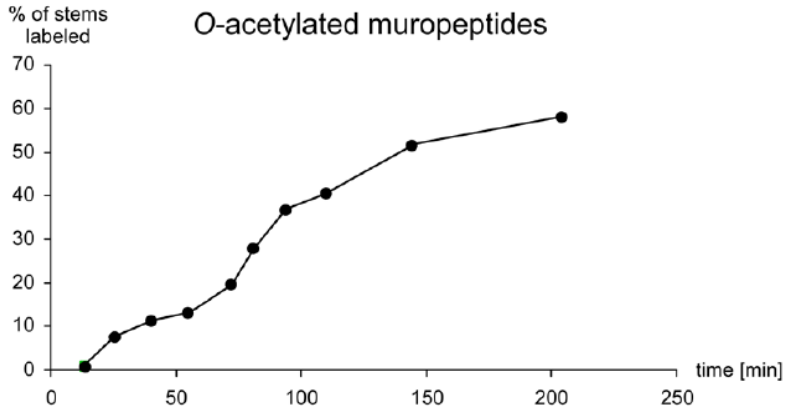


Figure 7.7: Plot of the average percentage of isotopically enriched *O*-acetylated muropeptides with respect to time after pulse.

7.4.5 Pulse-Chase Experiment

A pulse-chase experiment was conducted to provide further support for the conclusion that the amount of pentapeptide species is small. After growing cells for 38 min in labeled media, the average percentage of isotopically enriched pentapeptides (of m/z = 1080.507 and 1081.491) is 30%. The enrichment increases to about 75% when growing continuously in the labeled media for a total of 92 min. When the *L*-[$^{13}C_6$,$^{15}N_2$]lysine pulse is chased with media containing natural-abundance *L*-lysine for 54 min after the 38 min pulse, the percentage of isotopic enrichment is decreased to less than 2% (Figure 7.8). This decrease in enrichment suggests that the isotopically enriched pentapeptides are rapidly converted into other peptidoglycan species, being replaced by new subunits that are not enriched with *L*-[$^{13}C_6$,$^{15}N_2$]lysine. Thus, the isotopic enrichment is constantly being diluted over the course of the chase. The tripeptides (of m/z = 824.388), however, show little difference in the isotopic enrichment before and after the introduction of the chase. The lack of difference is consistent with our hypothesis that tripeptides without bridges are present in the mature cell wall where the *L*-[$^{13}C_6$,$^{15}N_2$]lysine is not diluted during the chase.

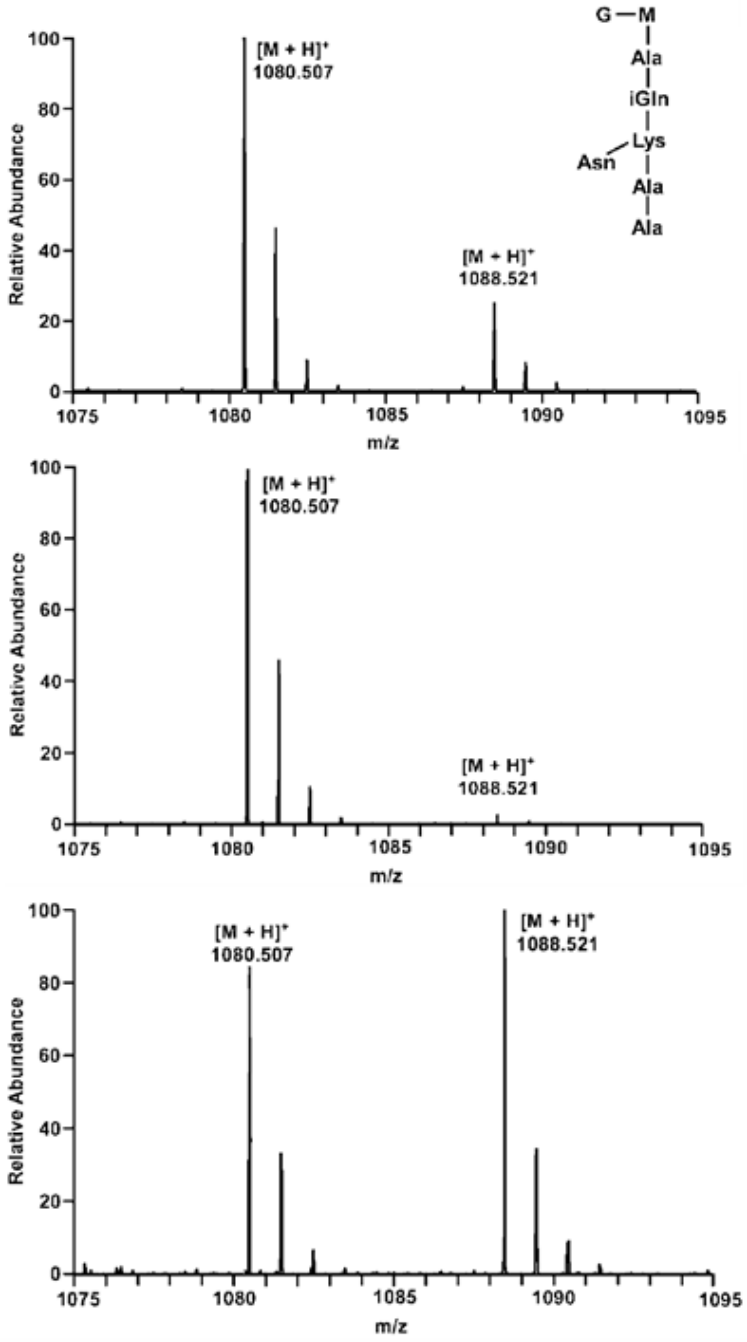


Figure 7.8: Mass spectra for an ion with an *m*/*z* of 1080.507 corresponding to a pentapeptide with a D-Asn bridge. (Top) Cells harvested and analyzed after a 38 min pulse, (middle) cells chased for 54 min with media containing only natural-abundance isotopes after an initial 38 min pulse, and (bottom) cells labeled continuously for 92 min.

7.4.6 Glycan Chains

There are three models regarding the arrangement of glycan chains in bacterial cell walls, and they are parallel (*17*), perpendicular (*18*), or randomly aligned (*19*) with respect to the cell membrane. Here we provide more experimental evidence to support the parallel model, based on the change of enrichment pattern of muropeptide dimers.

The isotopic enrichment of dimers could potentially reveal the organization of glycan chains because the constitution of dimers with respect to old and new peptidoglycan subunits varies in different models. When the glycan chains are perpendicular to the cell membrane, as suggested by the scaffold model, new peptidoglycan subunits are added at the ends of the glycan chains in the same plane. The dimers, formed from muropeptide in the same plane by cross-linking, are all newly incorporated (*18*). In this case, both *L*-lysine residues in dimers would be isotopically labeled after the introduction of the pulse. Our data show, however, that after 11 min in the heavy media, dimers with only one *L*-[$^{13}C_6$,$^{15}N_2$]lysine are observed (Figure 7.9, top, of m/z = 968.951), which does not agree with the scaffold model. The percentage of singly labeled dimers remains relatively constant throughout the growth (~10%). In contrast, doubly labeled dimers (Figure 7.9, of m/z = 972.958) represent a small fraction in the first time point but increase during the course of labeling, and reach ~60% at the last time point.

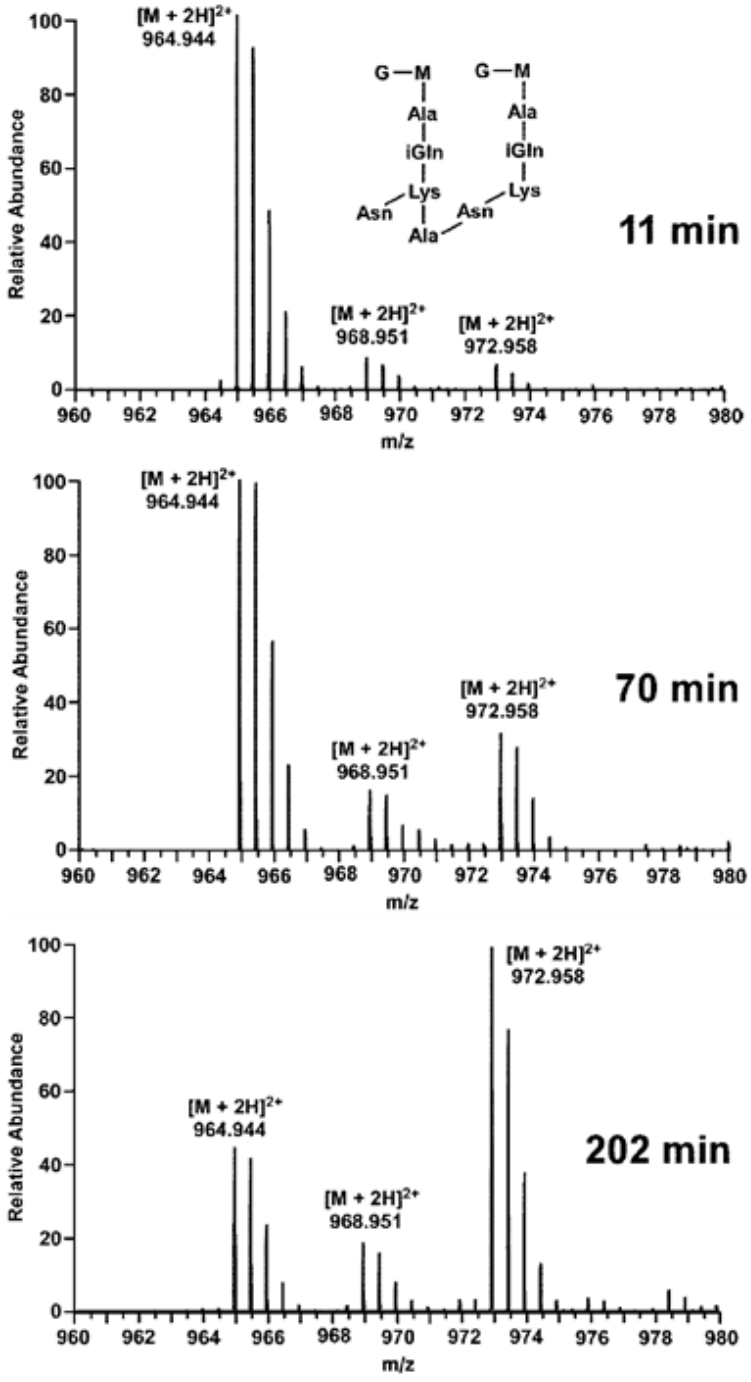


Figure 7.9: Mass spectra for the ion with an *m*/*z* of 964.944 corresponding to a doubly charged dimer with Asn bridges. (Top) After an 11 min pulse, (middle) after a 70 min pulse, and (bottom) after a 202 min pulse. The number of singly labeled dimers incorporating one *L*-[$^{13}C_6$,$^{15}N_2$]lysine remains relatively constant at all three time points.

The data are consistent with the parallel model where glycan chains are incorporated into the existing peptidoglycan via single-strand addition. The first newly synthesized glycan chain incorporating *L*-[$^{13}C_6$, $^{15}N_2$]lysine forms cross-links with the existing parallel glycan chain containing only natural-abundance amino acids. Thus, the dimers formed in between these chains contain only one *L*-[$^{13}C_6$,$^{15}N_2$]lysine. The amount of these dimers remains the same after this interface is formed, as indicated in Figure 7.9 (middle and bottom). We propose that the number of peptidoglycan layers may be estimated based on the model and the fraction of singly labeled dimers.

7.5 Conclusion

The combination of time-dependent isotopic labeling and mass spectrometry-based quantitation is shown to be a widely applicable approach for tracking the fate of molecules in bacterial cell walls. In this chapter, we describe the application of this method to understand the biosynthesis of bacterial peptidoglycan and to study the dynamics of peptidoglycan biosynthesis and maturation. The results reveal the order in which peptidoglycan subunits are post-insertionally modified to different structures. These modifications, although not fully understood with regard to their function, may relate to cell growth and division, and could be significant in microbial pathogenesis. The data here provide further evidence for the template model proposed by Schaefer and co-workers (*20, 21*), and extend the model by incorporating dynamic information. This method, in addition to the bottom-up MS strategy described in the previous chapter, may be utilized in a general approaches to study the peptidoglycan of any bacterial stains. As one of the applications, the investigation of peptidoglycan in *fem*-deletion mutants of methicilin-resistant *staphylococcus aureus* is discussed in the next chapter.

7.6 References

1. Patti, G. J., Chen, J., and Gross, M. L. (2009) Method revealing bacterial cell-wall architecture by time-dependent isotope labeling and quantitative liquid chromatography/mass spectrometry, *Analytical chemistry 81*, 2437-2445.
2. Patti, G. J., Chen, J., Schaefer, J., and Gross, M. L. (2008) Characterization of structural variations in the peptidoglycan of vancomycin-susceptible Enterococcus faecium: understanding glycopeptide-antibiotic binding sites using mass spectrometry, *Journal of the American Society for Mass Spectrometry 19*, 1467-1475.
3. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol Cell Proteomics 1*, 376-386.
4. Ong, S. E., and Mann, M. (2007) Stable isotope labeling by amino acids in cell culture for quantitative proteomics, *Methods in molecular biology (Clifton, N.J 359*, 37-52.
5. Milner, E., Barnea, E., Beer, I., and Admon, A. (2006) The turnover kinetics of major histocompatibility complex peptides of human cancer cells, *Mol Cell Proteomics 5*, 357-365.
6. Aumeran, C., Baud, O., Lesens, O., Delmas, J., Souweine, B., and Traore, O. (2008) Successful control of a hospital-wide vancomycin-resistant Enterococcus faecium outbreak in France, *Eur J Clin Microbiol Infect Dis 27*, 1061-1064.

7. Libisch, B., Lepsanovic, Z., Top, J., Muzslay, M., Konkoly-Thege, M., Gacs, M., Balogh, B., Fuzi, M., and Willems, R. J. (2008) Molecular characterization of vancomycin-resistant Enterococcus spp. clinical isolates from Hungary and Serbia, *Scandinavian journal of infectious diseases 40*, 778-784.

8. Park, I. J., Lee, W. G., Shin, J. H., Lee, K. W., and Woo, G. J. (2008) VanB phenotype-vanA genotype Enterococcus faecium with heterogeneous expression of teicoplanin resistance, *Journal of clinical microbiology 46*, 3091-3093.

9. Patti, G. J., Kim, S. J., and Schaefer, J. (2008) Characterization of the peptidoglycan of vancomycin-susceptible Enterococcus faecium, *Biochemistry 47*, 8378-8385.

10. Tong, G., Pan, Y., Dong, H., Pryor, R., Wilson, G. E., and Schaefer, J. (1997) Structure and dynamics of pentaglycyl bridges in the cell walls of Staphylococcus aureus by 13C-15N REDOR NMR, *Biochemistry 36*, 9859-9866.

11. Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., van Heijenoort, J., and Gutmann, L. (1996) Peptidoglycan structure of Enterococcus faecium expressing vancomycin resistance of the VanB type, *The Biochemical journal 313 (Pt 3)*, 711-715.

12. de Jonge, B. L., Gage, D., and Handwerger, S. (1996) Peptidoglycan composition of vancomycin-resistant Enterococcus faecium, *Microbial drug resistance (Larchmont, N.Y 2*, 225-229.

13. Signoretto, C., Lleo, M. M., Tafi, M. C., and Canepari, P. (2000) Cell wall chemical composition of Enterococcus faecalis in the viable but nonculturable state, *Applied and environmental microbiology 66*, 1953-1959.

14. Mann, M. (2006) Functional and quantitative proteomics using SILAC, *Nature reviews 7*, 952-958.

15. Bellais, S., Arthur, M., Dubost, L., Hugonnet, J. E., Gutmann, L., van Heijenoort, J., Legrand, R., Brouard, J. P., Rice, L., and Mainardi, J. L. (2006) Aslfm, the D-aspartate ligase responsible for the addition of D-aspartic acid onto the peptidoglycan precursor of Enterococcus faecium, *The Journal of biological chemistry 281*, 11586-11594.

16. Pfeffer, J. M., Strating, H., Weadge, J. T., and Clarke, A. J. (2006) Peptidoglycan O acetylation and autolysin profile of Enterococcus faecalis in the viable but nonculturable state, *Journal of bacteriology 188*, 902-908.

17. Holtje, J. V. (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli, *Microbiol Mol Biol Rev 62*, 181-203.

18. Dmitriev, B. A., Toukach, F. V., Schaper, K. J., Holst, O., Rietschel, E. T., and Ehlers, S. (2003) Tertiary structure of bacterial murein: the scaffold model, *Journal of bacteriology 185*, 3458-3468.

19. Koch, A. L. (1988) Biophysics of bacterial walls viewed as stress-bearing fabric, *Microbiological reviews 52*, 337-353.

20. Kim, S. J., Cegelski, L., Stueber, D., Singh, M., Dietrich, E., Tanaka, K. S., Parr, T. R., Jr., Far, A. R., and Schaefer, J. (2008) Oritavancin exhibits dual mode of action to inhibit cell-wall biosynthesis in Staphylococcus aureus, *Journal of molecular biology 377*, 281-293.

21. Kim, S. J., Matsuoka, S., Patti, G. J., and Schaefer, J. (2008) Vancomycin derivative with damaged D-Ala-D-Ala binding cleft binds to cross-linked

peptidoglycan in the cell wall of Staphylococcus aureus, *Biochemistry 47*, 3822-3831.

8. Mass Spectrometric Investigation of Peptidoglycan in *Fem*-deletion Mutants of Methicilin-resistant *Staphylococcus aureus*

8.1 Abstract

Mass spectrometric analysis has been performed on the peptidoglycan of a wild-type methicillin-resistant *Staphylococcus aureus* and its *fem*-deletion mutants, FemA and FemB. The bridge length changed from 4.3 glycyl residues (wide-type) to 2.2 (FemB) and 0.9 (FemA) and the cross-link decreased. Heterogeneous peptidoglycan fragments were observed in FemA and FemB mutants, likely due to the incomplete knockout of the relevant genes. Isotopic labeling experiment conducted on FemA reveals the independent processes of post-insertional modifications. The overall structural characterization suggests that the tertiary peptidoglycan structure in FemB and that in FemA represents a difference in the peptidoglycan structure of bacteria with long bridges and those with short bridges.

8.2 Introduction

Staphylococcus aureus is one of the most common causes of noscomial infections. Methicillin-resistant *S. aureus* (MRSA), in particular, is a leading pathogen presenting clinical challenge due to its resistant to most antibiotics (*1*). Vancomycin, "the drug of last resort", was effective towards infections caused by multiply resistant MRSA but has become less effective due to the emergence of vancomycin-resistant *S. aureus (2)*. Methicillin resistance in *S. aureus* is mainly caused by the methicillin-resistance determinant *mecA*, the gene for PBP2', an alternate penicillin-binding protein that has lower binding affinity (*3*). The resistance is also dependant on the *fem* (factors essential for methicillin resistance) factors through cell-wall metabolism (*4, 5*). Among the *fem* factors, *femX*, *femA* and *femB*, which encode proteins FemX, FemA and FemB, play important roles in the formation of the pentaglycine in the peptidoglycan (*6-8*). FemX (*6*) initiates the incorporation of the first glycyl unit, followed by the addition of Gly2-Gly3 by FemA (*7, 9*) and Gly4-Gly5 by FemB (*10*). *FemAB* inactivation results in a FemAB mutant that reduces the peptide to a monoglycine (*11*).

The deletion of *femA* and *femB* leads to the formation of peptidoglycan precursors that contain monoglycine and triglycine as the peptide bridge (*4, 5, 12*), thus may alter the tertiary structure of the *S. aureus* peptidoglycan. One of our major goals is to compare the peptidoglycan composition, analyzed by bottom-up mass spectrometry strategy, of these *fem*-deletion mutants with regular *S. aureus* that has a five-glycine bridge and *E. faecium* with one aspartic acid bridge and help gain better understanding of the three-dimensional structure of *S. aureus* peptidoglycan with difference bridge lengths. The previous solid-state NMR results show that the average bridging segment lengths are

5.0, 2.5 and 1.0 glycyl residues for wilde-type, *FemB* and *FemA*, respectively and the level of cross-linking decreases from FemB to FemA, suggesting a higher concentration of open glycyl segments (*13*). Another goal here is to characterize fine structures to support and complement NMR data by applying the mass-spectrometry based methods described in the last two chapters.

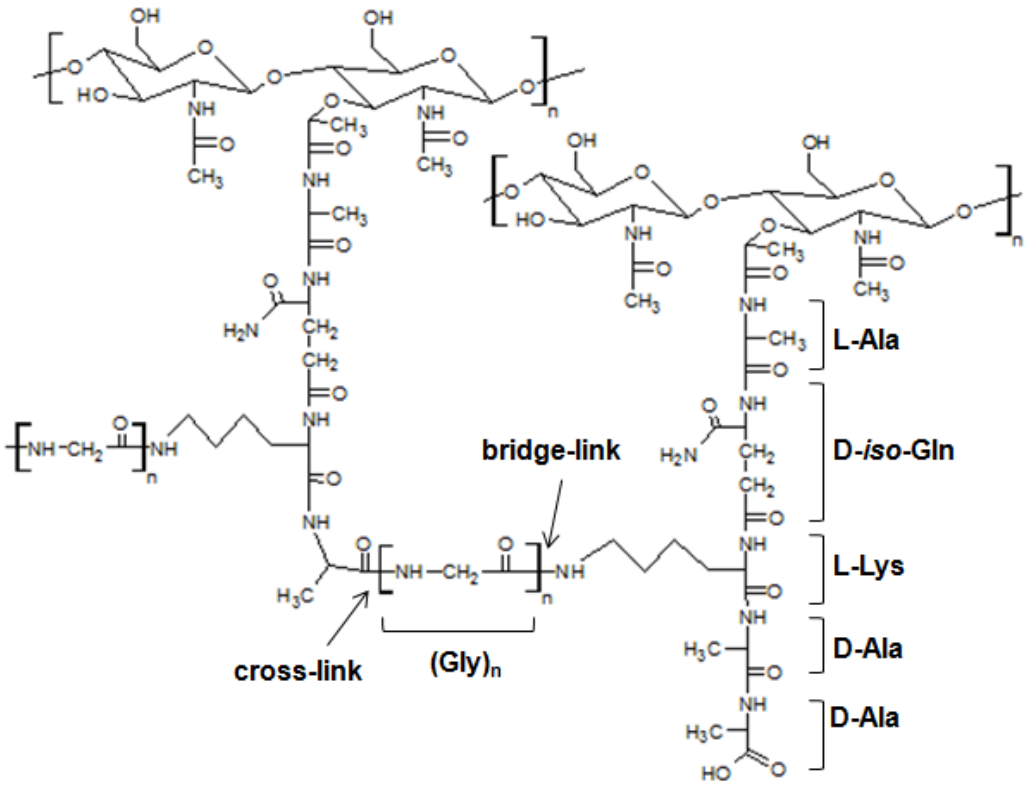


Figure 8.1: Chemical structure of the peptidoglycan stems in *S. aureus* and the *fem*-deletion mutants. Cross-linking between glycans occur through glycyl side chains connecting the carbonyl carbon of *D*-Ala of the 4th position of a stem to the □-nitrogen of *L*-Lys of the 3rd position of another stem. The bridge content in the side chain varies from pentaglycyl (MRSA), triglycyl (FemB) to monoglycyl (FemA).

8.3 Experimental

8.3.1 Sample Preparation

Start cultures of wild-type *S. aureus* (BB255) (*14*) and its *FemA* (UK17) (*7, 8*) and *FemB* (UT 34-2) (*10*) mutants were prepared by inoculating trypticase soy broth (TSB) in a test tube with a single colony from a nutrient agar plate. The overnight starter culture (1% final volume) was added to the sterile *S. aureus* standard medium (SASM), as described by Tong et al. (*15*). The cells were harvested at log phase at an optical density (OD) of 0.6 at 660 nm by centrifugation at 10000*g* for 10 min at 4°C in a Sorvall GS-3 rotor. Cell pellets were rinsed twice with ice-cold 40 mM triethanolamine hydrochloride (pH 7.0, adjusted with 1 M NaOH). The rinsed pellets were resuspended in 15 mL of the same buffer followed by rapid freezing and lyophilization.

The cell-wall isolates were obtained from a suspension of the rude cell-wall pellet in a minimum amount of sterile 10 mM triethanolamine hydrochloride buffer (pH 7.0, adjusted with 1 M NaOH) added dropwise with stirring to the boiling 4% sodium dodecyl sulfate (SDS). After 30 min boiling, the suspension was cooled for 2 h and was allowed to stand overnight at room temperature followed by centrifugation at 38000*g* for 1 h. Cell-walls were rinsed with buffer four times until no SDS could be observed. Then the pellet was resuspended in 0.01 M Tris buffer containing trypsin and α-chymotrypsin, incubated for 16 h at 37 °C, then sedimented and washed four times. The isolated cell walls were digested into muropeptides with lysozyme and mutanolysin following the protocol described by others (*16*) with slight modifications. All experiments were performed at neutral pH.

8.3.2 Time-dependent Isotopic Labeling of FemA Samples

SASM was prepared with natural abundance *L*-lysine replaced with uniformly labeled *L*-[^{13}C6,^{15}N2]lysine. The *FemA* cells were extracted from regular media when the absorbance at 660 nm was approximately 0.3 by centrifugation at 10,000 × g. Cells were then resuspended in the prewarmed SASM containing *L*-[^{13}C6,^{15}N2]lysine and incubated. Cells were harvested at various time points by centrifugation at 10,000g and boiled for 5 min in 40 mM triethanolamine hydrochloride (pH 7.0) to quench biological activity before being frozen and lyophilized. Isotopic enrichment is time independent after cells are harvested and boiled. We acquired 8 time points to investigate cell-wall development over approximately one cell cycle. The cell walls were then digested into muropeptides with lysozyme and mutanolysin (*16, 17*).

8.3.3 LC-MS/MS

Liquid chromatography/MS and MS/MS were performed by using a PicoView PV-500 (New Objective, Woburn, MA) nanospray stage attached to either an LTQ-FT mass spectrometer or an LTQ-Orbitrap mass spectrometer (ThermoFisher, San Jose, CA).

Muropeptide samples were loaded into an uncoated 75 µm i.d. fused-silica capillary column with a 15 µm picofrit tip (New Objectives, Woburn, MA), packed with C18 reverse-phase material (3 µm, 100 Å; Phenomenex, Torrance, CA) for 15 cm. The column was eluted at a flow rate of 250 nL/min for 10 min with 0.1% (vol/vol) formic acid in water and subsequently with a 60-min linear acetonitrile gradient (0%-40%) with 0.1% formic acid. The samples, as they emerged from the column, were sprayed into an

LTQ-FT mass spectrometer. Full mass spectra were recorded in the FT component of the instrument at 100,000 resolving power (at m/z = 400).

Accurate mass product-ion spectra of muropeptides were acquired by introducing the samples by nanospray as they eluted from the LC to an LTQ-Orbitrap mass spectrometer. To obtain major components' product-ion spectra, cycles consisting of one full FT-scan mass spectrum and five ensuing data-dependent MS/MS scans acquired by the Orbitrap (with a normalized collision energy setting of 35%) were repeated continuously throughout the elution with the following dynamic exclusion settings: repeat count, 3; repeat duration, 15 s; exclusion duration, 30 s. For MS/MS of minor components, a parent mass list was created and incorporated in the method.

8.4 Results and Discussion

8.4.1 Glycine Content in Muropeptides

Glycine content in peptidoglycan of wild-type *S. aureus* and the mutants reflects the average bridge length, and has been characterized using different approaches. The values are 1.2 for the FemA mutant, 2.6-2.9 for the FemB mutant, determined by amino acid analysis (*9, 10*). Recently Sharif et al. (*13*) measured the GEUs (glycin-equivalent unit) for whole cells of FemA and FemB to be 1.0±0.1 and 2.5±0.1 with the wild-type being normalized to 5.0. Due to the nature of these methods, the glycine content values are averaged among all peptidoglycan.

The LC-MS analysis, however, measures the relative amount of each muropeptide species. The overall glycine content is calculated by the following equation:

$$\frac{\sum n \cdot (muropeptide\ containing\ n\ glycines)}{\sum muropeptides}$$

The distributions of glycines in wild-type, FemA and FemB are shown in Fig. 8.2. The calculated average glycine content is 4.3 for wild-type, 0.9 for FemA mutant and 2.2 for FemB mutant. To compare these values to NMR measurement, wild-type glycine content is normalized to 5.0, thus glycine content in FemA and FemB are calculated to be 1.1 and 2.6, respectively. While it was suggested that there are 25% monoglycyl amd 75% triglycyl bridges in FemB peptidoglycan assuming these are the only two species present (*13*), here in the LC-MS measurement we are observing the distribution to be 12%, 20%, 6%, 60% and 2% for non-glycyl, monoglycyl, diglycyl, triglycyl and tetraglycyl bridges.

8.4.2 Bridge-links and Cross-links

MS results show that the degree of bridge-linking remains stable among the three strains, all being above 90%. The degree of cross-linking, however, decreases from wild-type (~70-80%) to FemB (~60-70%) and FemA (~50-60%), determined by quantitation method described in Chapter 5. We speculate that the estimation based on LC/MS analysis is less accurate than that based on NMR, owning to the highly heterogeneous muropeptides with varying ionization efficiency. MS is likely discriminating over the species with high degree of cross-linking, which disperse over a wide range of *m/z* and are ionized to a less extent. Figure 8.3 (a) exemplifies the complexity of peptidoglycan in wild-type. The heterogeneity is resulting from combinations of different muropeptide stems. Any given peak cluster corresponds to a particular oligomer, isotopically resolved by the mass spectrometer (Figure 8.3 (b)). While a number of muropeptides larger than

tetramer were observed for the wild-type strain, few of these were detected in the FemA peptidoglycan. The distribution of muropeptide species is dominated by dimers in FemA peptidoglycan, dimer and trimer in FemB peptidoglycan and trimer and above in the wild-type peptidoglycan.

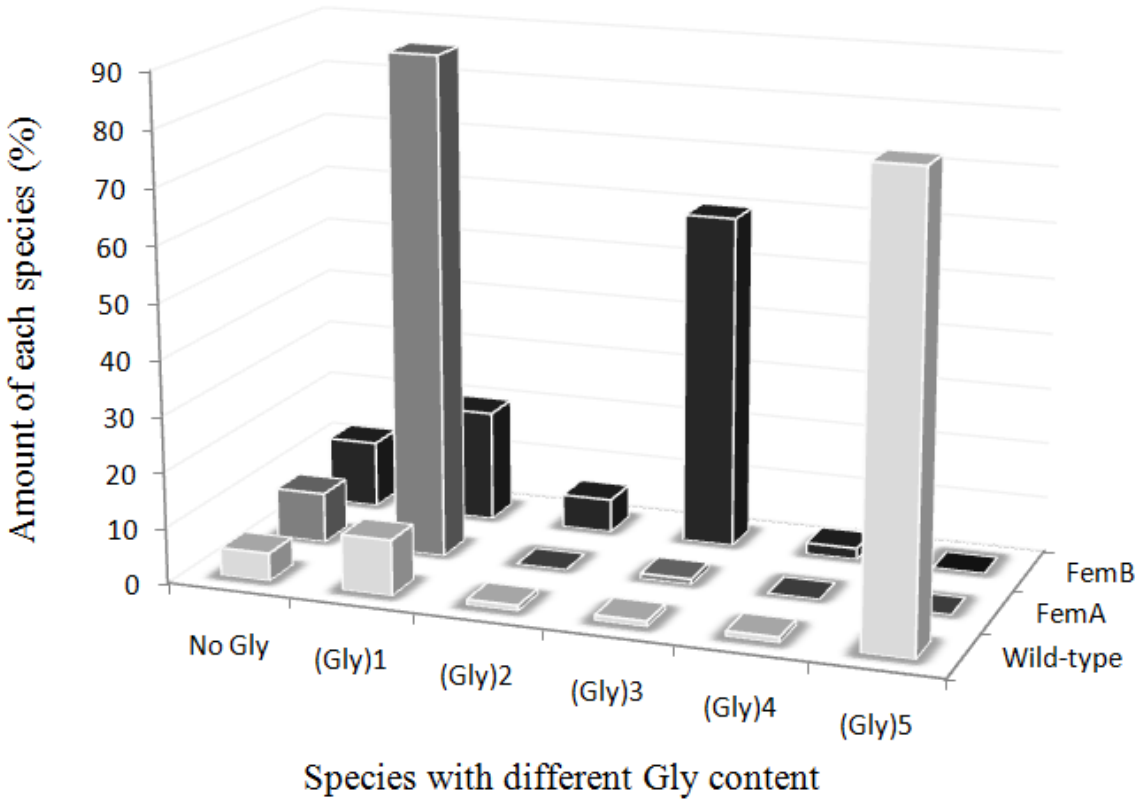


Figure 8.2: The percentage of each species with different Gly content in wild-type, FemA and FemB peptidoglycan monomers.

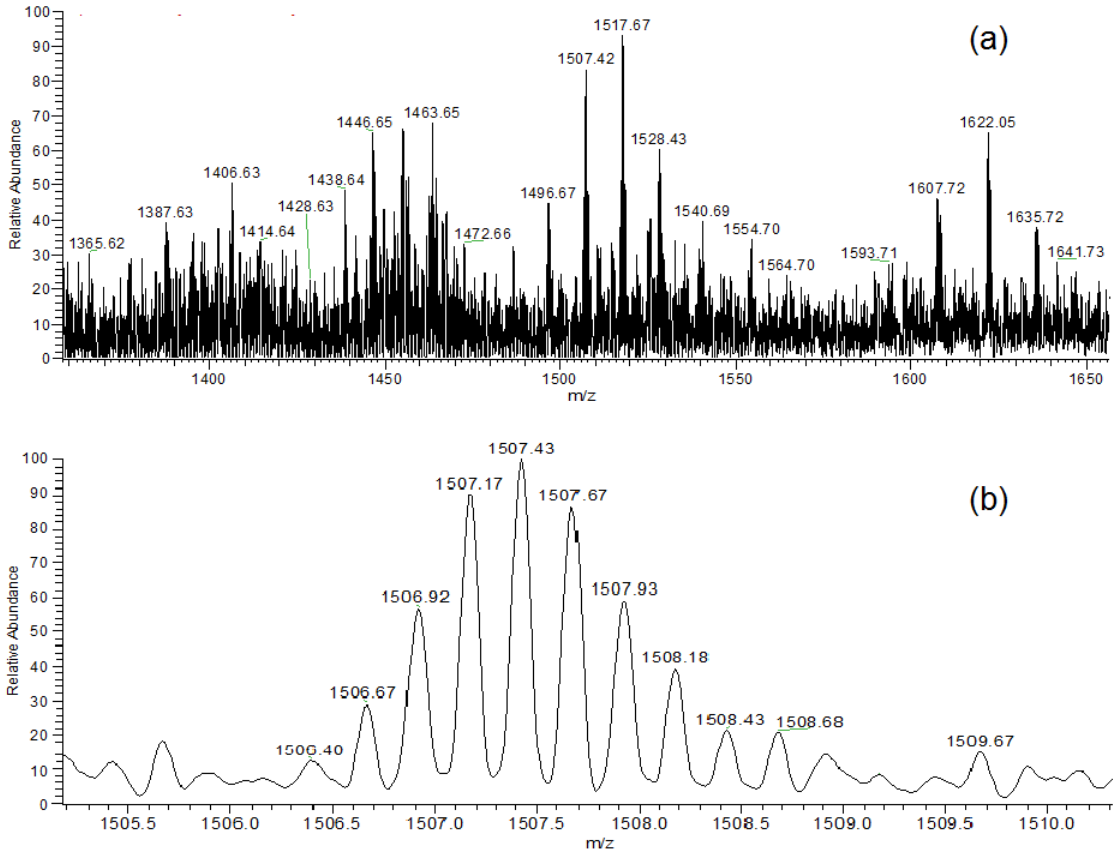


Figure 8.3: (a) A selected m/z range in the integrated mass spectrum over the retention time of wide-type peptidoglycan fragments larger than trimer. The peak clusters are corresponding to multiply-charged muropeptides with high degree of cross-linking. (b) A zoomed-in high-resolution spectrum of a peptidoglycan pentamer carrying four charges.

8.4.3 Gly-Gly-Gly in FemA

While *FemA*-deletion is believed to produce peptide bridge no longer than one glycine, Schaefer et al. suggests the existence of bridges containing three glycines in FemA mutant, supported by solid-state NMR data (unpublished). Here we detect with mass spectrometry a species with the accurate mass measurement corresponding to the exact mass of the acetylated muropeptide with three glycines in the bridge, as shown in Fig.8.4 (a). The structure is confirmed by MS/MS (Fig.8.4 (b)).

$(Gly)_3$ species, however, may not be homogenously distributed throughout the peptidoglycan. This is revealed by the different quantitation result obtained by whole-cell NMR analysis and bottom-up mass spectrometric analysis. We speculate that these species primarily exist in the region that is either unfavorable for digestion or lost during sample preparation before mass spectrometric analysis, thus yields less quantity than is determined by NMR where minimal loss occurs. The peptidoglycan heterogeneity in the whole cell of FemA is likely due to the incomplete knockout of *femA* by point mutagenesis (*8*).

8.4.4 Isotopic Labeling of FemA

We incorporated the isotopic labeling into the FemA growth cycle to investigate the dynamics of peptidoglycan processing and track the fate of precursors. The percentage of the species incorporating one or more *L*-[$^{13}C_6$,$^{15}N_2$]lysine isotopes was time-dependent. The peptidoglycan precursor, also known as Park's nucleotide (UDP-*N*-acetylmuramyl-l-alanyl-γ-d-glutamyl-meso-diaminopimelyl-*D*-alanyl-*D*-alanine), promptly incorporated *L*-[$^{13}C_6$,$^{15}N_2$]lysine and reached 94% enrichment after 10 min

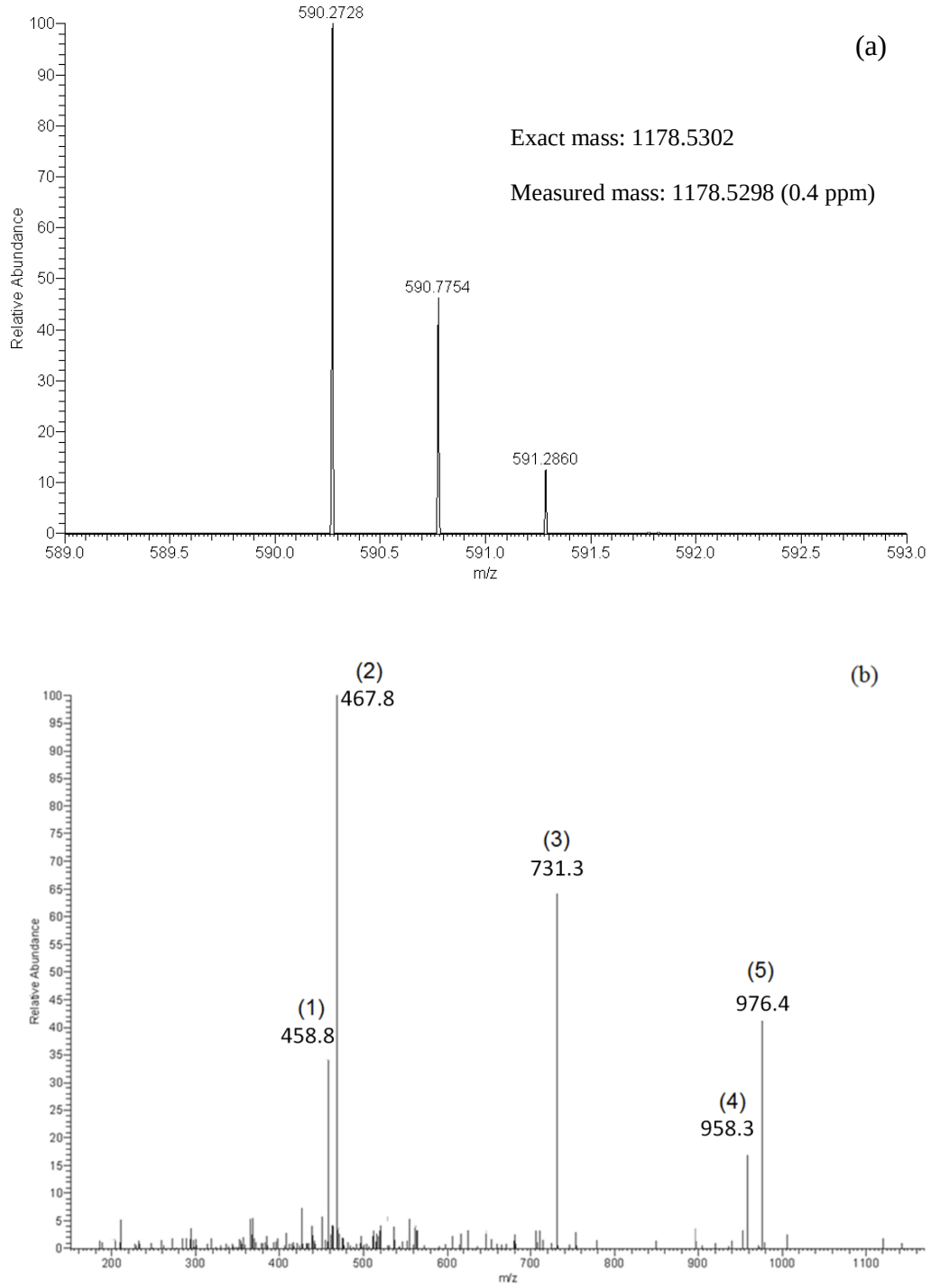
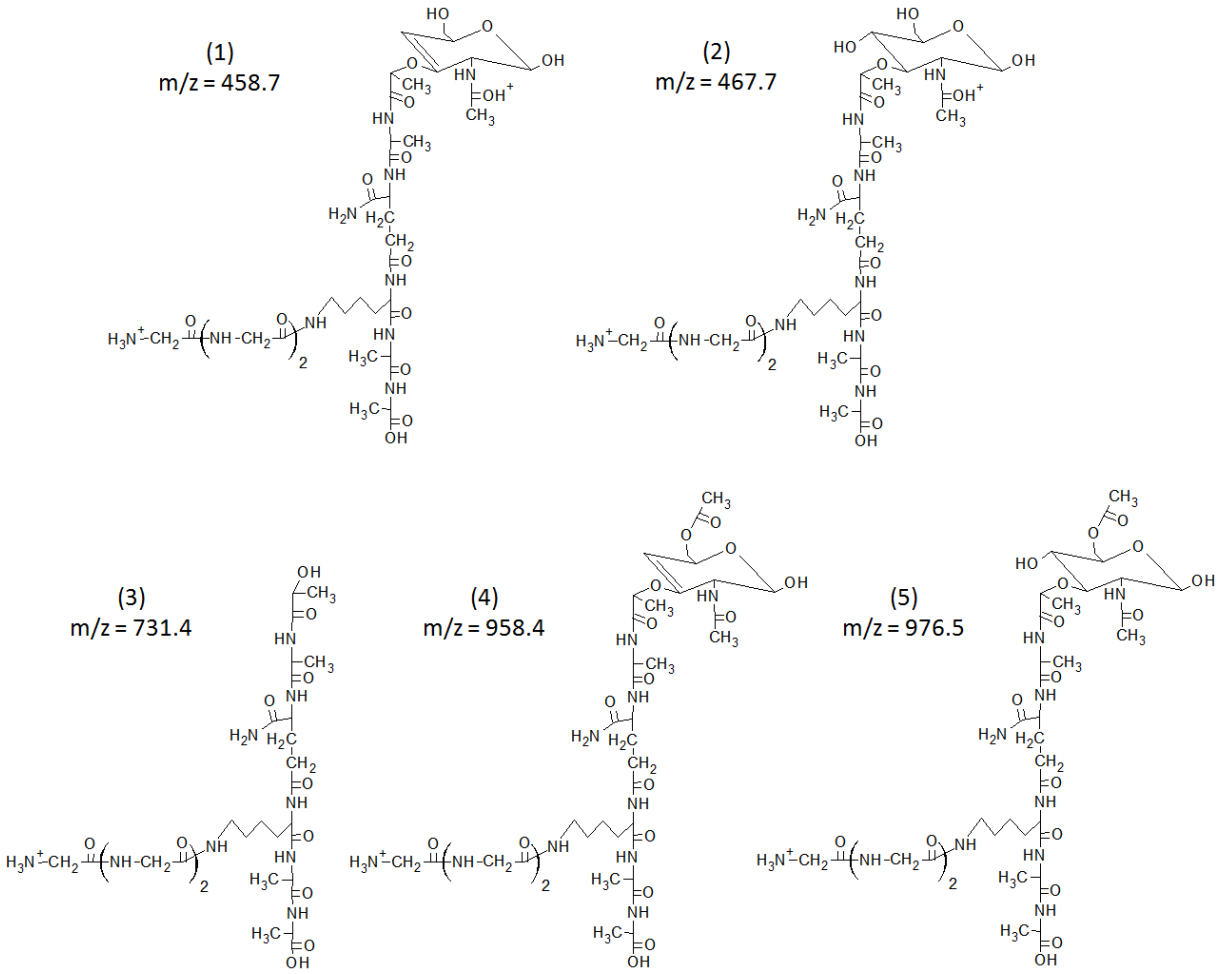


Figure 8.4: (a) Mass spectrum of the acetylated muropeptide from FemA containing three glycines. The accurate mass measurement is 0.4 ppm from the theoretical mass. (b) Product ion spectrum of the parent ion. The proposed structures of the fragments are listed in Scheme 8.1.



Scheme 8.1: Proposed structures of the fragment ions in the MS/MS spectrum in Figure 8.4(b). (1) and (2) are doubly charged.

incubation in labeled media (Figure 8.5, blue curve). Monomers with one glycine and two alanines, the most abundant species among all monomers, incorporated *L*-[$^{13}C_6$,$^{15}N_2$]lysine slightly more than the most abundant dimers (Figure 8.5, red and green curves) along the growth curve, indicating that dimers are produced later than monomers.

A comparison among monomers with no alanine, one alanine and two alanines reveals the post-translational modification sequence. It is shown in Figure 8.6 that monomers with two alanines are the first produced species, from which the other two species are modified. The species with one alanine is constantly less enriched than those with no alanine, indicating that the cleavage between two alanines by *D*,*D*-carboxypeptidase and between lysine and the adjacent alanine by *L*,*D*-carboxypeptidase are two independent processes, with the latter one being more progressive.

The incorporation of *L*-[$^{13}C_6$,$^{15}N_2$]lysine in dimers reveals some information regarding peptidoglycan growth model. Ten minutes after the pulse, there are ions representing dimer structures in which only one *L*-[$^{13}C_6$,$^{15}N_2$]lysine is incorporated (Figure 8.7, top). The number of doubly labeled dimers increases at intermediate times, but the percentage of structures with only one *L*-[$^{13}C_6$,$^{15}N_2$]lysine remains relatively constant compared with earlier time points (Figure 8.7, middle). After 90 minutes, more than half of the dimers (~60%) contain two *L*-[$^{13}C_6$,$^{15}N_2$]lysine residues, but structures with only one *L*-[$^{13}C_6$,$^{15}N_2$]lysine are still present at the same levels (~6%) as detected previously (Figure 8.7, bottom). The data suggest that the organization of FemA peptidoglycan agrees with the parallel model of glycan chain arrangement, for the same reason described in Chapter 7 (7.4.6). The structure with one *L*-[$^{13}C_6$,$^{15}N_2$]lysine,

however is less present in FemA (~6%) than in *E. faecium* (~10%), indicating that there more peptidoglycan layers in FemA than in *E. faecium*.

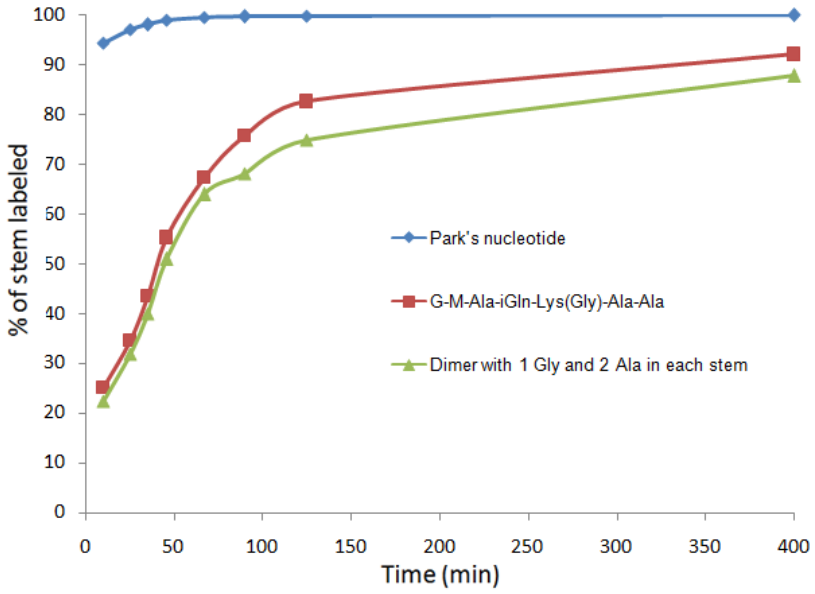


Figure 8.5: Plot of the percentage of isotopically enriched ions of Park's nucleotide (diamond), monomer with one glycine and two alanines (square) and dimers with one glycine and two alanines on each stem (triangle) with respect to time after pulse.

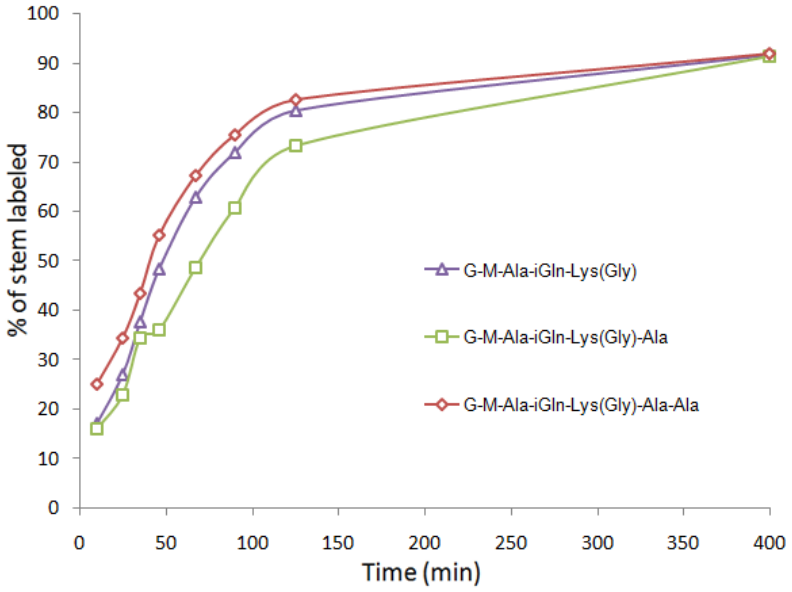


Figure 8.6: Plot of the percentage of isotopically enriched ions of monomers with no alanine (triangle), one alanine (square) and two alanines (diamond) with respect to time after pulse.

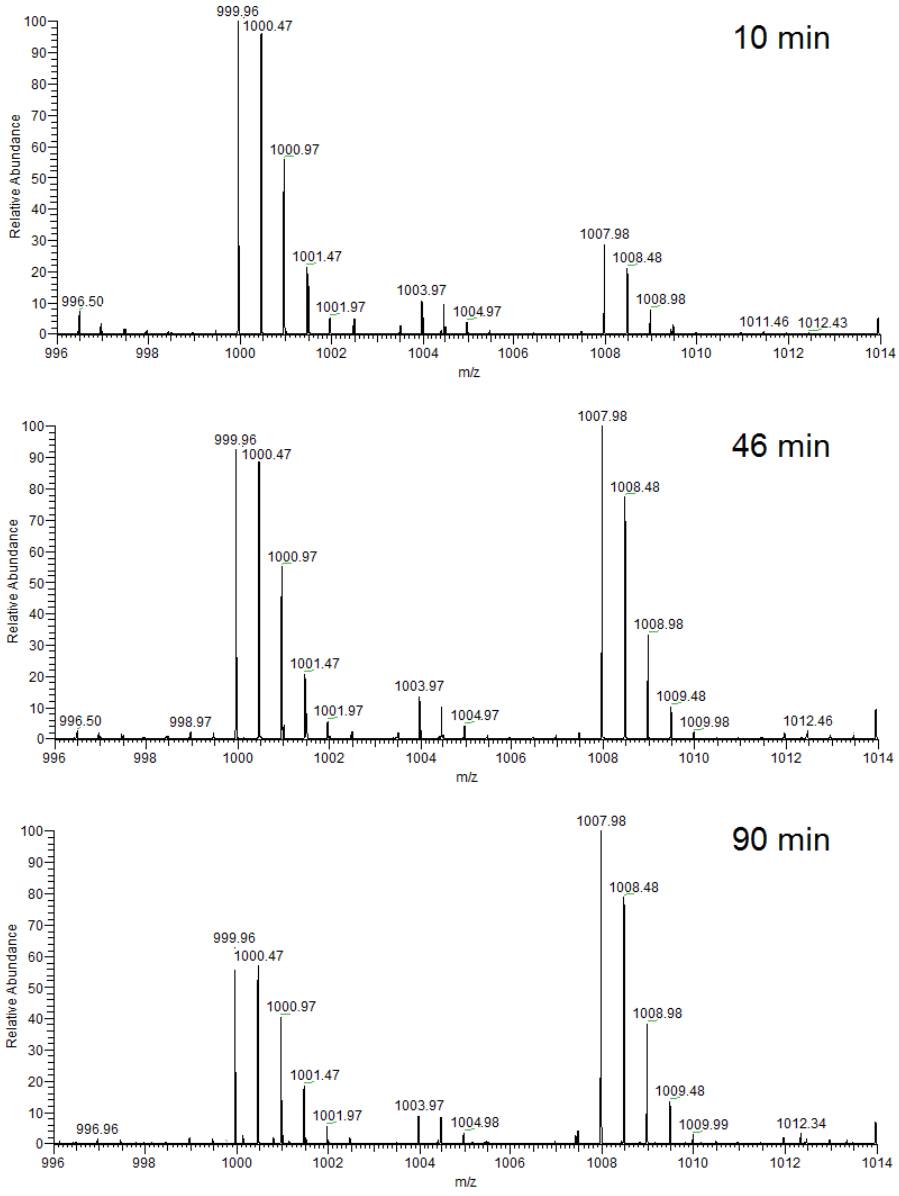


Figure 8.7: Mass spectra for the ion with an *m/z* of 999.96 corresponding to a doubly charged dimer with one glycine in each stem. (Top) After a 10 min pulse, (middle) after a 46 min pulse, and (bottom) after a 90 min pulse. The number of singly labeled dimers incorporating one *L*-[$^{13}C_6$,$^{15}N_2$]lysine remains relatively constant at all three time points.

8.4.5 *O*-Acetylation

The peptidoglycan of staphylococci has known to contain *O*-acetyl groups on about 60% of the *N*-acetylmuramic acid residues (*19*). The pulse-chase labeling previously conducted on *Staph. aureus* H suggests that *O*-acetylation is very closely linked with the addition of peptidoglycan units to the growing polymer *(20)*.

The *O*-acetylation extent is measured to be 54%, 64% and 63% for wild-type *S. aureus*, FemA and FemB, respectively. The percentage calculated for MRSA is likely to be underestimated owing to the bias of mass spectrometric detection on highly cross-linked species in MRSA, where *O*-acetylation occurs more often. The isotopic labeling experiment performed on FemA shows that the *O*-acetylated and non-acetylated species incorporates *L*-[$^{13}C_6$,$^{15}N_2$]lysine at a comparable rate (Figure 8.8), indicating that *O*-acetylation starts to occur in the early stage of maturation.

8.4.6 Peptidoglycan Architecture of Fem Mutants

The MS result is consistent with the observation by Schaefer and colleagues that significant reduction of peptidoglycan bridge length from MRSA to FemB and FemA mutants does not affect the ability of the Fem mutants to produce mature peptidoglycan structures (*13*). It also supports the hypothesis that a fundamental difference in the cell-wall tertiary structure may be present among these strains. The transition from tertiary structure in FemB peptidoglycan to that in FemA peptidoglycan may correlate with the difference in peptidoglycan structure of bacteria with long bridges and these with short bridges. The difference in tertiary structure may play a role in cell-wall biosynthesis and metabolism.

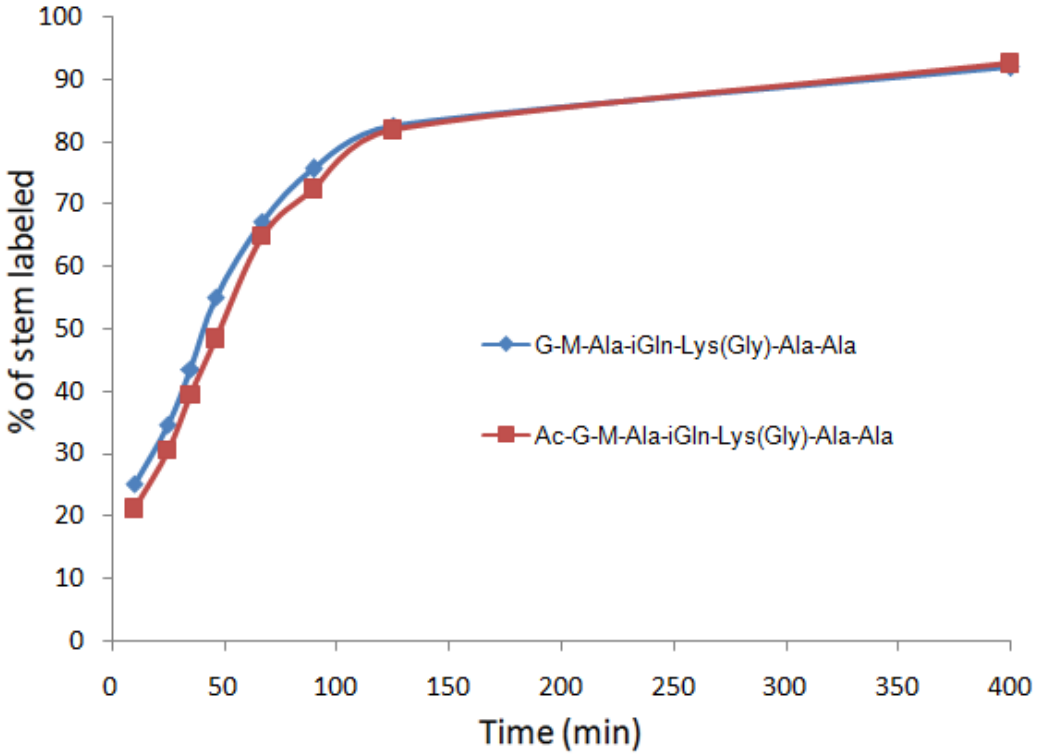


Figure 8.8: Plot of the percentage of isotopically enriched ions of monomers with two alanines and its acetylated species with respect to time after pulse.

8.5 Conclusion

In summary, mass spectrometry-based characterization of peptidoglycan provides detailed structural information of wild-type *S. aureus* and its *fem* mutants. Coupled with time-dependent isotopic labeling, it reveals the dynamics of peptidoglycan growth and post-insertional modification. As a complementary method to solid-state NMR, the MS strategy provides information regarding the cell-wall tertiary structure, which is an important subject in microbiology. It may also prove valuable for understanding the modes of action of glycopeptide antibiotics.

8.6 Acknowledgements

We thank Prof. Jacob Schaefer, Dr. Shasaad Sharif and Dr. Sung Joon Kim for the samples and the opportunity to collaborate on this project. This research was supported by the NCRR of NIH (Grant No. P41RR000954).

8.7 References

1. Hiramatsu, K., Cui, L., Kuroda, M., and Ito, T. (2001) The emergence and evolution of methicillin-resistant Staphylococcus aureus, *Trends in microbiology 9*, 486-493.
2. Hiramatsu, K., Hanaki, H., Ino, T., Yabuta, K., Oguri, T., and Tenover, F. C. (1997) Methicillin-resistant Staphylococcus aureus clinical strain with reduced vancomycin susceptibility, *The Journal of antimicrobial chemotherapy 40*, 135-136.

3. Hartman, B. J., and Tomasz, A. (1984) Low-affinity penicillin-binding protein associated with beta-lactam resistance in Staphylococcus aureus, *Journal of bacteriology 158*, 513-516.

4. Berger-Bachi, B., and Tschierske, M. (1998) Role of fem factors in methicillin resistance, *Drug Resist Updat 1*, 325-335.

5. Labischinski, H., Ehlert, K., and Berger-Bachi, B. (1998) The targeting of factors necessary for expression of methicillin resistance in staphylococci, *The Journal of antimicrobial chemotherapy 41*, 581-584.

6. Tschierske, M., Mori, C., Rohrer, S., Ehlert, K., Shaw, K. J., and Berger-Bachi, B. (1999) Identification of three additional femAB-like open reading frames in Staphylococcus aureus, *FEMS microbiology letters 171*, 97-102.

7. Ehlert, K., Schroder, W., and Labischinski, H. (1997) Specificities of FemA and FemB for different glycine residues: FemB cannot substitute for FemA in staphylococcal peptidoglycan pentaglycine side chain formation, *Journal of bacteriology 179*, 7573-7576.

8. Kopp, U., Roos, M., Wecke, J., and Labischinski, H. (1996) Staphylococcal peptidoglycan interpeptide bridge biosynthesis: a novel antistaphylococcal target?, *Microbial drug resistance (Larchmont, N.Y 2*, 29-41.

9. Maidhof, H., Reinicke, B., Blumel, P., Berger-Bachi, B., and Labischinski, H. (1991) femA, which encodes a factor essential for expression of methicillin resistance, affects glycine content of peptidoglycan in methicillin-resistant and methicillin-susceptible Staphylococcus aureus strains, *Journal of bacteriology 173*, 3507-3513.

10. Henze, U., Sidow, T., Wecke, J., Labischinski, H., and Berger-Bachi, B. (1993) Influence of femB on methicillin resistance and peptidoglycan metabolism in Staphylococcus aureus, *Journal of bacteriology 175*, 1612-1620.

11. Hubscher, J., Jansen, A., Kotte, O., Schafer, J., Majcherczyk, P. A., Harris, L. G., Bierbaum, G., Heinemann, M., and Berger-Bachi, B. (2007) Living with an imperfect cell wall: compensation of femAB inactivation in Staphylococcus aureus, *BMC genomics 8*, 307.

12. de Jonge, B. L., Sidow, T., Chang, Y. S., Labischinski, H., Berger-Bachi, B., Gage, D. A., and Tomasz, A. (1993) Altered muropeptide composition in Staphylococcus aureus strains with an inactivated femA locus, *Journal of bacteriology 175*, 2779-2782.

13. Sharif, S., Kim, S. J., Labischinski, H., and Schaefer, J. (2009) Characterization of peptidoglycan in fem-deletion mutants of methicillin-resistant Staphylococcus aureus by solid-state NMR, *Biochemistry 48*, 3100-3108.

14. Berger-Bachi, B. (1983) Insertional inactivation of staphylococcal methicillin resistance by Tn551, *Journal of bacteriology 154*, 479-487.

15. Tong, G., Pan, Y., Dong, H., Pryor, R., Wilson, G. E., and Schaefer, J. (1997) Structure and dynamics of pentaglycyl bridges in the cell walls of Staphylococcus aureus by 13C-15N REDOR NMR, *Biochemistry 36*, 9859-9866.

16. Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., van Heijenoort, J., and Gutmann, L. (1996) Peptidoglycan structure of Enterococcus faecium expressing vancomycin resistance of the VanB type, *The Biochemical journal 313 (Pt 3)*, 711-715.

17. de Jonge, B. L., Gage, D., and Handwerger, S. (1996) Peptidoglycan composition of vancomycin-resistant Enterococcus faecium, *Microbial drug resistance (Larchmont, N.Y 2,* 225-229.

18. Libisch, B., Lepsanovic, Z., Top, J., Muzslay, M., Konkoly-Thege, M., Gacs, M., Balogh, B., Fuzi, M., and Willems, R. J. (2008) Molecular characterization of vancomycin-resistant Enterococcus spp. clinical isolates from Hungary and Serbia, *Scandinavian journal of infectious diseases 40,* 778-784.

19. Mandelstam, M. H., and Strominger, J. L. (1961) On the structure of the cell wall of *Staphylococcus aureus*, *Biochemical and Biophysical Research Communications 5,* 466-471.

20. Snowden, M. A., Perkins, H. R., Wyke, A. W., Hayes, M. V., and Ward, J. B. (1989) Cross-linking and O-acetylation of newly synthesized peptidoglycan in Staphylococcus aureus H, *Journal of general microbiology 135,* 3015-3022.