Masters Theses   Graduate School

8-2006

The Integration of Oxidative Surface Mapping and Molecular Dynamics Simulation Techniques as a Strategy for Studying Protein Conformational Change

Kanan R. Vyas
University of Tennessee - Knoxville

Recommended Citation

Vyas, Kanan R., "The Integration of Oxidative Surface Mapping and Molecular Dynamics Simulation Techniques as a Strategy for Studying Protein Conformational Change. " Master's Thesis, University of Tennessee, 2006.
https://trace.tennessee.edu/utk_gradthes/1830

To the Graduate Council:

I am submitting herewith a thesis written by Kanan R. Vyas entitled "The Integration of Oxidative Surface Mapping and Molecular Dynamics Simulation Techniques as a Strategy for Studying Protein Conformational Change." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Life Sciences.

Robert Hettich, Major Professor

We have read this thesis and recommend its acceptance:

Hong Guo, Arnold Saxton

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

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

To the Graduate Council:

I am submitting herewith a thesis written by Kanan R. Vyas entitled "The Integration of Oxidative Surface Mapping and Molecular Dynamics Simulation Techniques as a Strategy for Studying Protein Conformational Change." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science with a major in Life Science.

_____Robert Hettich_____
Major Professor

We have read this thesis
and recommend its acceptance:

_____Hong Guo_____

_____Arnold Saxton_____

Accepted for the Council:

_____Anne Mayhew_____
Vice Chancellor and
Dean of Graduate Studies

(Original signatures are on file with official student records

The Integration of Oxidative Surface Mapping and Molecular Dynamics Simulation Techniques as a Strategy for Studying Protein Conformational Change

A Thesis presented for the Master of Science Degree
The University of Tennessee, Knoxville

Kanan R. Vyas
August 2006

Dedication

This thesis is dedicated to my fiancé, Amit N. Shah, for his moral support and encouragement, for his ability to make me smile on days that I didn't think I could and most of all, for his love.

Acknowledgements

I would like to thank those who have helped me complete my Master of Science degree in Life Science with a concentration in Genome Science and Technology.

I would like to especially thank my two advisors, Dr. Robert Hettich and Dr. Hong Guo, for their guidance in their respective specialties. I would also like to thank my committee (Dr. Robert Hettich, Dr. Hong Guo and Dr. Arnold Saxton) for their support, advice and encouragement in preparing me for a proper defense. I would finally like to thank Haobo Guo for his help with CHARMM, Christine Shook for her assistance with optimizing and conducting surface mapping experiments and Carlee McClintock for her assistance with performing tandem mass spectrometry.

Abstract

The range and number of new and unknown proteins is increasing at a staggering rate due to the recent genome sequencing projects. The next step in understanding how biological systems, even including the human body, work is by understanding the function of all the various proteins. Solving the structure of a protein is an important first step in elucidating its function; however, the study of its dynamic movements can specifically implicate regions involved in its function and even demonstrate the mechanism by which function is performed.

Molecular dynamics simulations are a powerful computational approach for visualizing the dynamic movement of proteins. Computational tools are predominantly theory based predictions. Therefore, they require validation by experimental results. Oxidative surface mapping is an experimental labeling method which can be used to identify "buried" vs. "solvent-accessible" regions in a folded protein. Movement in specific regions of a protein can be mapped and monitored using this method.

β-lactoglobulin is a well studied protein that undergoes a pH induced conformational change. It was chosen as the target protein for this study because it has been the focus of numerous studies in the past and much information is known about it. Even so, many aspects of this protein's structure still remain a mystery.

This thesis work is an attempt to integrate computational and experimental techniques as a strategy for studying the protein conformational change of a well studied protein system. The degree of overlap displayed by the integration of these two

techniques is limited, however it provides a foundation from which improvements can be

implemented for future attempts of studying protein systems using this approach.

Table of Contents

List of Tables

List of Figures

Chapter 1: The Advantage of Interfacing Experimental and Computational Approaches for Characterizing Protein Conformational Change

I. The Mismatch Between Protein Discovery and Protein Structural Characterization

The recent genome projects have generated an abundance of biological sequence data. Sequencing of the human genome is now complete; however, the full potential of the accomplishment continues to be elucidated. A large proportion of the genome is "junk DNA", which is thought to provide no direct function for the cell, (the fundamental unit of every living system) and less than 2% of the human genome encodes for proteins. Deriving meaningful knowledge from this data will take the work of countless scientists as protein sequences provide information that can be extracted by a multitude of methods to answer a variety of questions about the human body.

Determining the function of proteins is important in understanding how the human body works. A small fraction of proteins can be assigned a function by sequence similarity (Andrade et al., 1999). Structural similarity, on the other hand, can assign functions for significantly more proteins (Koppensteiner et al., 2000). Determining the structure of a protein from its sequence is therefore an important step in the path to elucidating its function (Domingues et al., 2000).

A variety of methods are currently employed for protein structure determination. X-ray crystallography and NMR are the most widely used experimental techniques for solving protein structure; however, computational approaches have also demonstrated promise in recent years. The Brookhaven Protein Data Bank (PDB) houses

1

experimentally determined protein 3D structures. To date, the PDB contains slightly more than 34,000 protein 3D structures; the majority of which were solved by X-ray crystallography. Though the number of structures submitted to PDB increases every year, the rate is not fast enough to keep up with the efficiency at which sequences are compiled. The experimental methods employed have limitations that affect the efficiency at which 3D structures can be solved. Computational approaches provide increased speed but are theory based and still require confirmation by experimental based methods.

Though structure prediction is an essential step in elucidating the function of a protein, it only provides information as a snapshot in time. Monitoring the structural movement of a protein as a function of time can specifically implicate regions important in its role (Karplus and McCammon, 2002). Furthermore, characterization of these regions of movement provides the higher degree of resolution necessary to clarify mechanisms for the conformational shifts, which can ultimately provide information about the protein's biological function (Christodoulou et al., 2004).

Computational techniques, such as molecular dynamics (MD) simulations, are powerful tools for studying the movement of proteins. They provide the ability to visualize proteins in motion at a resolution difficult to probe using experimental methods (Karplus and Kuriyan, 2005). Though a MD simulation can provide a considerable amount of information about a protein system, it is predominantly a theory based prediction and requires validation by experimental methods (Karplus and McCammon, 2002; Karplus and Kuriyan, 2005). Therefore, the attempt to integrate computational and experimental approaches provides a theoretical and experimental based solution to the study of protein conformational change.

The purpose of this thesis work is an attempt to integrate molecular dynamics simulations and an experimental technique with the premise that it will give new insight about a well studied model system.

II. Current Status of Experimental Approaches for Characterizing Protein Structure

Experimental techniques used to determine the structure of proteins have also been successfully utilized for visualizing their conformational change. The brief discussion that follows summarizes a few commonly used experimental techniques for studying conformational change, shedding light on the benefits and limitations of each.

X-ray crystallography has been employed to study conformational changes of proteins (Hakansson et al., 1997; Qin et al., 1998). Though researchers have had success with this method for detailed protein studies with a high degree of resolution, it has limitations and is time consuming. Culturing a crystal requires highly pure protein and can potentially take a long time depending on the proteins size, stability and structure. Many proteins are difficult to purify and concentrate, thus making them poor candidates for crystal growth. Additionally, crystallography provides the structure of a protein in the solid state as a snapshot in time. Therefore, multiple crystal structures harvested under different conditions are required to monitor any potential change in conformation. Growing crystals of the same protein under different conditions could be cumbersome as the protein under certain conditions may crystallize more readily than others. Because of this reason, a high level of patience is required when using this method for the study of protein conformational change. Finally, because crystals provide structures in the solid

state, specific reactions that require dynamic movement of the protein cannot be visualized (Hakansson et al., 1997). X-ray crystallography is an excellent first step in elucidating the structure and function of a protein; however, methods that allow visualization of movement in real time are better for extracting more detail.

NMR spectroscopy based methods have been successfully utilized as an experimental method to study the dynamic movement of proteins (Christodoulou, 2004; Sakurai and Goto, 2005). Its obvious benefit over x-ray crystallography is its ability to monitor protein conformational change in a time-dependent manner and in the solution-phase. However, not all proteins are good candidates for study by this method. Proteins must be abundantly available and highly pure for studies of their movement to be successful using this method. Additionally, proteins of study are restricted to those less than 40 kDa. Also, previous research has demonstrated the occurrence of severe signal broadening at neutral pH (Sakurai and Goto, 2005). This can potentially interfere with the accuracy of data analysis. Though NMR based methods can be easier to perform and provide information in the solution state their limitations generate a need for alternative methods for studying protein conformational change.

Various solution-phase structural labeling techniques have been used to study the structure of proteins. Chemical labeling methods such as hydrogen-deuterium exchange and oxidative surface mapping are powerful methods in understanding the structure and dynamics of proteins. Both methods are performed in the solution-phase and study the solvent-accessible surfaces of proteins by non-specifically labeling solvent exposed regions. Hydrogen-deuterium experiments can give information about a protein's backbone structure and hydrogen bonding properties (Sharp, 2003) but results are

difficult to extrapolate because the exchange rates are a sum of the hydrogen bonding, solvent-accessibility and back-exchange kinetics of the protein (Sharp et al., 2003). Surface mapping utilizing chemical modifications to probe solvent accessible surfaces is also a method that has been successful (Sharp et al., 2003). Many methods to modify solvent accessible amino acids are available, however non-specific methods have proved to be the most efficient. Preference is given to methods that can be performed with readily available materials, and are efficient and accurate in what one wants to study. Oxidative surface mapping of proteins via chemically generated hydroxyl radicals fits all three of these criteria (Sharp et al., 2003).

Given that most all of these experimental techniques have their limitations, coupling them with molecular dynamics simulations can together provide a stronger approach for studying protein conformational change, potentially allowing for a more thorough investigation than that provided by either method alone.

III. Selection of a Model Protein System for Examination of Conformational Changes

In choosing the model system of study, several important criteria were considered: the abundance and ease of preparation of the protein, the ability to mimic and accurately represent the protein in its native state, the presence of a conformational change of the protein system and previous successful studies that can be used to validate findings.

β-lactoglobulin is the most abundant globular protein found in the whey fraction of bovine milk (Dufour et al., 1994). Bovine β-lactoglobulin has as many as nine genetic

variants, of which variant A and B are the most common (Qin et al., 1998). The sequence of these two variants differ at only two amino acid residues: Asp64 in A is changed to Gly in B, and Val118 in A is changed to Ala in B (Qin, 1999); however, they experience a distinctly different degree of thermal stability and display independent self-association properties (Qin et al., 1999). The dairy industry produces large quantities of this protein for its use in the industrial processing of milk. In addition, it is used to form gels in food systems and can be engineered for other uses (Dufour et al., 1994). Its function, *in vivo*; however, still remains unknown. The two most common variants are readily available commercially (Sigma-Aldrich) for the purpose of experimental analysis and can be easily prepared for scientific studies as described in the next chapter.

Multiple studies of the bovine β-lactoglobulin protein have been conducted. These studies have employed various techniques including: x-ray crystallography (Qin et al., 1998), NMR (Sakurai and Goto, 2005), infrared spectroscopy (Dufour et al., 1994) and MD simulations (Eberini et al., 2004; Fogolari et al., 2005). The ability to study this protein by a variety of diverse experimental methods demonstrates the ease and accuracy by which its environment can be mimicked. This protein is an even more attractive candidate because of the fact that its function remains unknown even though analysis of its properties seems relatively uncomplicated.

The β-lactoglobulin protein structure is composed of eight antiparallel β-strands which forms a β-barrel, a ninth strand (strand I) that is implicated in dimerization, and one major and three minor α-helices (Uhrinova et al., 2000). It undergoes pH induced global and local transitions of its structure. Between pH 2 and pH 6 the protein undergoes

multiple aggregation and dissociation events (Taulier and Chalikian, 2001). Globally, it shifts from a monomer to a dimer to an octomer and back to a dimer just before pH 6 (Taulier and Chalikian, 2001).

As shown pictorially in Figure 1.1, the protein undergoes a pH induced local transition of its structure, termed the Tanford Transition (Tanford, 1959) roughly between pH 6 and pH 8. This transition involves the movement of the EF loop (the loop connecting β-strand E and β-strand F) of the protein (Qin et al., 1998). In acidic conditions, the EF loop closes off the entrance of the hydrophobic calyx (Qin et al., 1998). As the environment becomes more basic, the loop pulls away from the calyx opening, exposing the hydrophobic core of the protein (Qin et al., 1998). This pH induced movement of the EF loop suggests a possible function of the protein for transporting hydrophobic cargo. Another pH induced conformational change, around the α-helix region, has been visualized during MD simulations of the retinol binding protein (RBP, a protein structurally similar to β-lactoglobulin) (Gu and Brady, 1992). This suggests the presence of another region that undergoes a pH induced local transition. The Tanford Transition as well as the proposed movement near the α-helix region of the β-lactoglobulin protein is the focus of this study. The presence of a conformational change of the protein further confirms β-lactoglobulin is an appropriate protein candidate for this thesis study.

The β-lactoglobulin protein has been the focus of numerous studies since its first isolation in 1934 (Palmer, 1934). Though research on this protein system has been conducted for more than 70 years, its function still remains a mystery. Much has been

PDB ID 3BLG, pH 6.4   PDB ID 1BSY, pH 7.1   PDB ID 2BLG, pH 8.2
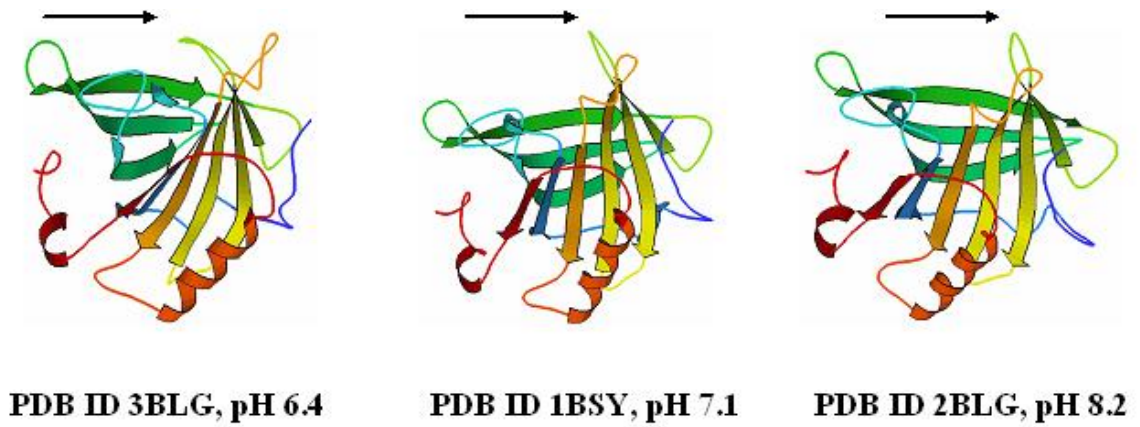
Figure 1.1: Structural Progression of the Tanford Transition.

Three crystal structures of bovine β-lactoglobulin A deposited into PDB from the same study (Qin et al., 1998). The structures at pH 6.4, 7.1 and 8.2 demonstrate the pH range spanning the Tanford Transition. The arrow points to the EF loop in each structure displaying how it pulls away from the opening of the hydrophobic calyx of the protein.

learned; however, about the properties of the protein. Studies of its structure and dynamics by a variety of methods have been performed and provide results that are consistently duplicated. For example; information from a NMR based dynamics study of the protein provided information about the rigidity of various regions of the β-lactoglobulin protein in solution state at a constant pH (Kuwata, 1999), the comparison of its structure with the structure of another protein in its superfamily suggested a plausible role for structurally similar domains (Papiz, 1986) and analysis of changes in conformation of its structure provided a deeper understanding of regions implicated in function of the protein system (Qin et al., 1998). These results provide a substantial pool of information about the protein to which results obtained from this study can be compared and validated.

IV. Selection of an Experimental Technique for Characterizing Protein Conformational Alterations

As discussed previously, computational approaches for the study of protein systems require confirmation by experimental results before they can be considered reliable. Molecular dynamics simulations of a protein system begin with an initial set of coordinates solved by either x-ray crystallography or NMR based structure determination methods. Since a partial integration between these experimental techniques and MD simulations is already attained simply by performing a MD simulations, and studies of the Tanford Transition of β-lactoglobulin utilizing x-ray crystallography (Qin, 1998) and NMR (Sakurai, 2005) based methods have been performed previously, integration of MD

simulation with chemical labeling methods was chosen as the strategy for this thesis study.

Analysis of the conformational change of the protein by chemical labeling has not been performed over a range of pH values. It has, however, been performed at a single pH value to support photochemical oxidation as a good method to probe the solvent accessible surface are of proteins (Sharp et al., 2004).

A variety of chemical labeling methods can be employed to study protein solvent accessibility. Due to the nature of the thesis study proposed, non-specific labeling techniques are preferred. Non-specific labeling provides better sequence coverage and hence better resolution of conformational shifts. Oxidative surface mapping via chemically generated hydroxyl radicals was chosen because of its accuracy, simplicity, efficiency and ability to be performed without the need for specialized equipment (Sharp et al., 2004). Additionally, this technique allows for non-specific labeling of the protein structure, thereby, eliminating the need to bombard the protein with a variety of different reagents.

V. Contributions That Can be Made to the Abundance of Knowledge Already Compiled for the Conformational Change of This Protein System.

This thesis will help elucidate the importance of combining computational and experimental approaches for the study of biological systems, but more specifically for the study of protein conformational change. Furthermore, it will serve as a foundation for future attempts at integrating two technologies and as a motivation for collaborations between researches from the two distinctly different disciplines.

The objective of this thesis, by integrating oxidative surface mapping and molecular dynamics simulation techniques, is to: determine the degree of integration possible between two distinct scientific approaches, learn about the limitations of both, and propose future refinement of integration as a strategy for studying protein conformational change.

Under the premise that integration of two techniques will provide more meaningful information than either one alone, the goal of this thesis work was to not only validate previous findings from studies of the β-lactoglobulin system but also contribute novel information to the abundant information that is already present.

This thesis study is divided into several chapters. Chapter 2 describes, in detail, the process by which oxidative surface mapping and molecular dynamics simulation techniques were applied to study the Tanford Transition of bovine β-lactoglobulin A. Chapters 3 and 4 present the results from both techniques and discuss their relevance to the conformational change of the protein. Results from these two chapters are compared and contrasted with crystal structures obtained at pH values of 6.2, 7.1 and 8.4; of which the pH 6.2 crystal was used for model building in MD simulation analysis. Chapter 5 demonstrates an attempt at integrating results from the two technologies. It includes a discussion of the benefits and drawbacks of the two methods and proposes improvements to the strategy for future attempts at studying protein conformational change.

Chapter 2: Materials and Methods for Monitoring Conformational Transitions in β-lactoglobulin A

This chapter describes our approach to integrate experimental and computational technologies to study conformational changes in proteins. The basic strategy is to probe the change in solvent accessible surface area of a well studied model system (β-lactoglobulin) as a function of pH using oxidative surface mapping, and then compare and contrast the findings with molecular dynamics simulations of the same system (Figure 2.1). Oxidative surface mapping was chosen because of its validity as a method to probe solvent accessible surface areas of proteins (Sharp et al., 2003). Molecular dynamics was paired with the oxidative surface mapping technique because of its ability to monitor change in a time dependent manner.

I. Oxidative Surface Mapping

β-lactoglobulin undergoes a conformational change called the Tanford Transition within a specific range of pH values (Tanford et al., 1959). Generation of hydroxyl radicals *in situ* with the native protein at a particular pH oxidizes amino acid residues in regions of the protein that are solvent exposed (Maleknia et al., 1999, 2001A-B, 2002). Mass spectrometry analysis of the protein after proteolytic digestion identifies those residues that are oxidized (Maleknia et al., 1999, 2001A-B, 2002). This technique will allow for the visualization of the conformational change in β-lactoglobulin by tagging those sites that become solvent accessible due to a change in conformation as a result of the proteins pH environment. The denatured protein serves as a positive control for all the possible oxidation sites within the protein and the
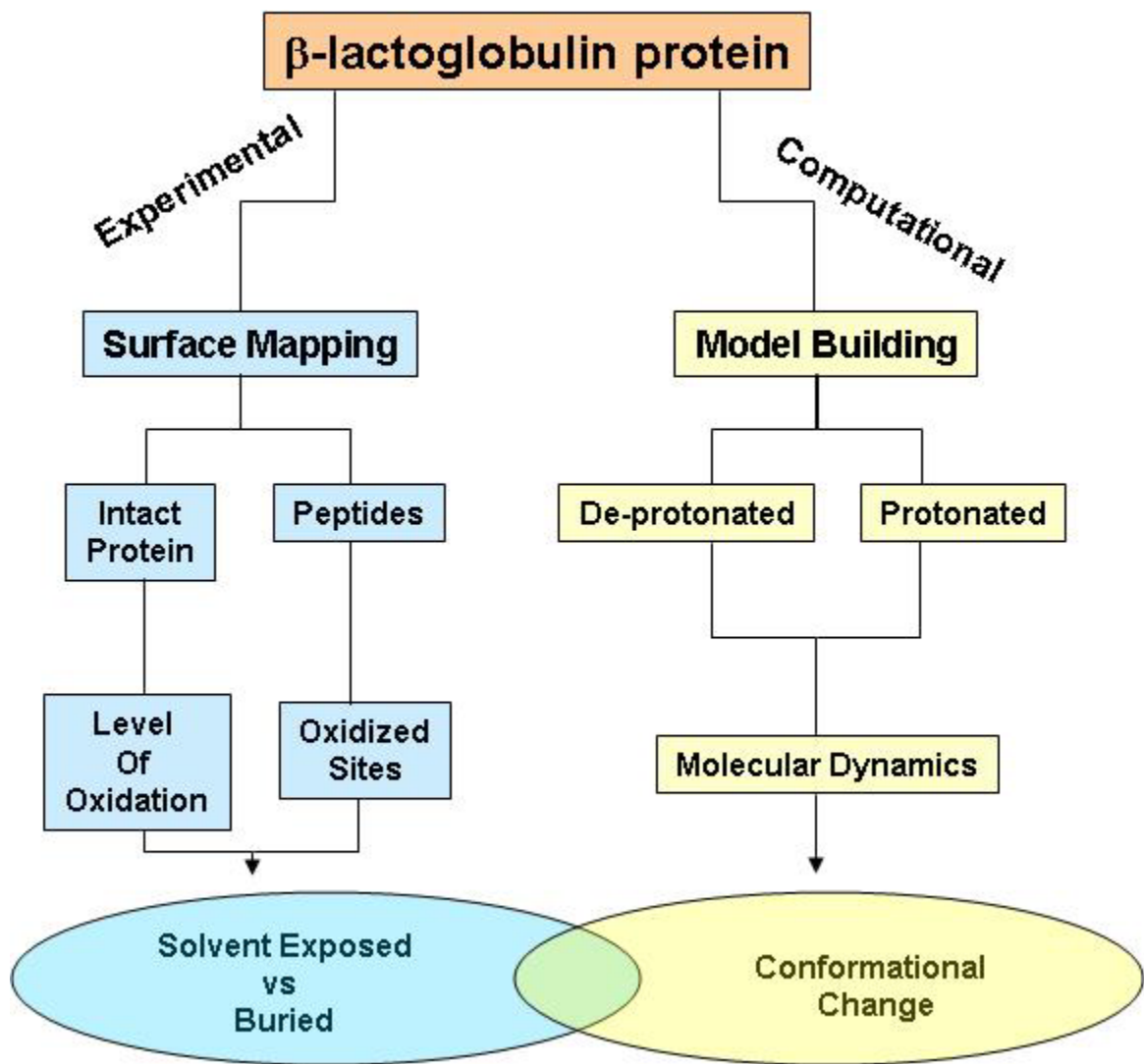
Figure 2.1: Scheme for the Integration of Experimental and Computational Approaches.

A diagram of the scheme utilized to study the conformational change of the β-lactoglobulin protein by coupling oxidative surface mapping and molecular dynamics simulation techniques.

lysozyme control experiments demonstrate that the pH dependent oxidation difference are due to the protein and not the hydroxyl radical.

A. Sample Preparation

A stock solution of β-lactoglobulin A (Sigma-Aldrich, St. Louis, MO) was prepared by reconstituting 2 mg of the lyophilized protein in 100% HPLC grade water (Burdick & Jackson, Muskegon, MI) to a concentration of 2 mg/mL. The stock was aliquoted into individual 2 mL Eppendorf tubes (VWR, West Chester, PA) such that each contained about 250 uL volume (0.5 mg concentration of protein). Another 750 uL of 100% HPLC grade water was added to each Eppendorf tube to bring the volume to 1 mL and concentration to 0.5 mg/mL.

The denatured sample was treated with 1 mL of a denaturing mixture of 6 M Guanidine HCl (Sigma-Aldrich) and 10 mM DTT (EMD Chemicals, Gibbstown, NJ) in 10 mL of 50 mM Tris (Sigma-Aldrich) 10 mM $CaCl_2$ (J.T. Baker, Phillipsburg, NJ) buffer and incubated for one hour in a 60°C water bath (Precision microprocessor controlled 280 series water bath, Thermo Electron Corporation, Waltham, MA).

The various pH samples were adjusted to their desired pH value by adding small amounts of 5 M HCl (Sigma-Aldrich) and/or 0.5 M NaOH (Sigma-Aldrich). The pH of each sample was verified by pH paper (EM Science, Gibbstown, NJ) and these samples were allowed to sit on the bench top until the incubation of the denatured sample so that oxidation of all samples could be done at one time. Sample preparation for the lysozyme (Sigma-Aldrich) experiments were done in the same manner as the pH adjusted β-lactoglobulin A samples.

B. Oxidation

An aliquot of 250 uL of 50% H_2O_2 (EMD Chemicals) was added to each Eppendorf tube to achieve an overall 10% H_2O_2 concentration (by volume). The oxidation reaction was initiated by exposing the tubes, with their caps off, to UV light in a Stratalinker (UVStratalinker 2400, Stratagene, La Jolla, CA) with a wavelength of 254 nm for 5 minutes.

Immediately after exposure, 2 uL of 100% Formic Acid (FA, EMD Chemicals) was added to each tube so that the protein would bind more efficiently to the stationary phase of a tC2Light SepPak (Waters, Milford, MA).

The oxidation reaction was quenched by extracting the sample with a tC2Light SepPak and allowing the H_2O_2 to be washed through while binding the oxidized protein. A tC2Light SepPak was used because the C2 stationary phase was sufficient for capturing intact proteins (*tC2*) and the concentration of protein in our sample was less than 1 mg (*Light*). Solutions were drawn up through a needle into a syringe, the needle was replaced with a tC2Light SepPak and the solution was passed through the SepPak by applying pressure to the syringe. To insure air was not ejected though the SepPak, solution was allowed to pass through until only a small layer of solution was visible in the syringe; except for the elution step where the solution was allowed to pass through to completion.

A new tC2Light SepPak was used for each sample and was conditioned by first passing 10 mL of a solution of 100% ACN/0.1% FA (ACN, Burdick & Jackson) followed by 10 mL of a solution of 100% HPLC grade water/0.1% FA. Next, the sample was passed through the SepPak twice. The SepPak was then washed by passing 10 mL of

100% HPLC grade water. Finally, elution was done into a clean 2 mL Eppendorf tube by 1.5 mL of 100% ACN. Care was taken to insure the first drop of the elution was captured as it contains the most protein.

Eluted samples were dried down in a SpeedVac (SpeedVacPlus SC210A, Savant, Santa Rosa, CA) using the low setting with the concentrator set to off until about 300 uL in volume. Progress was monitored every few minutes due to the high volatility of ACN. Based on previous experience, a 50% loss of sample was expected upon extraction through a SepPak.

C. Digestion

A 1 mL aliquot of the same denaturing mixture used to prepare the denatured sample was added to each of the Eppendorf tubes. The samples were incubated for 1 hour in a 60°C water bath. Upon incubation, each of the samples was transferred to a clean 15 mL Falcon tube (VWR) and diluted 6 fold by adding 6.5 mL of 50 mM Tris CaCl$_2$ buffer for a total volume of about 7.8 mL. Next, the pH of the samples was adjusted to pH 7 by adding minute amounts of 5 M HCl and/or .5 M NaOH. The pH of each sample was verified by pH paper. Dilution of the sample and adjustment of the pH were done to bring the sample conditions to less than 1M Guanidine and neutral pH, optimal conditions for trypsin digestion.

A 20 ug vial of lyophilized trypsin (Promega, Madison, WI) was reconstituted with 100 uL of the re-suspension buffer provided (50 mM Acetic Acid, Promega). 25 uL (5 ug) of this was added to each sample tube such that the ratio of trypsin to protein was 1:50. The sample tubes were tightly wrapped with parafilm to prevent evaporation and

placed on a rotator (Nutrating Mixer, VWR) in a 37°C incubator (Model 120, Barnstead/Lab-Line, Melrose Park, IL) for 18 hours. The following morning the same concentration of trypsin was added to each sample tube and the tubes were again placed on a rotator in a 37°C incubator; this time for 5 hours. Upon the second incubation the samples were centrifuged (Spinchron R, Beckman Coulter, Fullerton, CA) at 3000 rpm for 10 minutes. Immediately after centrifugation the supernatant was transferred into a clean 50 mL Falcon tube (VWR).

The supernatant was cleaned up by passing it through a tC18Light SepPak (Waters). A tC18Light SepPak was used because the C18 stationary phase was necessary for capturing peptides (*tC18*) and the concentration of peptides in the sample was less than 1 mg (*Light*). The same SepPak procedure was followed as in quenching of the oxidation reaction however the tC18Light SepPak was washed with a solution of 10 mL 100% HPLC grade water/0.1% FA after passing the sample through twice and elution was done with 2 mL of a solution of 100% ACN/0.1% FA. Formic acid was present in the wash and elution solutions because peptides need to be acidified more than intact proteins since they have less net charge.

D. ES-FTICR Analysis

All mass spectra were acquired by direct infusion with a 9.4-Tesla HiRes electrospray Fourier transform ion cyclotron resonance (ES-FTICR) mass spectrometer (IonSpec, Lake Forest, CA). Direct infusion was performed using a syringe pump (Harvard Apparatus, Holliston, MA) with a flow rate of 2.5 uL/min into an electrospray source (Analytica, Branford, CT). Before samples were run, mass scale calibration was

performed with ubiquitin to achieve a mass resolution of 50,000-160,000 full-width at half maximum (FWHM) and mass accuracy of ±3-5 parts per million (ppm) (Sharp et al., 2003). Collisional dissociation was performed as previously described (Sharp et al., 2003).

Each sample was prepared for mass spectrometry analysis by mixing water:acetonitrile (ACN):acetic acid in a 49:49:2 ratio (by volume). The spectra were taken in broadband mode in a mass to charge range of 400 to 2500 m/z units. The ions were allowed to accumulate in an external hexapole for 2.5 seconds before being pulsed through an ion guide to the analyzer cell of the instrument. Each spectrum was an average of 10 scans and the frequency and amplitude of each ion was extracted from the transient signal by the 256 fast Fourier transform (FFT) algorithm with the Hann window on. Manual analysis of the spectra was done to confirm oxidation had occurred in a time-dependent manner and to a meaningful degree.

E. LC-MS/MS Analysis

All (1D) liquid chromatography tandem mass spectrometry (LC-MS/MS) spectra were acquired by a Famos/Switchos/Ultimate High Performance Liquid Chromotography (HPLC) System (Dionex, Sunnyvale, CA) coupled to a LCQ-DECA quadrupole ion trap mass spectrometer (Thermo Finnigan, San Jose, CA) equipped with a nanospray source (Thermo Finnigan) (Uchiki et al., 2004).

Each peptide sample was prepared for tandem mass spectrometry analysis by drying down the C18light elution and redissolving in 100% HPLC water. The spectra were taken in the data dependent mode over a parent mass to charge range of 400 to 2000

m/z units for 2 hours. Dynamic exclusion was enabled (repeat count 2), where the four most abundant peaks in every MS scan were subjected to MS/MS analysis. The DBDigger 1.14 algorithm (Tabb, 2005) was used to search all MS/MS spectra as described in the next section. It was preferred over SEQUEST because of its ability to handle and detect multiple oxidation events in a single peptide.

<u>F. Data Analysis</u>

LC-MS/MS spectra were submitted to the database identification algorithm DBDigger. An in-house created database (BSA_HSA_plus_contams_and_variants.fasta), which contained 10 protein variants and 46 common contaminants was used for all searches. The search was performed using the following parameters: Parent Mass Tolerance, 3.0 Da; Fragment Ion Tolerance, 0.5 Da; Peptide Mass Range of 400-10000 Da; Theoretical Spectra of B and Y ions only; Protein Cleavage by trypsin; up to 2 missed cleavage sites; fully tryptic peptides only. DBDigger identifications were sorted and filtered by DTASelect (Tabb, 2001) as follows: 2-peptide level, delCN of at least 0.08, charge state of 1-4 and cross correlation scores (Xcorrs) of at least 23.76 (+1), 28.07 (+2) and 43.64 (+3). Manual validation of the DTASelect output files, by verifying the Xcorr score of the peptides with oxidation events were significantly above the filter for that particular net charge, were done to insure confidence in the results

II. Molecular Dynamics Simulations

Molecular dynamics simulations are enlisted to study the time-dependent properties of molecular systems (Karplus and McCammon, 2002). They allow for the

visualization of dynamics of atomic phenomenon which otherwise cannot be seen (Karplus and McCammon, 2002). The Tanford Transition of β-lactoglobulin is a well studied conformational change, but it can be verified and possibly further analyzed by this method as continued improvements to its methodology has advanced simulations of typical protein environments and can essentially better our understanding of how this system works (MacKerell et al., 1998).

A. Molecular Dynamics Software

All of the molecular dynamics simulations described in this study were generated by the CHARMM 33a (Brooks, 1983 *Chemistry at HARvard Molecular Mechanics developmental version 33a1,* 2006) program with an all atom potential energy function (MacKerell et al., 1998). CHARMM has become widely popular for modeling the structure and behavior of molecular systems and is a program that utilizes empirical energy calculations based on internal coordinate and pairwise nonbond interaction terms (Brooks et al., 1983*).*

B. CHARMM Potential Energy Function

Empirical energy calculations promote the conversion of geometric objects into physical objects allowing the dynamic study of protein systems. The simplicity of the potential energy function in empirical energy calculations is optimized to extract relevant information accurately and within an acceptable length of time. A potential energy function is described by the summation of various energy terms. The terms can be separated into internal energy terms and nonbonded interaction energy terms (Brooks et

al., 1983). Internal energy terms comprise of bond stretching, angle bending, dihedral torsion angle and improper torsion angle energy terms (Brooks et al., 1983). Nonbonded interaction energy includes terms for van der Waals and electrostatic interactions as well as a term for Urey-Bradley 1,3- distance. Figure 2.2 illustrates the source of these energy terms.

In equation form, the empirical energy function looks like:

$$\color{red}{\sum k_b(b-b_0)^2 + \sum k_\theta(\theta-\theta_0)^2 + \sum k_\phi(\phi-\phi_0)^2 + \sum k_\chi[1+\cos(\eta\chi-\delta)]} \color{blue}{+ \sum k_{UB}(S-S_0)^2 + \sum \left(\varepsilon_{ij}\left[(r_{ijmin}/r_{ij})^{12} - (r_{ijmin}/r_{ij})^6\right] + q_iq_j/\varepsilon_1 r_{ij}\right)}$$

It is derived by summation of the individual bonded (red) and nonbonded (blue) energy terms. $k_b, k_\theta, k_\phi, k_\chi$ and k_{UB} are the bond, angle, improper torsion angle, torsion (dihedral) angle and Urey-Bradley force constants, respectively; b, θ, ϕ, χ and S are the bond length, bond angle, improper torsion angle, torsion angle and Urey-Bradley 1,3-distance, respectively, the subscript 0 indicating values at equilibrium (MacKerell, 1998) ε_{ij} is the geometric mean of the Lennard-Jones well depth of atom i and j, r_{ijmin} is the arithmetic mean of the distance of i and j at the Lennard-Jones minimum, r_{ij} is the distance between atoms i and j, qi and qj are charges on atoms i and j respectively and ε_1 is the effective dielectric constant (MacKerell et al., 1998).

C. Model Building

1. Initial Coordinates and Modifications

Two models of the closed β-lactoglobulin structure were built using initial coordinates obtained from a crystal structure of β-lactoglobulin at pH 6.2 (PDBID:
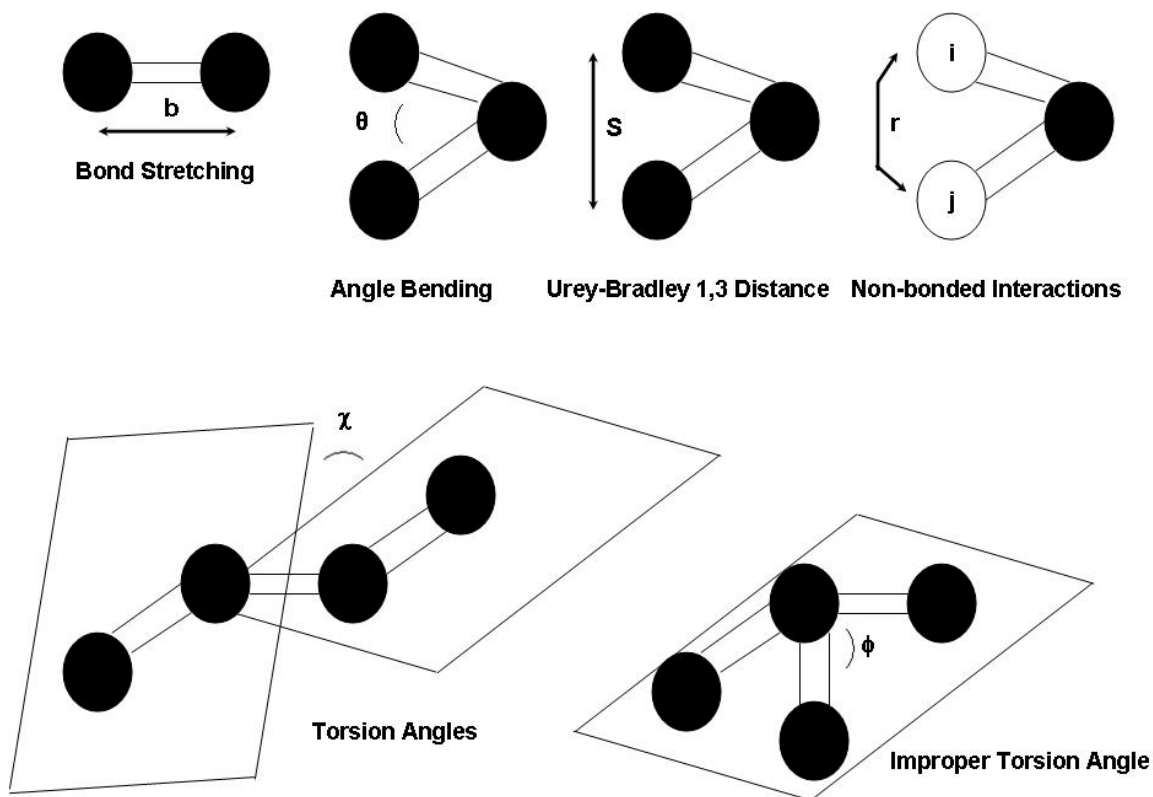
Figure 2.2: Bonded and Non-bonded Energy Terms.

Examples of bonded and nonbonded energy terms used in the all atom potential energy function in CHARMM. For bonded energy terms: **b** is the bond length between two atoms, θ is the angle between two bonds, χ is the torsion (dihedral) angle between atoms separated by three covalent bonds, ϕ is the improper torsion angle. The improper torsion angle is an angle that is not defines by four atoms that are connected sequentially. For nonbonded energy terms: **S** is the distance between atoms separated by two bonds and **r** is the distance between two atoms having charges q_i and q_j.

3BLG). Both models were identical except for the protonation state of Glu89. Histidine residues have a pKa of around 6.5 so their ionization states were manually determined. Upon a closer investigation of the structure, HIS146 was determined to be protonated on ND1 (*HSD*) and HIS161, protonated on NE2 (*HSE*). All other titratable residues were set to their ionized forms.

Initial coordinates from a PDB file are generally only of the heavy atoms in the system (Polygen, 1988). In order for CHARMM to compute the empirical energy functions, all atoms of the system need to have defined coordinates (Polygen, 1988). To satisfy this requirement the HBuild command was used to build and optimize the undefined hydrogen coordinates (Polygen, 1988). A patch was introduced for each disulfide bond (*Cys66-Cys160, Cys106-Cys119*). Finally, the atoms were centered about the moments of inertia of the structure and underwent a brief minimization by steepest descents (*SD*) for 5000 steps with a total energy change tolerance of 0.1 kcal/mol (*system exits minimization if total energy falls at or below tolerance value*) (Polygen, 1988).

2. Solvation

When building a model of a system, it is not only important to accurately represent the protein but also its environment. Though it is not yet possible to *perfectly* replicate the environment of typical proteins, models are built with proteins in an aqueous solvent environment since at least some portion of a typical protein is exposed to these conditions (Adcock and McCammon, 2006). Two types of methods are commonly used to solvate a system for molecular dynamics simulations. Implicit solvent gives the effect of solvation without the presence of actual water molecules by only introducing an

23

effective dielectric constant where explicit solvent consists of actual water molecules in addition to a dielectric constant.

The models in this study were solvated with explicit solvent of the TIP3P (transferable intermolecular potential 3P) (Jorgensen, 1983) potential function. An explicit solvent was chosen because it allows for a more detailed simulation of the system, providing data about specific interactions that together may facilitate the conformational change of β-lactoglobulin. The TIP3P potential function is the simplest of the explicit solvent potential functions in that it contains three interaction sites (Jorgensen et al., 1983). It was chosen for these models arbitrarily because of its simplicity but other potential function could have also been used (*ie. TIP4P, TIP5P, SPC*).

Within the explicit solvent models there are two different boundary conditions that can be applied to the solvent; periodic boundary conditions and stochastic boundary conditions. Periodic boundary conditions were applied to the solvent to eliminate surface effects from computations. The protein was placed in the middle of a cubic solvent box of dimensions 63Å x 58Å x 57Å for the deprotonated and 64Å x 58Å x 58Å for the protonated model system. The dimensions of the cubic solvent box differ slightly for both models most likely due to the minimization of the model system prior to solvation. To build the cubic solvent box, a coordinate file of a cube containing 216 3-site (*TIP3P*) water molecules was used. The coordinates of atoms in this cube were constructed from an equilibrated 4-site (*TIP4P*) model at 1 atm pressure and .0334 molecules/ampere with an edge length of 18.856 Å. These cubes were added until the calculated dimensions of the system were met. The distance between the protein and the outer boundary of the solvent box was set to 8 Å and any water molecule more than 8 Å away from the origin

or that overlapped (≤ *2.2 Å away*) with the protein or previously introduced water
molecules was deleted.

3. Neutralization

The net negative charge of the model systems was neutralized by replacing water
molecules in energetically favorable positions with positively charged sodium ions.
Chosen water molecules had to be at least 5.5 Å away from the protein and/or other
waters or ions. A total of 8 water molecules were replaced in the protonated model and 9
were replaced in the deprotonated model.

The ions were briefly minimized by SD for 10 steps and adopted basis Newton-
Raphson method (*ABNR*) for 25 steps. The final deprotonated model system contained
2595 protein atoms, 9 ions and ~19900 solvent atoms (water) and the final protonated
model system contained 2596 protein atoms, 8 ions and ~20700 solvent atoms (water).

4. Minimization

After fixing all bonds involving hydrogens to a tolerance of $1.0e^{-6}$ Å and
maximum iteration of 1500, the solvent was energy minimized. It was first minimized by
SD for 500 steps holding the protein and sodium ions fixed at a harmonic force constant
of 50 kcal/mol/$Å^2$. A constant dielectric non-bond energy of 1.0 kcal/mol/$Å^2$ was applied
with parameters as follows: distance cut off for generating the list of pairs = 14 Å;
distance cut off for smoothing function = 12 Å; maximum allowable distance to be
included in the image atom list = 14 Å.

The solvent was then energy minimized by ABNR for 5000 steps holding the protein and sodium ions fixed at a harmonic force constant of 20 kcal/mol/Å^2 with the same parameters as the SD minimization.

5. Heating

Each model system was heated for 10 ps using a 2 fs timestep starting from 50 K to a final temperature of 298 K. Heating was performed to bring the system to equilibrium in preparation for dynamics; hence, it was only performed for a short period of sampling time. A ±5 K deviation buffer was allowed for the final temperature and averages of the major energy values were calculated every 1 ps. Cutoff values were the same as in the minimization step. The non-bonded and image update lists were updated heuristically and no H-bond list was generated.

6. Dynamics and Data Analysis

Dynamics was performed with the same parameters as in the heating step except the temperature was held constant at 298 K and trajectories were generated in increments of 200 ps for a total of 3 ns. All data analysis was performed using various CHARMM scripts to extract and visualize interesting and potentially significant results.

III. Conclusions

The materials and methods outlined above provide a step-by-step protocol for an approach employing experimental oxidative surface mapping and computational

molecular dynamics simulation methods in visualizing the conformational change of β-lactoglobulin. An overall scheme of the approach to this study is shown in Figure 2.1. The rational for incorporating multiple approaches to study a well characterized model system is based on the premise that coupling of two methods will provide stronger insight than either one of the techniques alone. Subsequent chapters provide a discussion of the results obtained from each technique individually before making an attempt at coupling the results to extract more meaningful data.

Chapter 3: Oxidative Surface Mapping Measurements of pH Induced Conformational Transitions of β-lactoglobulin A

β-lactoglobulin undergoes a pH induced conformational change (Tanford et al., 1959). The EF loop of this protein covers the calyx of its hydrophobic core in acidic conditions but pulls away from the calyx opening under basic conditions.

Though the function of β-lactoglobulin still remains a mystery, its structural similarity with retinol-binding protein (Papiz, 1986) adopted its classification into the lipocalin superfamily, of which most members bind hydrophobic molecules and serve as carrier proteins. Simulations performed on retinol-binding protein displayed bending of the α-helix (Aqvist, 1986); hence care was taken to monitor the movement in this region as well as that of the loop.

Modification of proteins by reactive oxygen species (ROS) are the focus of several investigations because of their possible correlation with age and disease (Stadtman and Levine,2003). Inducible, stable, covalent modifications of proteins have been utilized with success in determining the higher order structure of proteins in solution (Sharp et al., 2003). Oxidative surface mapping using chemically generated hydroxyl radicals (Sharp et al., 2003) was employed to study the conformational change of the β-lactoglobulin protein in this study. The protein was induced to change conformation by subjecting it to various pH environments and allowing those residues exposed to solvent and hydroxyl radical to be oxidized. Using this approach, a more detailed, pH by pH , visualization of the change in conformation of the β-lactoglobulin protein is possible.

I. Oxidation by Hydroxyl Radical

The oxidation reaction is initiated by exposing pH-adjusted protein samples containing 10% H_2O_2 (by volume) to UV light. Hydrogen peroxide is a reactive oxygen species. When attacked by a photon of UV light, a H_2O_2 molecule undergoes photodissociation into two hydroxyl radicals. A radical is a highly unstable group of atoms that contains an unpaired electron. The hydroxyl radical is highly reactive and seeks to stabilize itself by pairing its unpaired electron. It achieves this by "stealing" a hydrogen atom from various regions of the β-lactoglobulin protein or from other molecules, such as H_2O present in the sample constituents. The amino acid or molecule that gives up its hydrogen atom is now converted into a radical and has the potential of eventually being oxidized from the solvent (water) or dissolved molecular oxygen. Hydroxyl radicals can also pair their unpaired electron by reacting with one another. This not only stabilizes these highly reactive radicals but also reforms H_2O_2. This method of oxidation is appropriate as it has been successfully utilized for the surface mapping of proteins (Sharp et al., 2003; Sharp et al., 2005) and its yield is not influenced within the pH range of this study (Hubner and Roduner,1998).

II. Oxidizable Amino Acids

Hydroxyl radicals can cause modifications of proteins by attacking side chains of residues or by forming covalent cross linkages and can cause protein fragmentation by attacking backbone bonds (Garrison, 1987). The rates at which these modifications occur are different with side chain modifications occurring more rapidly than backbone bond modifications and the degree of modifications is dependent on competition from sample

constituents, the amino acid composition and conformation of the protein (Garrison, 1987).

All amino acids are susceptible to oxidative modification; however, aliphatic, sulfur-containing and aromatic amino acids predominate over others. Attack by a hydroxyl radical abstracts a hydrogen and initiates a chain of events that eventually lead to an oxidative modification. Hydroxyl radicals preferentially abstract hydrogens from those amino acids that are left the most stable after attack. This is why aliphatic, sulfur-containing and aromatic amino acids are most susceptible to hydroxyl radical attack. Table 3.1 summarizes the reactivity rates of the twenty amino acids for reactions with hydroxyl radicals in aqueous solution (Buxton et al., 1988). In the following discussion, aliphatic amino acids refer to isoleucine (Ile), leucine (Leu), Valine (Val) and alanine (Ala); sulfur containing amino acids refer to cysteine (Cys) and methionine (Met); and aromatic amino acids refer to phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp).

Oxidation at aliphatic side chains originates with hydroxyl radical attack and abstraction of a non-labile hydrogen, leaving a carbon centered radical in the side chain. Next, the carbon radical is rapidly converted to a peroxyl radical by reacting with molecular oxygen in solution (Stadtman and Levine,2003). This peroxyl radical instigates further reactions eventually resulting in a hydroxyl substitution at the side chain (Stadtman and Levine,2003). The hydroxyl substitution adds an oxygen atom to the side chain producing a mass shift of about 16 atomic mass units (15.9949 amu) which can be detected and localized by mass spectrometric analysis.

Oxidation at side chains of sulfur containing amino acids occur by addition at the sulfur atom. Methionine residues can potentially undergo two oxidation events at their

Table 3.1: Reactivity Rates of Amino Acids.

Amino Acid Side Chain	k (L mol^{-1} s^{-1})
Ala	7.7×10^7
Arg	3.5×10^9
Asn	4.9×10^7
Asp	7.5×10^7
Cys*	3.4×10^{10}
Gln	5.4×10^8
Glu	1.6×10^8
Gly	1.7×10^7
His*	5.0×10^9
Ile*	1.8×10^9
Leu*	1.7×10^9
Lys	3.5×10^8
Met*	8.3×10^9
Phe*	6.5×10^9
Pro*	4.8×10^8
Ser	3.2×10^8
Thr	5.1×10^8
Trp*	1.3×10^{10}
Tyr*	1.3×10^{10}
Val	7.6×10^8

Reactivity rates of the twenty amino acids with hydroxyl radical in aqueous solution. Those amino acids indicated in red, blue and green are aliphatic, sulfur containing and aromatic in nature, respectively. Starred amino acids were found oxidized in a previously reported study (Sharp, 2003).

sulfur atom generating methionine sulfoxide and methionine sulfone (Vogt, 1995). Each addition event produces a mass shift of about 16 amu for a potential mass shift of 32 amu (31.9898 amu) for each methionine residue. Because previous studies have reported instances of methionine oxidation regardless of solvent accessibility or direct hydroxyl radical attack, (Sharp et al., 2003; Kiselar et al., 2002) oxidation events at methionine side chains are not necessarily considered indicative of solvent accessibility. Cysteine, another sulfur containing residue, can also undergo multiple oxidation events at its sulfur atom (Paget and Buttner, 2003). Oxidation is initiated by abstraction of labile hydrogens via hydroxyl radical attack followed by additions at the sulfur atom by interactions with molecular oxygen, producing sulfenic, sulfinic and sulfonic acid derivatives (Paget and Buttner, 2003). With a total of three potential oxidation events, each cysteine residue can potentially cause a mass shift of 47.987amu.

Oxidation at aromatic amino acid side chains occurs by direct addition of the hydroxyl radical into their unsaturated cyclic ring. This addition creates a side chain carbon radical which reacts with molecular oxygen in solution to form a peroxyl radical and eventually results in a hydroxyl addition (Garrison, 1987).

The above mechanisms are explained assuming abundance of molecular oxygen however, in the absence of molecular oxygen, carbon centered radicals which are in close structural proximity to one another can react to stabilize themselves producing covalent cross linkages (Garrison, 1987).

A potential site of hydroxyl radical attack other than the side chains of amino acid residues is at the α-carbon of the protein backbone (Garrison, 1987). Attack at this atom leads to the eventually cleavage of the peptide bond but its frequency is low due to the

abundance of molecular oxygen and water molecules which can quickly quench hydroxyl radicals before they can make it to the protein backbone, highly reactive side chains which stabilize those hydroxyl radicals that weren't quenched by molecular oxygen, the steric inaccessibility of the α-carbon hydrogen and the planar state of the amino acid side chain (Hawkins and Davies, 2001).

III. Identification of Oxidation Events

Electrospray ionization transfers protonated peptides from solution into the gas phase. Without the presence of at least one charge, peptides cannot be detected by the mass spectrometer. Protons attach where binding is most stable, hence basic amino acids are the most desirable. After ionization, the peptide can be induced to dissociate in the mass spectrometer by collision with helium gas. Figure 3.1 displays the components of the LC-MS/MS used to perform tandem mass spectrometry on the samples prepared for this study. Samples are taken up by the auto-sampler and directed to flow over the microcolumn. The microcolumn contains C18 stationary phase, which creates a highly hydrophobic medium and binds peptides easily. When a gradient of high water to high organic solvent is applied to the microcolumn by the HPLC, peptides that are shorter or less hydrophobic are washed off of the column first and analyzed by the mass spectrometer before those that are longer or more hydrophobic. Dissociation of the peptide can occur at any amide bond and even though only one event occurs per peptide, the presence of multiple copies of the same peptide and the ability to cleave at more than one amide bond results in multiple fragments of the same peptide. The most common ions observed from collisional induced dissociation of peptides are b and y ions. B ions
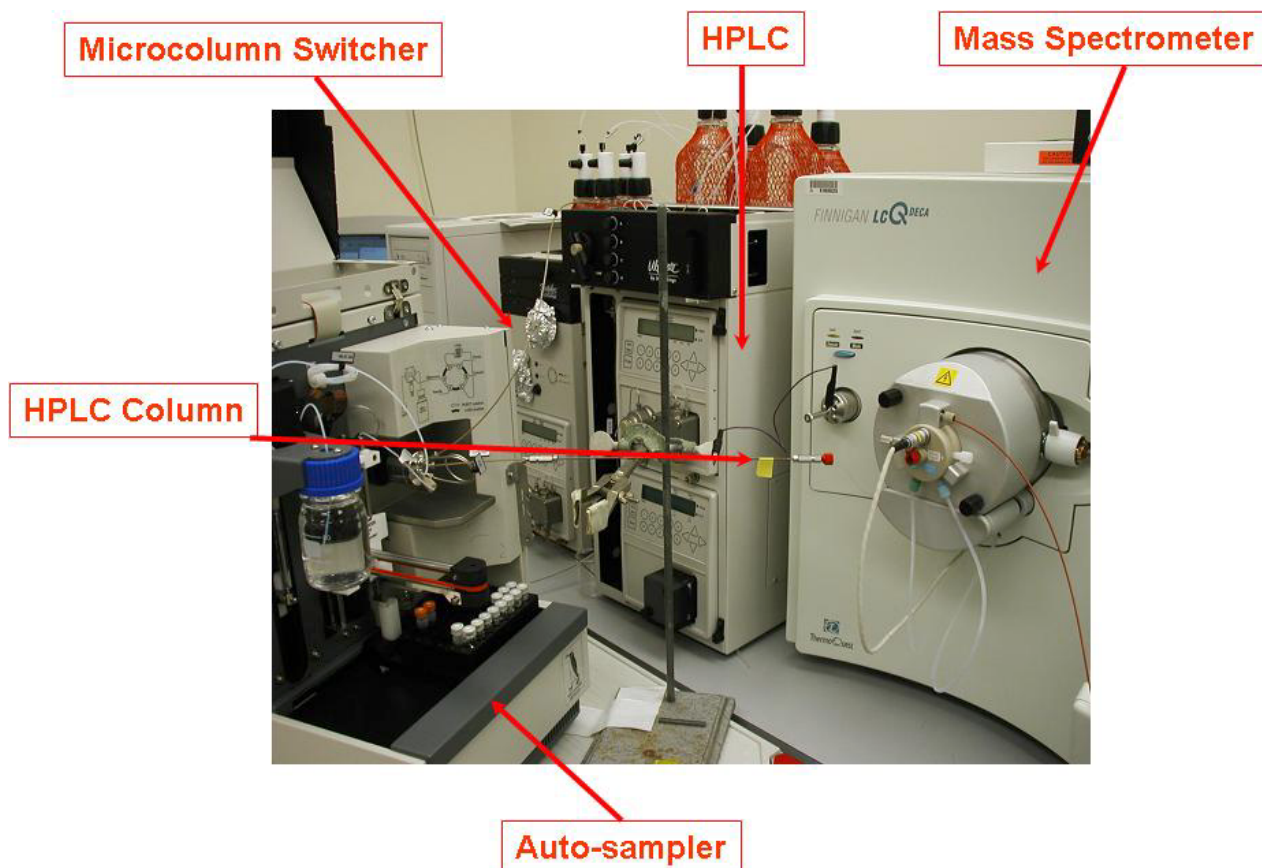
Figure 3.1: LC-MS/MS Components.

The different components of a Famos/Switchos/Ultimate High Performance Liquid

Chromotography System coupled to a LCQ-DECA used to perform LC-MS/MS on the

samples in this study. The samples are taken up by the auto-sampler and allowed to flow

over the microcolumn. The HPLC applies a gradient of high water to high organic

solvent allowing those peptides that are shorter or less hydrophobic to be analyzed by the

mass spectrometer before those that are longer or more hydrophobic. The mass

spectrometer takes a full scan of the ionized peptides that enter its analyzer after which it

isolates and subjects the four most intense peaks to dissociation by helium gas.

are the result of fragmentation at the amide bond such that the charge is retained on the N terminus and Y ions are the result of fragmentation at the amide bond such that the charge is retained on the C terminus. Manually, the sequence of a peptide is determined by visual inspection of the peaks from a mass spectrum, most of which are generally b or y. The sequence of a peptide can be determined by addition or subtraction of the monoisotopic masses of amino acids from peaks of a mass. Any peak which is +16, +32 or +48 Daltons more than the monoisotopic mass of that residue identifies a single, double or triple oxidation event on that amino acid. Figure 3.2 displays a sample MS/MS spectrum of an oxidized peptide. Only the B and Y ions are displayed in the MS/MS spectrum as other ions were filtered out during the DBDigger search. The methionine residue is oxidized in this peptide and is demonstrated by the 16 Da shift in monoisotopic mass for the b10 and y17 ions. Table 3.2 provides the monoisotopic mass of each individual amino acid (http://haven.isb-sib.ch/tools/isotopident/htdocs/aa-list.html) as a reference. Figure 3.3 shows a sample output of the DTASelect software for the pH 7 sample. The MS/MS spectrum shown in Figure 3.2 (**B**) is of the first peptide highlighted by the purple circle.

IV. Surface Mapping Measurements

Oxidative surface mapping of the β-lactoglobulin protein at various pH values including those encompassing the Tanford Transition (pH 7 to pH 8), revealed results that appear to correlate with the change in conformation of the protein. Table 3.3 provides a summary of the results from which several overall observations can be made.

A.
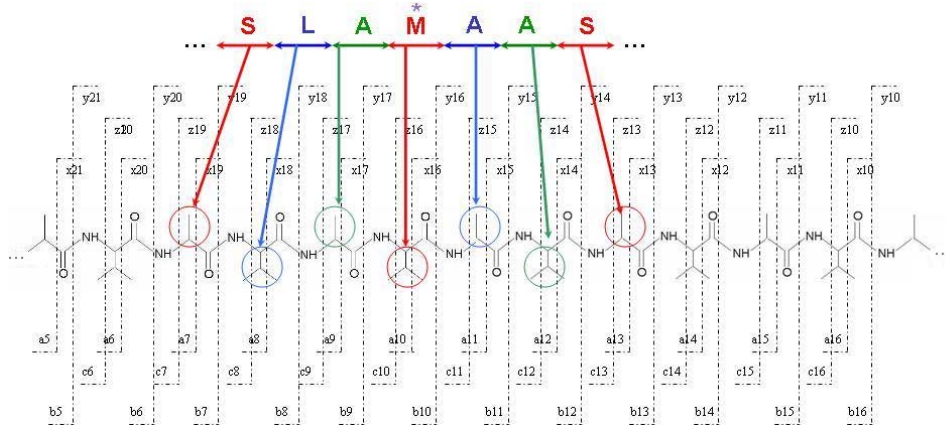
B.
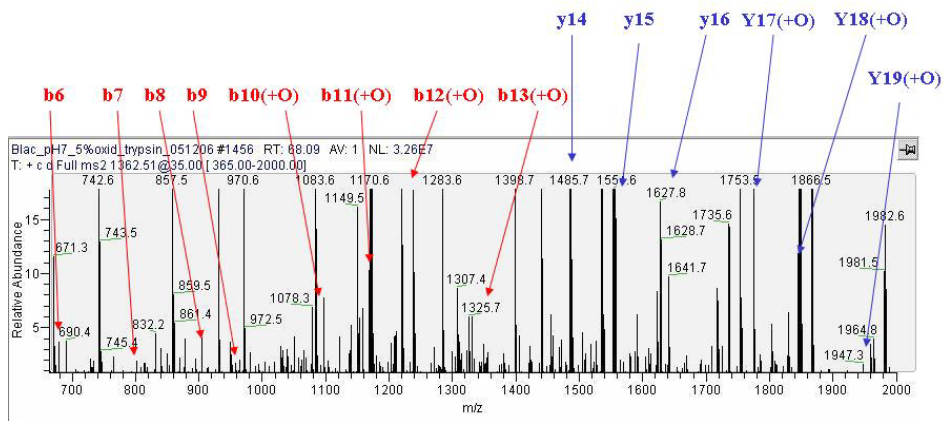
Figure 3.2: Sample Peptide Fragmentation Spectrum.

A partial diagram of the cleavage sites of peptide 15-40 of β-lactoglobulin A (**A**). B and

Y ions are the most common and the methionine residue has a single oxidation event. The

partial peptide sequence displayed is of the precursor ion of the MS/MS spectrum (**B**).

MS/MS spectrum of peptide 15-40 of β-lactoglobulin A from quadrupole ion trap (**B**).

Peptides are labeled according to the site of cleavage and which of the resulting

fragments retained the charge. The methionine had an oxidation event as its monoisotopic

mass shifted by 16 Da.

36

Table 3.2: Monoisotopic Mass of Amino Acids.

Amino Acid	Monoisotopic Mass (Da)
Ala	71.037
Arg	156.101
Asn	114.043
Asp	115.027
Cys	103.009
Gln	128.059
Glu	129.043
Gly	57.021
His	137.059
Ile	113.084
Leu	113.084
Lys	128.095
Met	131.04
Phe	147.068
Pro	97.053
Ser	87.032
Thr	101.048
Trp	186.079
Tyr	163.063
Val	99.068

The monoisotopic mass of the twenty amino acids used for manual determination of the sequence of a precursor peptide after dissociation of a precursor ion and analysis of the MS/MS spectra (as illustrated in Figure 3.2).

Figure 3.3: Sample DTASelect Output.

Sample DTASelect Output for a β-lactoglobulin sample at pH 7. The first peptide sequence on the list (peptide 15-40) has the highest score and also contains an oxidation event on Met24. This demonstrates the computational alternative to the manual approach displayed in Figure 3.2 for this same precursor peptide.

Table 3.3: Oxidative Surface Mapping Results for β-lactoglobulin A as a Function of pH.

AA	pH 2	pH 5	pH 6	pH 7	pH 8	Urea-denatured
Leu 1	+1o	+1o		+1o		
Ile 2	+1o	+1o		+1o		
Met 7	+1o	+1o	+1o,+2o	+1o	+1o	+1o
Leu 10		+1o			+1o	
Met 24	+1o	+1o	+1o	+1o	+1o	+1o
Ile 29						+1o
Leu 54						+1o
Leu 57						+1o
Trp 61					+1o	+1o
Cys 66			+3o	+3o	+3o	+3o
Ile 71			+1o		+1o	+1o
Ile 72					+1o	+1o
Ile 78					+1o	
Phe 82						+1o
Cys 106						+1o
Met 107			+1o		+1o	+1o
Leu 117					+1o	
Cys 121			+3o		+3o	+3o
Leu 122						+1o
Pro 126						+1o
Leu 133		+1o			+1o	+1o
Phe 136				+1o	+1o	+1o
Leu 140					+1o	
Phe 144		+1o		+1o		
Met 145	+1o	+1o, +2o	+1o	+1o, +2o	+1o,+2o	+1o,+2o
Leu 149	+1o	+1o		+1o		
Phe 151	+1o		+1o	+1o	+1o	+1o
Leu 156					+1o	+1o
Cys 160	+1o	+1o, +2o, +3o	+3o	+1o, +2o, +3o	+3o	+3o
Ile 162		+1o		+1o		

The results from oxidative surface mapping of the protein at various pH values (2, 5, 6, 7 and 8) and under denatured conditions. The table lists those amino acids that were oxidized in any of the samples, however the β-lactoglobulin sequence contains more oxidizable amino acids. +1o, +2o and +3o represent 1, 2 or 3 oxidation events, respectively.

First, some residues are oxidized independent of pH. Taking into consideration that methionine oxidations are ambiguous, Cys160 is oxidized in the structure at all pH values. This residue forms a disulfide bond with Cys66 and is located in the C-terminal region of the protein. Though this region experiences considerably more flexibility then the β-strands (Kuwata et al., 1999) and is highly solvent exposed, oxidation of Cys160 across the pH spectrum of study suggests no significant solvent accessibility differences. Inversely, certain residues are found unoxidized throughout the spectrum of study, notably Leu54. An oxidation event on this residue was only seen in the denatured sample, again suggesting no significant solvent accessibility differences occur in this region.

Second, several residues exhibit pH sensitivity. For example, the highly reactive Trp61 and Phe136 are not oxidized under acidic conditions but become oxidized at pH 8 and pH 7, respectively; and successively thereafter. This implies that a change in conformation near Trp61 occurs between pH 7 and pH 8 such that the residue becomes de-protected from solvent and is allowed access to hydroxyl radical attack. A similar scenario is implied for Phe136, however the change in conformation that exposes this residue occurs between pH 6 and pH 7.

Lastly, the denatured sample displays the greatest degree of oxidation. This agrees with expectation as all potential sites of oxidation are solvent exposed under these conditions. The results however, don't show complete oxidation of the protein. Most of the residues that aren't oxidized are those that are not reactive, namely (Ile and Leu) however the highly reactive Phe144 was not found oxidized in this sample. Possible explanations for this could be: hydroxyl radicals generated via photodissociation of H_2O_2 got quenched by water/molecular oxygen (in the sample constituents) before they could

reach this residue, proteins containing oxidation events at this residues could have been washed through the SepPak after oxidation or digestion causing its abundance to drop to inadequate levels for detection by mass spectrometry or sequence dependent oxidation of residues may be inherent such that oxidation occurs preferentially when Phe is next to amino acid X but not when it is next to amino acid Y.

The discussion that follows focuses specifically on the movements of the EF loop and α-helix region. Figure 3.4 provides the β-lactoglobulin sequence (**A**) and structure (**B**) for reference. Comparisons between results obtained from a study of the same system conducted by Qin et al. (1998), from which the initial crystal coordinates were obtained, are included where appropriate.

V. pH Induced Global Transitions of the β-lactoglobulin Protein

A monomer at pH 2, the β-lactoglobulin protein undergoes dimerization via β-strand I near pH 3 (Taulier and Chalikian, 2001). Between pH 4 and 5, the protein undergoes further aggregation into an octomer before dissociating back to its dimer form just after pH 5 (Taulier and Chalikian, 2001). The protein remains in its dimer form at physiological pH until it undergoes irreversible unfolding above pH 9 (Taulier and Chalikian, 2001). A NMR based solution structure study of the β-lactoglobulin protein at pH 2 revealed that the monomeric structure at this pH is similar to that of each subunit of the dimer before the Tanford Transition (Kuwata et al., 1999).

A.

LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLR

VYVEELKPTPEGDLEILLQKWENDECAQKKIIAEKTKIPA

VFKIDALNENKVLVLDTDYKKYLLFCMENSAEPEQSLVCQ

CLVRTPEVDDELEKFDKALKALPMHIRLSFNPTQLEEQCH

I

B.

N-terminus, (residues 1-15)
Strand A (residues 16-27)
A-B Loop, (residues 28-45)
Strand B (residues 46-49)
B-C Loop (residues 50-52)
Strand C (residues 53-63)
C-D Loop (residues 64-69)
Strand D (residues 70-78)
D-E Loop (residues 79-83)
Strand E (residues 84-85)
E-F Loop (residues 86-89)
Strand F (residues 90-98)
F-G Loop (residues 99-100)
Strand G (residues 101-109)
G-H Loop, (residues 110-118)
Strand H (residues 119-124)
H-α helix Loop (residues 125-130)
α-helix (residues 131-142)
α-helix-I Loop (residues 143-150)
Strand I (residues 151-157)
C-terminus, (residues 158-162)

Figure 3.4: β-lactoglobulin A Sequence and Structure.

The protein sequence (**A**) for bovine β-lactoglobulin A. The stretch of sequence highlighted in yellow indicates residues of focus for change in conformation of the loop and that highlighted in green indicates residues of focus for change in conformation of the α-helix. The protein structure (**B**) of bovine β-lactoglobulin A with amino acid designations.

VI. pH Induced Local Transitions of the β-lactoglobulin Protein

A repeatedly visualized transition occurs at roughly pH 7. Its occurrence near physiological pH suggests it may be important in protein function. Movement of the protein is localized to the EF loop. At pH values below the Tanford Transition, the EF loop covers the calyx of its hydrophobic core and at pH values following the Tanford Transition, the EF loop has pulled away from the calyx opening exposing the rim and interior of its hydrophobic core.

Bending of the α-helix was seen from a molecular dynamics simulation of RBP (Aqvist et al., 1986), a protein structurally similar to β-lactoglobulin, and the orientation of the α-helix was found to be different than that of a x-ray structure when a solution structure dynamics was performed (Kuwata et al., 1999). These previous studies suggest local transitions near this region may also occur.

A. EF Loop Region

To confirm the movement of the EF loop using this technique, focus was turned to residues 61-78. These residues comprise β-strand D and part of C as well as the CD loop. Strands D and C are near the rim of the hydrophobic calyx and the loop that connects them is in close proximity to the EF loop before the Tanford Transition.

The experimental data indicates movement of the loop occurs between pH 7 and pH 8. This sharp pH transition is most likely because Glu89 has a pKa of 7.3. Trp61 and Ile71, 72 and 78 which were not oxidized at pH 7 become oxidized at pH 8. The Trp61 is highly reactive with hydroxyl radical, any change in the solvent accessibility of its side chain will instigate hydroxyl radical attack; hence it serves as a good probe for loop

movement. Cys66 which is located in the CD loop, is initially oxidized at pH 6 possibly due to its disulfide bond with Cys160 which is in a highly flexible and solvent exposed region of the protein structure. Findings by Qin et al. (1998) describe the solvent accessible surface area (SASA) of the disulfide Cys66 and Cys160 as being 35 fold greater than the one between Cys106 and Cys119. This is probably why only oxidations at the more exposed cysteine residues were detected in the pH induced samples. Ile71 which is located on the D strand near the rim of the hydrophobic calyx, is also oxidized at pH 6. This indicates that the Tanford Transition begins to initiate at this pH value but we may have missed detection of this peptide in the pH 7 sample. Other changes seen in the Qin study were not detected from the surface mapping measurements. The partial burial of Asp85 was not detected from the surface mapping measurements due to the relative unreactivity of this side chain with hydroxyl radical. The exposure of Leu87, though more reactive than Asp85, was also not observed in the surface mapping data. Surface mapping measurements were however able to confirm the inaccessibility of tyrosine residues as none of the four present in the β-lactoglobulin sequence experienced oxidation through the pH spectrum of study. Qin's analysis implicates the reorientation of the side chain of Met107 during the Tanford Transition. This was seen from the oxidative surface mapping measurements as initiating between pH 5 and pH 6. The reorientations, however, of the less reactive Ile84 and Leu39 are not observed. Those amino acid residues implicated by Qin as experiencing greater than 30 Å^2 shift in their solvent accessibility were not detected because of their relative unreactive side chains. Other differences are detected from the terminal regions. Leu1 which is highly exposed in the

structural studies by Qin, is never found oxidized. Conversely, Cys160 which is found oxidized in all samples, is determined to have only 9 $Å^2$ or solvent accessibility at pH 8.2.

B. α-Helix Region

To monitor the movement of the α-helix, focus was turned to residues 130-140. These residues comprise the length of the α-helix which is in close proximity to strands A, H, and I. In agreement with the second binding site described by Qin et al. (1998), the experimental data clearly shows initial movement of the α-helix away from the protein body at pH 5 and complete movement away by pH 8 as indicated by the oxidation all potentially oxidizable amino acids in that region. Leu133, Phe136 and Leu140 are located in the inner cleft of the α-helix facing Cys121 and Leu117. The oxidation of both of these residues by pH 8 suggests a pulling away rather than a denaturing of the α-helix. According to Qin, the increased reactivity of Cys121 at higher pH may not be directly related to the Tanford Transition but can be attributed to the change in environment surrounding the SH group of this residue. As the proteins environment becomes more basic the area surrounding the SH group of Cys121 becomes more negatively charged. Though the solvent accessibility of this residue does not change, the increase in negative charge surrounding its SH group increases its susceptibility to modification.

VII. Denatured Control

This experiment served as a positive control for all the possible oxidation sites on the protein. It also confirmed the degree of oxidation is a result of solvent accessibility. Under these conditions the entire β-lactoglobulin protein is completely exposed to

solvent and hydroxyl radical attack. The degree of oxidation should be the greatest here compared to that at any pH environment. The results confirmed this as the denatured sample shows oxidation events throughout the length of the protein and to a greater degree than for any pH environment. Additionally, amino acids that were not oxidized within the pH spectrum of study were found to be under these conditions. Not all possible oxidation sites were shown to be oxidized. Possible explanations for this are: the peptides were washed through the SepPak after oxidation or digestion, the signal for the peptide was not picked up by the mass spectrometry or the spectrum was not captured in the filter when the LCQ output was run through DTASelect. These explanations not only apply to the denatured sample but also to those of the pH spectrum of study.

VIII. Lysozyme Control

This control experiment was performed to demonstrate that the pH dependence of the β-lactoglobulin protein is due to the protein and not the hydroxyl radical. Hydroxyl radical yield is not affected by the pH spectrum of this study (Hubner and Roduner,1998). Its yield is affected at very high (>10) or very low (<2) pH values. Lysozyme does not undergo a pH dependent conformational change so oxidative surface mapping of this protein at pH 2 and pH 7 was performed to demonstrate no change in the degree of oxidation. The results confirmed pH dependence is due to the protein as the lysozyme experiments at both pH environments gave similar results (data not shown).

IX. Conclusions

The results obtained from the study of the β-lactoglobulin protein in various pH environments demonstrated the conformational shift termed the Tanford Transition. In addition, movement of the α-helix region was visualized as solvent accessibility and oxidized sites in that region increased between pH 2 and pH 5.

The conformational shift of the β-lactoglobulin protein was localized to certain regions of the protein as shown from the results. Those amino acids that remained either always oxidized or never oxidized throughout the pH spectrum of study suggested no significant change in solvent accessibility of these regions.

Chapter 4: Molecular Dynamics Simulations of Protonated and Deprotonated Glu89 Models of β-lactoglobulin A Starting From the pre-Tanford Transition Conformation

In order to visualize the conformational change of β-lactoglobulin, two models were built using initial coordinates obtained from a crystal structure of the protein at pH 6.2 (PDB ID: 3BLG). This particular coordinate file was chosen because of its success in demonstrating the Tanford Transition, upon slight perturbation, during molecular dynamics (MD) simulations (Eberini et al., 2004).

Tanford was the first to discover interesting behavior at a carboxylic side-chain in the EF loop; it had an unusually high pKa of 7.3 and appeared to trigger movement of the EF loop upon titration such that it was buried when in an acidic environment but exposed in a basic environment (Tanford et al., 1959). This carboxylic side-chain was later identified as being that of glutamic acid 89 (Glu89) (Brownlow et al., 1997; Qin et al., 1998). Previous molecular dynamics simulations have implicated the protonation state of the Glu89 residue as an important trigger for the Tanford Transition (Fogolari et al., 2005; Eberini et al., 2004). As a result of this finding, the two models built for this study were identical except for the protonation state of this amino acid.

These models were built under the premise that the structure with the deprotonated Glu89 amino acid side chain, starting from a structure not favorable to the protonation state of Glu89, would undergo the Tanford Transition and the structure with the protonated Glu89 amino acid side chain would remain unchanged. The fluctuations in and around the α-helix were also monitored as previous studies have implicated this

48

region in movement and even suggest it as a location of a second binding site (Monaco et al., 1987).

I. Molecular Dynamics Simulation Results

3ns and 2.4ns molecular dynamics simulations were performed on the deprotonated and protonated models (built as described in detail in Chapter 2), respectively. As predicted, the Tanford Transition was observed during the simulation of the deprotonated model but not of the protonated model. Movements in the α-helix region were negligible upon visual examination of both simulations. Figure 4.1 provides a visual display of the changes that were observed from both MD simulations.

The following discussion is segmented into an analysis of solvent accessibility and free energy; based on the change in average distance between residues in and around regions of suspected or confirmed conformational change. The discussion includes a comparison with results from previous molecular dynamics studies of the same protein system (Fogolari et al., 2005; Eberini et al., 2004) as well as solvent accessibility data presented in the study from with the initial coordinates were obtained (Qin et al., 1998).

A. Solvent Accessibility

Solvent accessibility changes were explored using the GetArea 1.1 (Fraczkiewicz and Braun, 1998) software available online through the Sealy Center for Structural Biology at the University of Texas Medical Branch (http://www.scsb.utmb.edu/cgi-bin/get_a_form.tcl). This software takes in Cartesian coordinates in PDB format and

α-Helix   **EF-loop**

Protonated

Cα RMSD Total 1.076Å
Cα RMSD Helix .487Å
Cα RMSD Loop .299Å

De-protonated

Cα RMSD Total 1.645Å
Cα RMSD Helix .288Å
Cα RMSD Loop 1.398Å
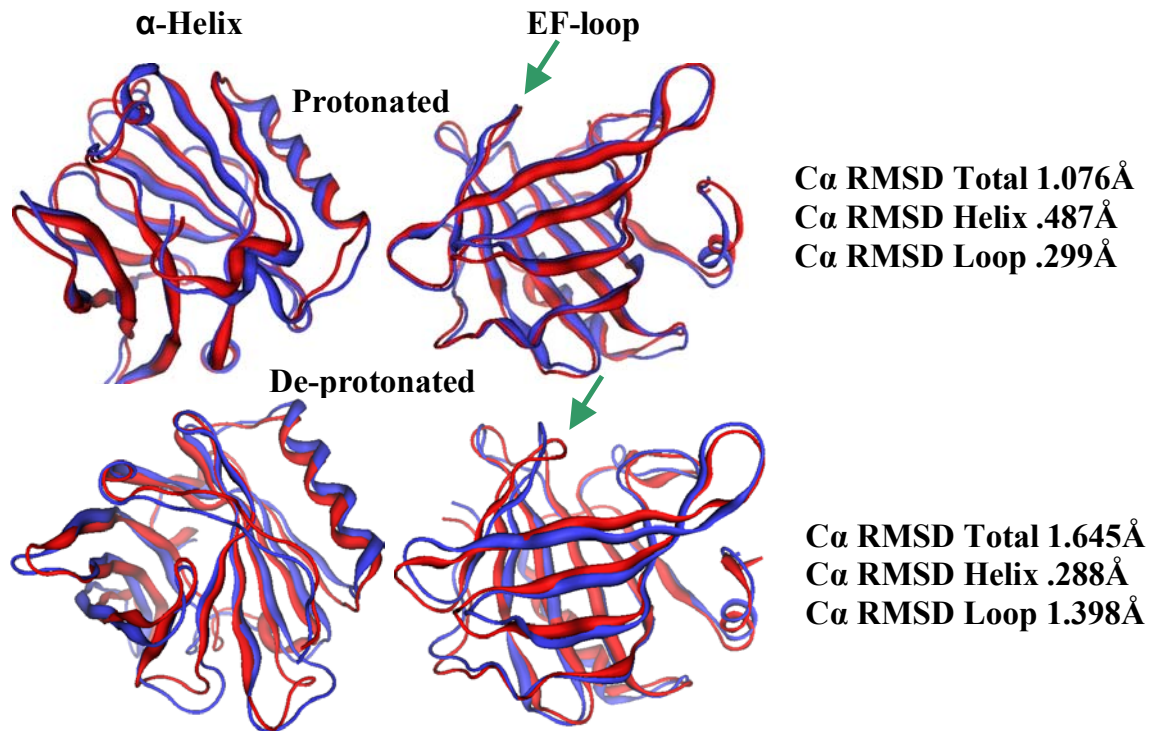
Figure 4.1: MD Simulation Structures.

MD simulation structures from both the protonated and deprotonated Glu89 models. The left column displays the α-helix and the right column displays the EF-loop region. The structures in the top row are results from the protonated model and those in the bottom row are results from the deprotonated model. Red represents the conformation before simulation and the blue represents the conformation after a 3.0 (deprotonated) or 2.4ns (protonated) simulation. RMSD between Cα atoms of the respective red and blue structures using MOE are presented. Total, Helix and Loop refer to residues 1-162, 85-90 and 130-140, respectively.

provides the solvent accessible surface area (SASA) of a protein. Output can be varied to include different levels of detail ranging from per atom to whole protein SASA calculations. The software was set up such that it would output SASA in a per residue format and provide the following information: residue name, residue number, residue SASA, apolar SASA, backbone SASA, side chain SASA, ratio of side chain SASA to the "random coil" value of that residue and its location (inside or outside) based on the previous ratio (<20% is considered in, >50% is considered out). The radius of the water probe was left at the default value of 1.4Å.

Before analyzing changes in SASA of residues from the MD simulations, the GetArea method of calculating SASA was validated using the crystal structures of the "closed" (PDB ID:3BLG), "intermediate" (PDB ID:1BSY) and "open" (PDB ID:2BLG) forms of the protein; all deposited into PDB from the same study (Qin et al., 1998). For GetArea to be a useful method of monitoring SASA changes from the dynamics simulations, it needed to correctly display the change in SASA of these three structures. The GetArea results from the crystal structures can be found included in Tables 4.1 and 4.2. Comparisons between these results and those described in the study from which the structures were derived (Qin et al., 1998) are discussed below and serve as a validation.

The GetArea results agreed substantially with what was found in the study. Glu89 was the only amino acid to display a pH induced, buried to solvent exposed transition. A deprotonated Glu89 is unable to form a hydrogen bond with Ser116 to stabilize itself in the "closed" conformation. As a result, the protein is forced to move to a more stable conformation. Pulling away of the EF loop from the opening of the hydrophobic calyx is a result of the protein's attempt to attain a more stable conformation for the perturbation

51

Table 4.1: SASA Results From the MD Simulation of the Deprotonated Glu89 Model.

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg	Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg
Leu1																			Lys83																		
Ile2																			Ile84																		
Val3																			Asp85																		
Thr4																			Ala86																		
Gln5																			Leu87																		
Thr6																			Asn88																		
Met7																			Glu89																		
Lys8																			Asn90																		
Gly9																			Lys91																		
Leu10																			Val92																		
Asp11																			Leu93																		
Ile12																			Val94																		
Gln13																			Leu95																		
Lys14																			Asp96																		
Val15																			Thr97																		
Ala16																			Asp98																		
Gly17																			Tyr99																		
Thr18																			Lys100																		
Trp19																			Lys101																		
Tyr20																			Tyr102																		
Ser21																			Leu103																		
Leu22																			Leu104																		
Ala23																			Phe105																		
Met24																			Cys106																		
Ala25																			Met107																		
Ala26																			Glu108																		
Ser27																			Asn109																		
Asp28																			Ser110																		

52

Table 4.1 Continued

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg
Ile29																		
Ser30																		
Leu31																		
Leu32																		
Asp33																		
Ala34																		
Gln35																		
Ser36																		
Ala37																		
Pro38																		
Leu39																		
Arg40																		
Val41																		
Tyr42																		
Val43																		
Glu44																		
Glu45																		
Leu46																		
Lys47																		
Pro48																		
Thr49																		
Pro50																		
Glu51																		
Gly52																		
Asp53																		
Leu54																		
Glu55																		

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg
Ala111																		
Glu112																		
Pro113																		
Glu114																		
Gln115																		
Ser116																		
Leu117																		
Val118																		
Cys119																		
Gln120																		
Cys121																		
Leu122																		
Val123																		
Arg124																		
Thr125																		
Pro126																		
Glu127																		
Val128																		
Asp129																		
Asp130																		
Glu131																		
Ala132																		
Leu133																		
Glu134																		
Lys135																		
Phe136																		
Asp137																		

Table 4.1 Continued

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg
Ile56																		
Leu57																		
Leu58																		
Gln59																		
Lys60																		
Trp61																		
Glu62																		
Asn63																		
Asp64																		
Glu65																		
Cys66																		
Ala67																		
Gln68																		
Lys69																		
Lys70																		
Ile71																		
Ile72																		
Ala73																		
Glu74																		
Lys75																		
Thr76																		
Lys77																		
Ile78																		
Pro79																		
Ala80																		
Val81																		
Phe82																		

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1bsy	2blg
Lys138																		
Ala139																		
Leu140																		
Lys141																		
Ala142																		
Leu143																		
Pro144																		
Met145																		
His146																		
Ile147																		
Arg148																		
Leu149																		
Ser150																		
Phe151																		
Asn152																		
Pro153																		
Thr154																		
Gln155																		
Leu156																		
Glu157																		
Glu158																		
Gln159																		
Cys160																		
His161																		
Ile162																		

The 1000 structures of each trajectory were averaged to produce one structure per trajectory file. Each trajectory covered 200 ps of sampling time, hence 15 average structures were generated from the 3 ns MD simulation of the deprotonated model. The table above displays the residue name/number in the left most column followed by the per residue SASA results of a BLG crystal structure at pH 6.2 (3BLG). Next, the per residue SASA for the average structure of each trajectory are provided (indicated by 1-15). Finally, the per residue SASA of the BLG crystal structures at pH 7.1 (1BSY) and pH 8.2 (2BLG) are included. Residues considered to be solvent exposed, buried or partial by GetArea are indicated by blue, red and no color, respectively.

Table 4.2: SASA Results from the MD Simulation of the Protonated Model.

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg	Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg
Leu1																Lys83															
Ile2																Ile84															
Val3																Asp85															
Thr4																Ala86															
Gln5																Leu87															
Thr6																Asn88															
Met7																Glu89															
Lys8																Asn90															
Gly9																Lys91															
Leu10																Val92															
Asp11																Leu93															
Ile12																Val94															
Gln13																Leu95															
Lys14																Asp96															
Val15																Thr97															
Ala16																Asp98															
Gly17																Tyr99															
Thr18																Lys100															
Trp19																Lys101															
Tyr20																Tyr102															
Ser21																Leu103															
Leu22																Leu104															
Ala23																Phe105															
Met24																Cys106															
Ala25																Met107															
Ala26																Glu108															
Ser27																Asn109															
Asp28																Ser110															

Table 4.2 Continued

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg	Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg
Ile29																Ala111															
Ser30																Glu112															
Leu31																Pro113															
Leu32																Glu114															
Asp33																Gln115															
Ala34																Ser116															
Gln35																Leu117															
Ser36																Val118															
Ala37																Cys119															
Pro38																Gln120															
Leu39																Cys121															
Arg40																Leu122															
Val41																Val123															
Tyr42																Arg124															
Val43																Thr125															
Glu44																Pro126															
Glu45																Glu127															
Leu46																Val128															
Lys47																Asp129															
Pro48																Asp130															
Thr49																Glu131															
Pro50																Ala132															
Glu51																Leu133															
Gly52																Glu134															
Asp53																Lys135															
Leu54																Phe136															
Glu55																Asp137															

57

Table 4.2 Continued

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg
Ile56															
Leu57															
Leu58															
Gln59															
Lys60															
Trp61															
Glu62															
Asn63															
Asp64															
Glu65															
Cys66															
Ala67															
Gln68															
Lys69															
Lys70															
Ile71															
Ile72															
Ala73															
Glu74															
Lys75															
Thr76															
Lys77															
Ile78															
Pro79															
Ala80															
Val81															
Phe82															

Name	3blg	1	2	3	4	5	6	7	8	9	10	11	12	1bsy	2blg
Lys138															
Ala139															
Leu140															
Lys141															
Ala142															
Leu143															
Pro144															
Met145															
His146															
Ile147															
Arg148															
Leu149															
Ser150															
Phe151															
Asn152															
Pro153															
Thr154															
Gln155															
Leu156															
Glu157															
Glu158															
Gln159															
Cys160															
His161															
Ile162															

The 1000 structures of each trajectory were averaged to produce one structure per trajectory file. Each trajectory covered 200 ps of sampling time, hence 12 average structures were generated from the 2.4 ns MD simulation of the protonated model. The table above displays the residue name/number in the left most column followed by the per residue SASA results of a BLG crystal structure at pH 6.2 (3BLG). Next, the per residue SASA for the average structure of each trajectory are provided (indicated by 1-15). Finally, the per residue SASA of the BLG crystal structures at pH 7.1 (1BSY) and pH 8.2 (2BLG) are included. Residues considered to be solvent exposed, buried or partial by GetArea are indicated by blue, red and no color, respectively.

in the chemistry of the Glu89 side chain. This causes Ser116 to experience partial exposure of its side chain. The exposure of Ser116 was also confirmed by the GetArea results. All but one amino acid comprising the surface of the hydrophobic interior of the calyx did not experience a change in solvent accessibility. Ile71, displayed an increase in solvent accessibility probably due to the movement of Glu89 away from the rim of the cavity where it is located. This was also seen from the surface mapping measurements. The side chain transition of Asp85 from being solvent exposed to being partially buried due to the movement of the loop was visualized by both Qin et al. and via the GetArea data. Similarly, Glu131, which becomes partially buried at pH 7.1 but solvent exposed again at pH 8.2 was confirmed by GetArea. Lys91, Glu65 and Glu114 which were all detected as experiencing changes in SASA of more than 30Å by Qin et al., were not confirmed as experiencing changes by GetArea most likely because the change fell within the range of solvent accessibility characterization (in, out or partial) of the GetArea algorithm. They were, however, confirmed as being exposed (Glu65 and Glu114) and partial (Lys91). This is a caveat of using two different methods for SASA calculations. Glu108 was buried in all three structures supporting Qin et al.'s claim of there being a salt bridge between it and Leu1, which was established as being solvent exposed by both. Finally, Qin et al. found the side chains of Leu87, Leu39 and Lys60 buried at pH 6.2 but partially exposed at basic pH values and this was also seen from the GetArea results. Comparisons between the GetArea results and the Qin study were promising and provided enough confidence for using GetArea as the method to study change in residue SASA from MD simulations.

The 3ns (2.4 for the protonated model) molecular dynamics simulation was done in trajectories of 200ps each for a total of 15 (12 for the protonated model) trajectory files. Each trajectory file contained 1000 structures for a total of 15000 (12000 for the protonated model) structures. InsightII software (Accelrys) was used to produce an average structure from the 1000 structures of each trajectory file. A total of 15 (12 for the protonated model) average structures were generated by this method. An average structure was used for SASA analysis because it eliminates any artifacts presented by a single structure and provides a better representative structure of the sample space. The GetArea output of the 15 (12 for the protonated model) average structures is summarized in Table 4.1 (Table 4.2 for the protonated model).

1. Overall Analysis

Overall analysis revealed no detection of any large changes in solvent accessibility for the sampling times monitored. GetArea software did not describe the movement of any amino acid from a solvent exposed to buried state or vice versa. Differences in degree of solvent accessibility were detected for several residues in both models. The following discussion provides an overall comparative analysis and most likely explanations of the results obtained followed by a detailed investigation of regions implicated in confirmed or suggested conformational change.

Changes in the solvent accessibility of residues were visible upon comparison of SASA results from both simulations. They were detectable throughout the protein, however, the specific residues and their direction of change were unique to each simulation. Keeping in mind that proteins are in constant dynamic motion during

61

simulations, observations from data obtained from the simulations were made on a broad scale.

Some residues did not experience any change in their solvent accessibility state. Amino acids 21-27 started and remained buried throughout the MD simulations of both models. Conversely, residues 63-65 remained solvent exposed throughout both MD simulations. These results agree with the GetArea calculations for these residues from the crystal structures and suggest that a change in the protonation state of Glu89 does not instigate any significant change in solvent accessibility in their region of the protein structure. More undisturbed residues were seen from the simulation of the protonated model than for the deprotonated model. This is expected as movement was observed from the simulation of the deprotonated model but not from the protonated model.

Some residues experience change when they shouldn't. Though Thr49 and Tyr99 are characterized as partial and buried by the GetArea output for the crystal structures at pre- and post-Tanford Transition pH values, they appear buried and partial, respectively, throughout the MD simulations of both models. This phenomenon could be the effect of minimization of the model before performing MD simulations or a limitation of the GetArea algorithm in its exposure classification.

Some residues experience fluctuations in their solvent accessibility during the simulation sample time. Though both models do not display the same residues undergoing fluctuations, the percentage of residues that experience fluctuations are similar. Because of this, fluctuations cannot be attributed to the movement of the EF loop but are probably the result of the constant dynamic movement of the protein.

Some residues experience increased exposure. For example Ile71 and Ser116 display a transition from buried to partial in MD simulations of both models. Ile71 is located on the rim of the hydrophobic cavity. As the EF loop pulls away from the opening of the hydrophobic cavity, this residue becomes unprotected from solvent. The simulation of the deprotonated model showed solvent accessibility change (buried to partial) for this residue but exposure of this residue in the protonated model remained the same (partial). This residue was characterized as "buried" in the crystal structure at pH 6.2 but was shown as "partial" throughout the MD simulation. This again could be the result of minimization of the protonated model before performing MD simulations. Though it was not characterized as buried at its starting state, it is important to note that it did not display any change in solvent accessibility. Ile71 was also implicated in increased solvent exposure from the surface mapping measurements. Increased exposure was demonstrated as a result of the change in conformation of the protein between pH 5 and 6 when the EF loop initiates its movement. Ser116, which hydrogen bonds with Glu89 when it is protonated, experiences an increase in solvent accessibility during both simulations. This increase occurs much faster in the simulation of the deprotonated than the protonated model. This residue forms a hydrogen bond with Glu89 when it is protonated. This hydrogen bond appears to prolong significant movement of the EF loop, however the dynamic movement of the protein eventually results in the cleavage of the hydrogen bond and exposure of the serine side chain.

An important observation from both MD simulations was that Glu89 was categorized by GetArea calculations as buried throughout the sampling time monitored. This result is expected during the simulation of the protonated Glu89 model; however,

simulation of the deprotonated model should have exhibited a shift from buried to completely solvent exposed for this residue. Movement of the EF loop was observed during this MD simulation; however, it must have not been large enough for the GetArea solvent accessibility categorization to change. Therefore, though movement was observed during the MD simulation of the deprotonated model, it was not substantial enough to display change in solvent accessibility by GetArea calculations.

2. EF Loop Movement

The solvent accessibility change of Glu89, which should have been displayed during the simulation of the deprotonated model, was not present. In fact, it was classified as "buried" for the entire 3 ns sampling time. The fact that the pre- and post-Tanford Transition crystals display this change from GetArea calculations suggests that factors in addition to the protonation state of Glu89 are involved in attaining the "complete" Tanford Transition. Though Glu89 appeared to undergo a significant shift in exposure during the MD simulation, its SASA values were still within the "buried" classification in GetArea.

Other amino acid side chains that implicate movement of the EF loop are Ser116 and Ile71 which both experienced an increase in exposure and were previously discussed.

3. α-Helix Movement

The α-helix was not implicated in movement upon visual inspection of both MD simulations. The GetArea results from both simulations for residues in this region also

did not display any significant change in SASA. If the α-helix does undergo a pH induced change, the protonation state of Glu89 does not appear to contribute.

B. Free Energy Calculations

The free energy of a protein system is lowest when the protein is in its most stable conformation. Free energy calculations were performed to demonstrate the most stable, or preferred, distance between C-alpha atoms on the EF loop (85-90) and that of Ile71 which is located on strand D, near the rim of the hydrophobic cavity as well as between C-alpha atoms on the α-helix (130-140) and that of Gln120, which is located on the H strand.

Ile71 was chosen as an anchor point for the loop distance calculations because it is located on the rim of the hydrophobic cavity which becomes exposed as the loop pulls away during the Tanford Transition. Additionally, its localization to a β-strand, which is more rigid than a loop, prevents its contribution in distance measurements. Gln120 was chosen as the anchor point for the α-helix distance calculations because it is located on a strand that appears to be more rigid than others during dynamics (Kuwata et al., 1999). More sensitivity of distance measurements in this region were desired because of the suggested movement of the α-helix (Monaco et al., 1987).

As in the SASA analysis, the investigation of conformational change using this method was focused on regions of known or suspected movement, namely the EF loop and α-helix region, respectively. The average distance between the C-alpha atoms of the six amino acids that make up the EF loop (85-90) and the C-alpha atom of Ile71 as well as the average distance between the C-alpha atoms of the 11 amino acids that make up

the α-helix (130-140) and the C-alpha atom of Glu120 were monitored as a function of time (individual data not shown). Figure 4.2 displays these results for the EF loop and α-helix regions from the molecular dynamics simulation of both the de- and protonated models.

Calculation of free energy was based on the probability density of the measurements. The following equation was used to determine the free energy from the dynamics simulation:

$$\omega = -k_B T \ln\rho$$

where k_B is the Boltzmann constant, T is the temperature in Kelvin and ρ is the probability density. This equation is useful in measuring free energy when N (number of particles in the system), V (volume of the system) and T (temperature of the system) are constant. Figure 4.3 displays the free energy distance calculations for the EF loop and α-helix regions from the molecular dynamics simulation of both the deprotonated and protonated models.

The 3 ns and 2.4 ns simulation times of the deprotonated and protonated models; respectively, are not a measure of the change in conformation of the protein model in real time but during sampling time. The deprotonated model sampled a conformation of its model protein every 0.2 ps resulting in a total of 15,000 structures sampled during the 3 ns of sampling time. The protonated model also sampled a conformation of its model protein every 0.2 ps but because its sampling time was only for a total of 2.4 ns, a total of 12,000 structures were sampled. The total structures sampled become the population from which the probability density is calculated during free energy measurements.

A.
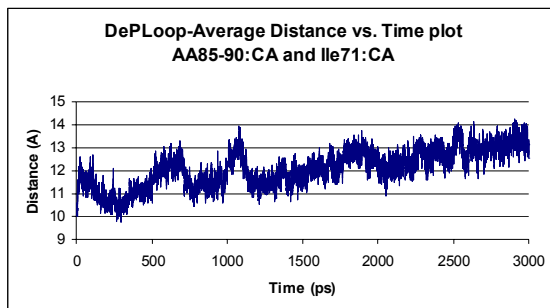
B.
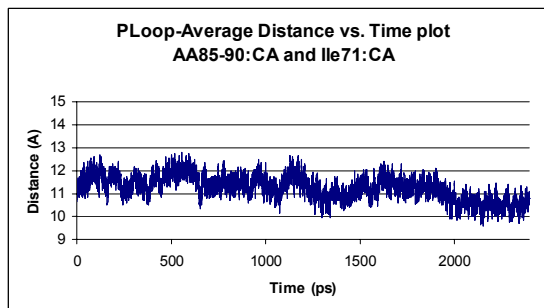
C.
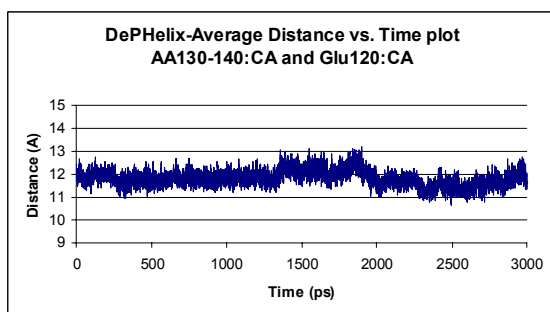
D.
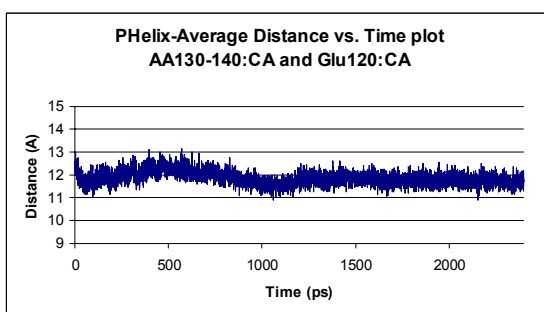
Figure 4.2: Average Distance Plots.

The average distance plotted as a function of time. The top row (**A** and **B**) are of the loop region and the bottom row (**C** and **D**) are of the helix region. The left column (**A** and **C**) are from the dynamics simulation of the deprotonated model and the right column (**B** and **D**) are from the dynamics simulation of the protonated model.

A.

DePLoop-Free energy vs. Average Distance profile

B.

PLoop-Free energy vs. Average Distance plot

C.

DePHelix-Free energy vs. Average Distance profile

D.

PHelix-Free energy vs. Average Distance profile
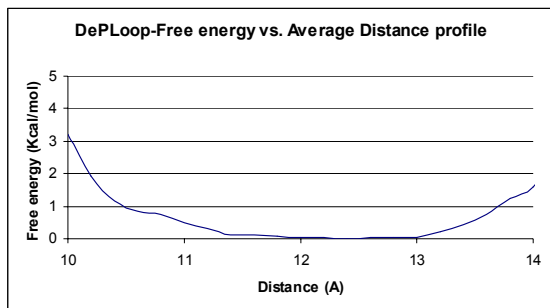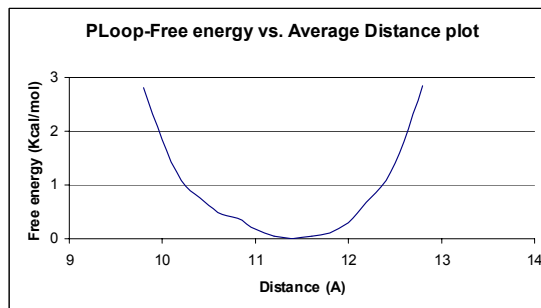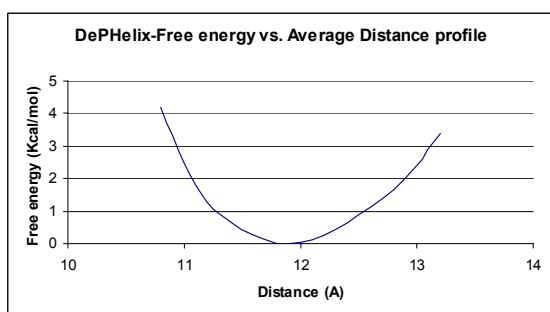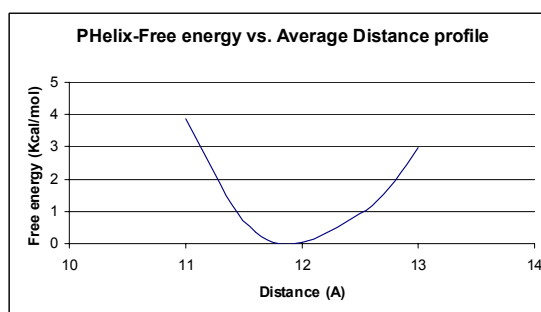
Figure 4.3: Free Energy Profiles.

The free energy plotted as a function of average distance. The top row (**A** and **B**) are of the loop region and the bottom row (**C** and **D**) are of the helix region. The left column (**A** and **C**) are from the dynamics simulation of the deprotonated model and the right column (**B** and **D**) are from the dynamics simulation of the protonated model. simulation.

1. Deprotonated Glu89 Model

The loop-average distance plot (Figure 4.2A) displays an increase during the 3ns dynamics simulation. The average distance begins around 10Å and increases to about 13Å by the end of the 3ns sampling time. The increase in distance does not appear to occur as a smooth trend. It displays a couple rapid fluctuations before 1ns after which it seems to come to equilibrium after 2.5ns. Random fluctuations are a caveat of simulations. In MD simulations, proteins are subjected to constant dynamic motion. As a result, meaningful data should be extracted from overall observations rather than localized regions of sampling time. The most meaningful information to take away from this plot was the quantitative increase in distance that supports what was visualized during the simulation of this model.

The helix-average distance plot (Figure 4.2C) does not display any significant movement of the helix during the 3ns dynamics simulation. The average distance begins around 11.5Å and stays pretty much at this level for the majority of the sampling time. A sudden increase in distance is seen around 1.5 ns; however, this increase is only about 1 Å and eventually returns to 11.5 Å range values so it is not considered significant. This change could be a reflection of the protein attempting to get over an energy barrier.

The loop-free energy profile (Figure 4.3A) shows that the protein is in its most stable conformation when the average distance between the C-alpha atoms of residues 85-90 and the C-alpha atom of residue Ile71 is about 12.5Å. The helix-free energy profile (Figure 4.3C) shows that the protein is in its most stable conformation when the average distance between the C-alpha atoms of residues 130-140 and the C-alpha atom of residue Gln120 is about 11.8Å.

69

2. Protonated Glu89 Model

The loop-average distance plot (Figure 4.2B) does not display any broad scale change in average distance during the 2.4ns dynamics simulation. This is confirmatory of what was visualized from the dynamics simulation of this model and expected when Glu89 is protonated. The EF loop region of this model does not show any signs of movement away from the opening of the hydrophobic cavity.

The helix-average distance plot (Figure 4.2D), similar to the one of the deprotonated model, does not display any significant movement of the helix during the 2.4ns dynamics simulation. The average distance begins and stays close to 11.5Å for the majority of the sampling time suggesting no movement of the α-helix occurs under these conditions.

The loop-free energy profile (Figure 4.3B) shows that the protein is in its most stable conformation when the average distance between the C-alpha atoms of residues 85-90 and the C-alpha atom of residue Ile71 is about 11.5Å. This stable distance is less than that for the deprotonated structure, supporting the fact that the protonation state of Glu89 effects the conformation of the EF loop. The difference between the stable distances for the deprotonated and protonated structures is not very large. This could be an artifact of the averaging of distances but also suggests that factors other than the protonation state of Glu89 are important for the complete Tanford Transition.

The helix-free energy profile (Figure 4.3D) shows the same results as for the deprotonated model. The protein is in its most stable conformation when the average distance between the C-alpha atoms of residues 130-140 and the C-alpha atom of residue

Glu120 is about 11.8Å. This indicates that if movement of the α-helix region does occur, it is not induced by the change in the protonation state Glu89 residue.

The sampling of distance space is a lot less for the α-helix than for the EF loop in both models. This implies that the EF loop moves more than the α-helix; which is confirmed by both visual examination of the MD simulations and comparisons between the loop and helix average distance plots. Comparing the loop-free energy profiles of both models, the sampling distance space is far greater for the deprotonated than the protonated model. Equilibrium simulations exhibit less sampling of conformational space than a simulation that is not at equilibrium (Fogolari et al., 2005). The deprotonated model is not at equilibrium as the closed β-lactoglobulin structure prefers to have its Glu89 residue protonated. When a protein is not at equilibrium, it undergoes successive changes in its conformation until it adopts a structure that is most stable at its given condition. The deprotonated Glu89 model undergoes these successive changes in sampling average distance values from less than 10 to greater than 14Å until it finds its most stable conformation. The larger sampling distance space displayed by the deprotonated model than the protonated model could be a result of this phenomenon or because the deprotonated model in general samples a larger space than the protonated model. Further investigation would be required to explain these findings.

II. Conclusions

MD simulations of the deprotonated and protonated models implicated the de-protonation of Glu89 as a trigger for the Tanford Transition and validated the models. A transition of the protein was obvious from visual inspection of the deprotonated model

71

simulation. Solvent accessibility measurements of the simulations provided an additional level of detail; identifying specific residues associated with movement and confirming previously reported results (Eberini et al., 2004). However, a complete Tanford Transition was not confirmed due to disagreement between SASA calculations of critical residues from the simulation and the crystal structures at pH 7.1 and 8.2. This suggested other factors are necessary for the complete "closed" to "open" conversion. Additionally, the methodology employed to measure SASA was probably not at an adequate level of detail as the one obvious accessibility difference (Glu89) was not displayed by the data. The free energy profiles confirmed the affinity for the deprotonated model to adopt a more stable conformation of the EF loop at a greater distance away from the opening of the cavity than the protonated model, confirming de-protonation of Glu89 as the trigger for the Tanford Transition, but also suggested that other factors are necessary in achieving the target "open" conformation as the difference between the most stable average distance between the protonated and deprotonated models was not very large. No significant movement of the α-helix was observed from either techniques. Overall, the results obtained from the two techniques agree with the expected behavior of the models built.

Chapter 5: An Attempt at Integrating Oxidative Surface Mapping and Molecular Dynamics Simulation Results and Conclusions Based on Overall Findings

This thesis work was designed as an attempt to explore how to integrate two dramatically different disciplines (experimental and computational) that were used to examine protein conformational change. Oxidative surface mapping was employed as the experimental component and molecular dynamics simulations as the computational. The effectiveness of correlating these two techniques was evaluated by enlisting the study of a well known protein, β-lactoglobulin A, which undergoes a pH dependent conformational transition.

Results from the techniques described in Chapter 2 individually agree with what has been published about the model protein system; however, the ability to integrate these two disciplines in order to obtain meaningful results and/or additional details about the system has not been demonstrated prior to this work. The following discussion provides a brief review of the information provided by each technique on the study of bovine β-lactoglobulin A, after which the ability, extent and effectiveness in integrating the two techniques is examined.

I. Summary of Results

Oxidative surface mapping provided results as a function of the SASA of individual residues. Changes in the structure of the protein were mapped by inducing the β-lactoglobulin protein to conform to its structure in various pH environments and then

chemically labeling solvent-exposed and reactive amino acid side chains. This gave insight about conformational changes the protein undergoes; however, due to the limited number and different reactivity rates of the most commonly oxidized amino acid side chains, this method (as described in the materials and methods chapter) for measuring conformational change of proteins provided limited structural resolution. Further optimization of the methodology could prove otherwise. Enlisting the help of less selective labeling methods would provide better sequence coverage in stretches of amino acids that are relatively unreactive with hydroxyl radical. This would provide a higher degree of structural resolution and a deeper understanding of which residues are implicated in movement.

The Tanford Transition was characterized by oxidation events at a number of amino acid side chains, around the EF loop, that were not exposed in the acidic environments but had dramatically altered solvent accessibility at pH 8. Previous studies report the movement of the EF loop as the only significant difference in conformation as a function of pH however, oxidation events at amino acid side chains in the cleft between the α-helix and body of the protein initiated at pH 5, proposing the presence of an additional region implicated in change due to pH.

Molecular dynamics simulation results presented information about the SASA and free energy of the system. The per-residue SASA of an average structure of each trajectory was determined and compared between trajectories of sampling time. Changes in SASA for residues in regions of proposed mobility served as evidence for change in conformation of the protein. Several residues expected to undergo change in their SASA were displayed as doing so using this SASA measurement method. Surprisingly, the sole

74

amino acid that should display a shift from completely buried to completely solvent exposed (Glu89) was not shown as doing so by the SASA calculation results.

Free energy calculations were performed on average distance measurements at both the loop and α-helix regions. The protonated model served as a negative control since the protein does not undergo the Tanford Transition when Glu89 is in its protonated state. Therefore, MD simulation results from the deprotonated model are the most relevant for discussion. The deprotonated model experienced the Tanford Transition and demonstrated stability of its structure at a loop to D-strand average distance greater than that for the stable structure of the protonated model. This difference was not very large suggesting that other factors are necessary in achieving the target "open" conformation. No significant movement of the α-helix region however, was observed from free energy analysis.

II. Overall Observations

Both techniques demonstrated movement of the EF loop away from the opening of the hydrophobic calyx, however, the extent of movement of the loop appears to differ between techniques. Each technique provides complementary but not necessarily overlapping information. Therefore, seamless integration of the results from both methods can be performed to only a limited degree. The absence of a significant degree of overlap in the results is not necessarily a negative outcome. It allows for a deeper understanding of the system, elucidates limitations of each method and suggests improvements that can be made to better approach a similar problem in the future.

III. Novel Information

The inability of the MD simulation of the deprotonated Glu89 model to display the "complete" Tanford Transition implied the presence of other factors involved in the conformational change. Though the additional factors that contribute are not immediately apparent, the surface mapping measurements provide hints. Movement of the α-helix was observed from the oxidative surface mapping data. This suggests that a change in pH of the environment affects residue(s) in this region instigating a pulling away of the α-helix from the body of the protein. Additionally, movement of the EF loop to a degree larger than that observed from the MD simulation is observed from the surface mapping measurements implying that other titratable residues in the region of the EF loop may also undergo changes in their chemistry as the pH becomes more basic. This scenarios is very likely as a shift in the pH of the environment of the β-lactoglobulin protein would implicate other residues to change charge. The models that were built were kept simple, therefore their simulations resulted in the inability to visualize the "complete" Tanford Transition.

IV. Limitations

Oxidative surface mapping is blind to residues that do not readily react with hydroxyl radicals. A simple solution to this situation could be to increase the concentration of hydroxyl radicals in solution, thereby increasing the probability of the less reactive amino acids to become oxidized. Increased hydroxyl radical concentration would not only increase the probability of relatively unreactive amino acids of being oxidized, but could also pose the threat of overly oxidizing proteins, thereby possibly

distorting the native folded structure. This would result in extensive oxidation of the protein such that the distinction between a solvent exposed and buried residue would become obsolete. Additionally, increased hydroxyl radical concentration could implicate the cleavage of the backbone of the protein. Cleavage of the protein into peptides by hydroxyl radical attack would result in the inability to accurately probe the structure of the protein as regions that should be buried in the intact protein would now become exposed due to cleavage of the backbone. Because of these possibilities, oxidation methods are optimized to keep a delicate balance. Though this method of oxidation is rapid and does not require special equipment or reagents, its potential lack of sensitivity in the study of proteins that undergo subtle but significant conformational shifts around amino acid sequences where reactivity is limited could render it limited in scope and ineffective in studying certain proteins. Conversely, continued optimization of this method of surface mapping could improve its efficiency and increase its scope of protein systems. Specifically, less selective labeling methods would provide deeper insight about movement in regions that are not labeled by oxidative surface mapping.

Although the field of molecular dynamics simulations has progressed over the last ten years, this approach still has some limitations for probing the complete conformational alterations of proteins. Simulations are performed based on parameters defined by the user. Thus, limitations are not as much the fault of the algorithm as the user who builds the model, sets up the parameters, and runs the simulations. Nonetheless, MD simulations are a powerful tool to visualize the effects of specific perturbations to a protein system. The simulation of the deprotonated Glu89 β-lactoglobulin protein displayed the pH induced movement of the EF loop. Upon closer inspection of the

simulation by way of monitoring SASA changes of individual residues and measuring the free energy change, the movement of the EF loop, though significant upon visual inspection, stabilizes only about 1 Å away from the stable position of the EF loop for the protonated Glu89 model. As per GetArea SASA calculations of the crystal structures, Glu89 is the only residue to undergo complete reversal of its buried state as a result of the Tanford Transition. Movement of the EF loop away from the opening of the hydrophobic core of the protein is not significant enough to display change in the SASA data of this residue, in fact this residue is characterized as buried throughout the MD simulation. Molecular dynamics simulations lack the ability to behave heuristically. Simulations are performed based on the parameters initially given. The simulation of the deprotonated model of β-lactoglobulin only displays the initial trigger not the "complete" Tanford Transition as described by the crystal structures. This suggests that the presence of other factors is necessary and displays the limitations of the method but identifies that they are the direct result of initial parameters. Also, MD simulations are limited to the structural space they can survey during the sampling time allowed. Longer sampling times potentially allow proteins to conform to structures of local energy minimums and display a larger structural space. The "complete" Tanford Transition could be displayed simply by monitoring longer sampling times of the deprotonated model.

V. Future Strategy

The presence of other factors involved in the Tanford Transition is apparent upon comparison between the results obtained from oxidative surface mapping and molecular dynamic simulation techniques. In addition to Glu89, the protonation state of additional

residues in the vicinity of the EF loop and α-helix can be perturbed to probe which additional residues undergo change in their chemistry and contribute to the conformational shift of the protein. Simulations performed on these models could confirm their involvement if an increased degree of agreement between surface mapping measurements is observed. Additionally, free energy simulations based on potential of mean force and umbrella sampling can be employed in future analysis to provide results of regions that cannot be sampled by regular molecular dynamics.

The surface mapping technique can be further improved to cover less reactive amino acid residues and to obtain better sequence coverage. Protein samples under denatured conditions can be used to monitor effects of optimization.

VI. Concluding Remarks

Although our initial attempt to integrate these specific experimental and computational techniques revealed a limited degree of overlap, it provided a better understanding of the conformational change of the β-lactoglobulin protein and suggested alternative approaches and refinements for future attempts of this strategy. Implementation of a revised strategy could provide further information about the protein system and continued heuristic improvements to the strategy for studying protein conformational change.

Literature Cited

Adcock SA, McCammon JA. (2006) Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. ***Chem. Rev.*** **106:** 1589-1615.

Andrade MA, Brown NP, Leroy C, Hoersch S, de Daruvar A, Reich C, Franchini A, Tamames J, Valencia A, Ouzounis C, Sander C. (1999) Automated genome sequence analysis and annotation. ***Bioinformatics*** **5:** 391-412.

Aqvist J, Sandblom P, Jones TA, Newcomer ME, van Gunsteren WF, Tapia O. (1986) Molecular dynamics simulations of the Holo and Apo forms of Retinol Binding Protein: Structural and dynamical changes induced by retinol removal. ***J. Mol. Biol.*** **192:** 593-604.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. ***J. Comp. Chem.*** **4:** 187-217.

Brownlow S, Cabral JHM, Cooper R, Flower DR, Yewdell SJ, Polikarpov I, North ACT, Sawyer L. (1997) Bovine β-lactoglobulin at 1.8Å resolution-still an enigmatic lipocalin. ***Structure*** **5:** 481-495.

Buxton GV, Greenstock CL, Helman WP, Ross AB. (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:** 513-886.

Christodoulou J, Larsson G, Fucini P, Connell SR, Pertinhez TA, Hanson CL, Redfield C, Nierhaus KH, Robinson CV, Schleucher J, Dobson CM. (2004) Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes. ***Proc. Natl. Acad. Sci.*** **101:** 10949-10954.

Domingues FS, Koppensteiner WA, Sippl MJ. (2000) The role of protein structure in genomics. ***FEBS Letters*** **476:** 98-102.

Dufour E, Robert P, Bertrand D, Haertle T. (1994) Conformation changes of β-lactoglobulin: An ATR infrared spectroscopic study of the effect of pH and ethanol. ***J. Prot. Chem.*** **13:** 143-149.

Eberini I, Baptista AM, Gianazza E, Fraternali F, Beringhelli T. (2004) Reorganization in Apo- and Holo- -Lactoglobulin Upon Protonation of Glu89: Molecular Dynamics and pKa Calculations. ***Prot.: Struct. Func. Bioinf.*** **54:** 744-758.

Fogolari F, Moroni E, Wojciechowski M, Baginski M, Ragona L, Molinari H. (2005) MM/PBSA Analysis of Molecular Dynamics Simulations of Bovine -Lactoglobulin: Free Energy Gradients in Conformational Transitions? Proteins: ***Struct. Func. Bioinf.*** **59:** 91-103.

Fraczkiewicz R, Braun W. (1998) Exact and Efficient Analytical Calculations of the Accessible Surface Areas and Their Gradients for Macromolecules. *J. Comp. Chem.* **19:** 319-333.

Garrison WM. (1987) Reaction Mechanisms in the radiolysis of peptide, polypeptides, and proteins. *Chem. Rev.* **87:** 381-398.

Gu W, Brady JW. (1992) Molecular dynamics simulations of the whey protein β-lactoglobulin. *Prot. Enginer.* **5:** 17-27.

Hakansson K, Doherty AJ, Shuman S, Wigley DB. (1997) X-Ray crystallography reveals a large conformational change during Guanyl transfer by mRNA capping enzymes. *Cell* **89:** 545-553.

Hawkins CL, Davies MJ. (2001) Generation and propagation of radical reactions on proteins. *Biochim. Biophy. ACTA* **1504:** 196-219.

Hubner G, Roduner E. (1998) EPR investigation of OH" radical initiated degradation reactions of sulfonated aromatics as model compounds for fuel cell proton conducting membranes. *J. Mater. Chem.* **9:** 409-418.

Jorgensen WL, Chandrasekhar J, Madura JD. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79:** 926-935.

Karplus M, McCammon JA. (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **9:** 646-652.

Karplus M, Kuriyan J. (2005) Molecular dynamics and protein function. *Proc. Natl. Acad. Sci.* **102:** 6679-6685.

Kiselar JG, Maleknia SD, Sullivan M, Downard KM, Chance MR. (2002) Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. *Int. J. Radiat. Biol.* **78:** 101-114.

Koppensteiner WA, Lackner P, Wiederstein M, Sippl MJ. (2001) Characterization of novel proteins based on known protein structures. *J. Mol. Biol.* **296:** 1139-1152.

Kuwata K, Hoshino M, Forge V, Era S, Batt CA, Goto Y. (1999) Solution structure and dynamics of bovine β-lactoglobulin A. *Prot. Sci.* **8:** 2541-2545.

MacKerell Jr. AD, Bashford D, Bellott M, Dunbrack Jr. RL, , Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher III WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin

D, and Karplus M (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. Bio.* **102:** 3586-3616.

Maleknia SD, Brenowitz M, Chance MR. (1999) Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. *Anal. Chem.* **71:** 3965-3973.

Maleknia SD, Ralston CY, Brenowitz MD, Downard KM, Chance MR. (2001A) Determination of macromolecular folding and structure by synchrotron X-Ray radiolysis techniques. *Anal. Biochem.* **289:** 103-115.

Maleknia SD, Downard KM. (2001B) Unfolding of apomyoglobin helices by synchrotron radiolysis and mass spectrometry. *Eur. J. Biochem.* **268**: 5578-5588.

Maleknia SD, Kiselar JG, Downard KM. (2002) Hydroxyl radical probe of the surface of lysozyme by synchrotron radiolysis and mass spectrometry. *Rapid Commun. Mass Spectrom.* **16:** 53-61.

Monaco HL, Zanotti G, Spadon P. (1987) Crystal Structure of the Trigonal Form of Bovine Beta-lactoglobulin and of its Complex with Retinol at 2.5Å Resolution. *J. Mol. Biol.* **197:** 695-706.

Paget MSB, Buttner MJ. (2003) Thiol-based Regulatory Switches. *Annu. Rev. Genet.* **37:** 91-121.

Palmer AH. (1934) The preparation of a crystalline globulin from the albumin fraction of cow's milk. *J. Biol. Chem.* **104:** 359-372.

Papiz MZ, Sawyer L, Eliopoulos EE, North ACT, Findlay JBC, Sivaprasadarao R, Jones TA, Newcomer ME, Kraulis PJ. (1986) The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature* **324:** 383-385.

Polygen Corporation. (1988) CHARMM User's Guide.

Qin BY, Bewley MC, Creamer LK, Baker HM, Baker EN, Jameson GB. (1998) Structural basis of the Tranford Transition of bovine β-lactoglobulin**. *Biochem.* **37:** 14014-14023.

Qin BY, Bewley MC, Creamer LK, Baker EN, Jameson GB. (1999) Functional implications of structural differences between variants A and B of bovine beta-lactoglobulin. *Prot. Sci.* **1:** 75-83.

Sakurai K, Goto Y. (2005) Dynamics and mechanism of the Tanford Transition of bovine β-lactoglobulin studied using heteronuclear NMR spectroscopy. *J. Mol. Biol.* **356:** 483-496.

Sharp JS, Becker JM, Hettich RL. (2004) Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Anal. Chem.* **76:** 672-683.

Sharp JS, Becker JM, Hettich RL. (2003) Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry. *Anal. Biochem.* **313:** 216-225.

Sharp JS, Guo J, Uchiki T, Xu Y, Dealwis C, Hettich RL. (2005) Photochemical surface mapping of C14S-Sml1p for constrained computational modeling of protein structure. *Anal. Biochem.* **340:** 201-212.

Stadtman ER, Levine RL. (2003) Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. *Amino Acids* **25:** 207-218.

Tabb DL, McDonald WH, Yates JR III. (2001) DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Prote. Res.* **1:** 21-26.

Tabb DL, Narasimhan C, Strader MB, Hettich RL. (2005) DBDigger: Reorganized Proteomic Database Identification That Improves Flexibility and Speed. *Anal. Chem.* **77:** 2464-2474.

Tanford C, Bunville LG, Nozaki Y. (1959) The reversible transformation of β-lactoglobulin at pH 7.5. *J. Am. Chem. Soc.* **81:** 4032-4035.

Taulier N, Chalikian TV. (2001) Characterization of pH-induced transitions of β-lactoglobulin: Ultrasonic, densimetric, and spectroscopic studies. *J. Mol. Biol.* **314:** 873-889.

Uchiki T, Dice LT, Hettich RL, Dealwis C. (2004) Identification of phosphorylation sites on the Yeast Ribonucleotide Reductase Inhibitor Sml1. *J. Biol. Chem.* **279:** 11293 - 11303.

Uhrinova S, Smith MH, Jameson GB, Uhrin D, Sawyer L, Barlow PN. (2000) Structural changes accompanying pH-induced dissociation of the β-lactoglobulin dimer. *Biochem.* **39:** 3565-3574.

Vogt W. (1995) Oxidation of Methionyl Residues in Proteins: Tools Targets and Reversal. *Free Radic Biol Med.* **18**: 93-105.

Vita

Kanan R. Vyas was born on May 8, 1977 in Hayward, California to parents Ramesh N. Vyas and Niranjana R. Vyas. Three years later, she became the big sister of Ankit R. Vyas. She was raised, for the most part, in Fremont, California where she attended Hopkins Junior High School and Mission San Jose High School before moving to Davis, California to attend The University of California at Davis. She graduated from UC Davis in 2000 with a Bachelor of Science degree in Microbiology.

After attaining a Certificate in Bioinformatics from the University of California at Santa Cruz and several years of work experience in both industry (Affymetrix Inc. and Perlegen Science Inc.) and academia (Stanford University School of Medicine), she joined The University of Tennessee/Oak Ridge National Laboratory Graduate School of Genome Science and Technology. In 2006 she graduated with a Master of Science degree in Life Science with a concentration in Genome Science and Technology.

Her plans after graduation are to marry her long time fiancé, Amit. N. Shah, in September 2006.