Title

Carbon-deuterium bonds as an infrared probe of protein dynamics, local electrostatics and folding

Permalink

Author

Sagle, Laura B.

Publication Date

2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Carbon-Deuterium Bonds as an Infrared Probe of Protein Dynamics, Local

Electrostatics and Folding

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Chemistry

by

Laura B. Sagle

Committee in charge:

        Professor Andrew McCammon, Chair
        Professor Floyd Romesberg
        Professor Patricia Jennings
        Professor Elizabeth Komives
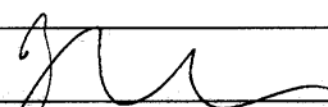   Professor Douglas Magde
        Professor Jose Onuchic

2006

The dissertation of Laura B. Sagle is approved, and it is

acceptable in quality and form for publication on microfilm:

Patricia A Jennings

Elizabeth A. Komives

Douglas Magde

Jack McCammon

Chair

University of California, San Diego

2006

iii

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Thanks to the Baldridge group, who greatly aided in many difficult computational studies. Professor Baldridge, who is wonderful to work with and has also been a great source of support.

Thanks to the Okamura/Feher group, who I've learned so much from and have always treated me with kindness.

Thanks to Professors McCammon and Komives, who not only allowed me to change groups, but also stood behind me every step of the way. Through every extension, every lapse in salary and every time I needed advise or just to talk. Thank you both so much, you certainly went above and beyond what any graduate student could expect.

Thanks to the co-authors on the papers whose work appears as chapters 3 and 4 in this thesis: Jörg Zimmermann, Phillip Dawson, and Floyd Romesberg. The Journal of the American Chemical Society is also acknowledged for granting permission to include the published results in this thesis.

Laura B. Sagle

Ph.D. Chemistry   Winter 2006
University of California, San Diego
Advisors: Floyd Romesberg and Andrew McCammon

M.S. Chemistry   June 2000
University of California, San Diego
Advisors: Melvin Okamura and George Feher

B.S., Chemistry   June 1996
University of California, Santa Cruz
Chemistry, Biochemisty Track
Advisor: David Kliger

Research Experience
The Scripps Research Institute, San Diego, Professor Floyd Romesberg, Advisor
3/2001 – present

Incorporated carbon-deuterium bonds into cytochrome c through its semi-synthesis and used these bonds as IR probes. Equilibrium folding experiments as well as oxidation induced changes to the protein were measured in a site-specific manner. Measurements involved optimization of the FTIR and careful baseline correction. The analysis of the results required deuterium frequency calculations (using Gaussian98), fitting the data to models using Gaussians, and the calculation of thermodynamic parameters.

University of California, San Diego, Professor Melvin Okamura, Advisor
3/1999 – 3/2001

Investigated proton transfer pathways to the Qb site in the bacterial reaction center protein through effects induced by exogenous metal ions. Samples were natively expressed and purified using biochemical methods. Measurement of binding constants involved laser excitation and monitoring of electron transfer rates. Trends of Kd values vs pH were analyzed.

POINT Biomedical Corporation, Dr. Tom Ottoboni, Research Director
10/1996 – 6/1998

Developed assays to measure fat and sugar content in purified collagen for an ultrasound contrasting agent. Worked on the development of a urological drug delivery device consisting of an emulsion and water soluble gel. Tested different emulsions for stability and ability to hold the free-base drug as well as different water soluble gels for brittleness and viscosity.

University of California, Santa Cruz, Professor David Kliger, Advsor
9/1994 – 6/1996

> Tested for excitonic coupling/energy transfer between the heme chromophores of hemoglobin a. Sample preparation was carried out in an oxygen-free environment and CD and OD spectra of different concentrations of hemoglobin as well as different percentages of CO bound were compared and modeled using Matlab. A thesis paper was written on the data that was collected and analyzed.

Professional Affiliations and Awards
American Chemical Society, Student Member (1997-present)
Biophysical Society, Student Member (1999-present)
Protein Society, Student Member (2004-present)
Sigma Xi, Associate Member (1997-present)
Tau Chapter, Member (1999-present)
GAANN (Greater Advancement in Areas of National Need) Fellowship (1998-2001)
LJIS (La Jolla Interfaces in Science) Fellowship (2002-2004)

Publications
Sagle, Laura B., Zimmermann, Jörg, Dawson, Philip E., and Romesberg, Floyd E., "Sequential Unfolding of Cytochrome *c* in Urea" in preparation.

Cremeens, Matthew E., Fujusaki, Hiroshi, Zimmermann, Jörg, **Sagle, Laura B.**, Matsuda, Shigeo, Dawson, Philip E., Straub, John E., and Romesberg, Floyd E., "Efforts Toward Developing Direct Probes of Protein Dynamics" in preparation.

Sagle, Laura B., Zimmermann, Jörg, Dawson, Philip E., and Romesberg, Floyd E., "Redox-Coupled Dynamics and Folding in Cytochrome *c*" submitted to *JACS*.

Sagle, Laura B., Zimmermann, Jörg, Dawson, Philip E., and Romesberg, Floyd E., "A High Resolution Probe of Protein Folding" *Journal of the American Chemical Society*. **2004**, 126(11), 3384-3385.

Paddock, M.L., **Sagle, L**., Tehrani, A., Beatty, J.T., Feher, G., Okamura, M. Y., "Mechanism of Proton Transfer Inhibition by Cd^{2+} Binding to Bacterial Reaction Centers: Determination of the pK_a of Functionally Important Histidine Residues" *Biochemistry*. **2003**, 42(32), 9626-9632.

Adelroth, P; Paddock, ML; **Sagle, LB**; Feher, G; Okamura, MY, "Identification of the Proton Pathway in Bacterial Reaction Centers: Both Protons Associated with Reduction of Q(B) to Q(B)H(2) Share a Common Entry Point" *Proceedings of the National Academy of Sciences of the United States of America*. **2000,** *97(N24)*, 13086-13091.

Goldbeck, RA; **Sagle, L**; KimShapiro, DB; Flores, V; Kliger, DS, "Evidence for Heme-Heme Excitonic Coupling in the Soret Circular Dichroism of Hemoglobin" *Biochemical and Biophysical Research Communications*. **1997,** *235(N3)*, 610-614.

Presentations
Protein Society Conference, 2004
"Carbon-Deuterium Bonds as High Resolution IR Probes of Protein Folding", poster presentation

Biophysical Society Conference, 2002
"Binding of Cd^{2+} and Ni^{2+} to the Reaction Center Displaces a Proton", poster presentation

ABSTRACT

Carbon-Deuterium Bonds as an Infrared Probe for Protein Dynamics, Local

Electrostatics and Folding

by

Laura B. Sagle

Doctor of Philosophy in Chemistry

University of California, San Diego, 2006

Professor Andrew McCammon, Chair

The new technique being developed in the Romesberg laboratory of incorporating carbon-deuterium bonds into proteins and using them as infrared probes is further explored. Carbon-deuterium bonds are incorporated into horse heart cytochrome c through its semi-synthesis in which only the C-terminal 39 residues are accessible. Chapter 3 describes a project investigating redox-liked differences in cytochrome c by the incorporation of C-D bonds at six residues throughout the protein. It is found that when the protein is oxidized, there are both electrostatic changes as well as a greater amount of unfolded protein present only on the proximal side of the heme. The lack of consistent linewidth changes, indicating greater flexibility of the protein in the oxidized state, along with distinct changes in the amount of unfolded protein present

CHAPTER 1. INTRODUCTION

1.1 Carbon-Deuterium Bonds as an Infrared Probe in Proteins

It is truly an exciting time to be partaking in biophysical research. In recent years, there has been an outpouring of new biophysical techniques that have greatly increased our knowledge of how biological systems work. Due to paramount advances in interdisciplinary research it is now possible to monitor biological processes taking place in living cells in real time and gain insight into the path that an oxygen molecule takes on its way to the inner core of the hemoglobin protein. (Schotte, 2003 and Gaietta, 2002) Proteins have been shown to be involved in nearly every biological process ranging from DNA replication to the clotting of blood. (Voet and Voet, 1995) Their main function is to catalyze a given chemical reaction which at a fundamental level involves the selective breaking of one chemical bond over another. In order for the chemical bond to break, it must first be selectively stretched, and this stretching is best viewed with vibrational spectroscopy. The Romesberg group is currently developing a new biophysical technique in which carbon-deuterium bonds are incorporated into proteins and used as an infrared probe (Chin, 2001, 2002 and Sagle, 2004, 2005). The deuteration of a C-H bond shifts the vibrational frequency to a region of the spectrum that is unobscured by other protein vibrations (see Figure 1.1). This allows for a non-purturbative, direct measurement of a single bond within a large protein molecule, with no size limitations on the protein of interest. Through the use of chemical synthesis methods for making proteins it is possible to insert carbon-deuterium bonds site-selectively anywhere throughout
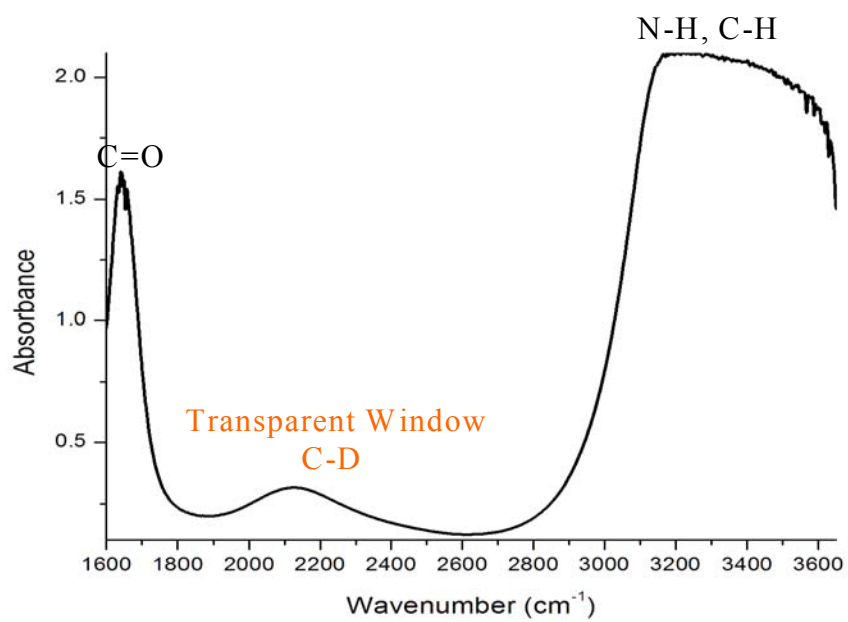
Figure 1.1. FTIR spectra of cytochrome *c* in buffer highlighting the saturated amide regions and the transparent window, where carbon-deuterium bonds absorb.

the protein, making the study of both local and general effects possible. The fast time resolution inherent to infrared spectroscopy coupled with the high molecular resolution (a single bond within a protein can be measured) should make this a very powerful technique for measuring protein dynamics, protein-protein interactions, catalysis and folding. This thesis consists of the description and further development of this technique with applications to understanding the structure-function relationship and folding of cytochrome *c*.

Like most new biophysical techniques, this technique is interdisciplinary in nature and lies at the interface between three areas of chemistry: infrared spectroscopy, chemical synthesis of proteins, and biochemistry (or biological significance). Infrared spectroscopy and the central role that carbon-deuterium bonds have played in understanding the physical chemistry of small molecules has laid the groundwork for this technique. Since the technique is dependent on synthetic access to the protein, a brief review of this field along with some other possible systems of study will also be discussed. Lastly, the ultimate goal of any new biophysical technique is to learn something new about the biological system at hand, so a discussion of cytochrome *c* as a model system will follow.

1.1.1 Spectroscopy of Carbon-Deuterium Bonds and Previous Infrared Measurements in Proteins

Infrared spectroscopy is one of the oldest and most powerful tools in determining the structure of compounds ranging from small molecules to proteins (Bellamy, 1958).

Infrared spectroscopy is not only widely used for structure determination, but through the years, physical chemists have made use of its fast timescales and ability to freeze out different conformations of a small molecule under study (Meyer, 1969). Infrared spectroscopy has played a central role in understanding the physical nature of intermolecular and intramolecular coupling, hydration phenonmenom and hydrogen bonding, and energy redistribution and relaxation. In a wide range of studies, deuteration has been a key facilitator in understanding these important issues. The oldest application of the deuteration of a C-H bond is in assigning spectral peaks, especially when several overlapping peaks are present (Suzuki, 1963). The assignment of IR spectra and deuteration of specified groups has been a tremendous aide for physical chemists in search of a more detailed understanding of the effects giving rise to splitting patterns, frequency shifts and changes in intensity. Recent work in which the C-H overtone spectra of toluene-d_0, toluene-d_1, and toluene-d_2 was measured illustrates the insight gained involving small molecule dynamics and coupling. This work revealed that coupling between the C-H stretch and the methyl torsion motion is significant compared to the coupling between adjacent C-H oscillators (Rong, 2002). Several recent IR absorption studies have also been devoted to understanding the C-H hydrogen bond and this type of hydrogen bond has been suggested to exist in biological molecules. In one study, deuterated chloroform was used as a probe to measure the hydrogen bonding interactions between chloroform and various ether compounds. It was shown that the intensity enhancement of the stretching vibration of the hydrogen bond donor (C-D) is proportional to the strength

of the hydrogen bond (Goutev, 2001). Intermolecular vibrational relaxation (IVR) has been at the forefront of understanding how chemical reactions proceed in small molecules, and deuteration has been an especially useful tool in these studies. In a pioneering set of studies by Hynes the overtone spectra of benzene and perdeuterobenzene were compared and an explanation was offered for why the spectral linewidths of benzene were so much broader than those of perdeuterobenzene. It was revealed that Fermi resonance and Coriolis coupling between the C-H stretch and ring modes (such as the C-C stretches and the C-C-H wags) were responsible for the broad lines in benzene. In perdeuterobenzene these pathways of coupling are essentially shut down due to the shift in energy of the C-D bond (Sibert, 1984). A more recent study involved a time resolved pump-probe and quantum beat measurements on the O-D stretch of d_1 and d_6-phenol. They found that not only was the lifetime of the O-D stretch five times different depending on the number of C-D bonds in the ring, but the pathway of coupling was also very different. The quantum beat measurements showed that the O-D stretch in d_1-phenol was coupled to three other states, and d_6-phenol was only coupled to one other state (Becker, 2003). Small molecule studies of IVR are of particular interest in the context of this biophysical technique, since it is believed that proteins have evolved to funnel energy in certain ways to perform catalysis.

At the same time physical chemists were developing infrared spectroscopy techniques to better understand issues concerned with coupling and reactivity, infrared spectroscopy of biological molecules was also being developed. In 1950, Elliot and

Ambrose made the first observation of a biological molecule using an infrared spectrometer. They compared the infrared spectrum of a peptide that when crystallized under different conditions existed two different conformations, an α–helix and a β–sheet. The difference in spectral features that they reported were highlighted in one region of the spectrum, later known as the amide I band, which is still today the most widely studied region of the infrared spectrum concerning biological molecules (Elliot, 1950). Figure 1.2 shows how the amide I band is currently used for secondary structure determination. The splitting pattern exhibited by the β-sheet conformer was later understood by a series of calculations carried out by Miyazawa in 1960. In the β-sheet conformation, the carboxyl groups are in close proximity and couple, yielding a splitting pattern, whereas the distance between the carboxyl groups is much larger in the α-helix conformation, so no coupling occurs (Miyazawa, 1960). In the 1960's, the full assignment of the amide I, II and III bands was carried out. The amide I band consists mostly of the carboxyl stretch with minor contributions from the CN stretch and the CCN and NH bends. The amide II band is comprised of the CN and CC stretches, as well as the NH and CO bends. And the amide III band is made up of the CN stretch and NH bend with minor contributions from the CC stretch and CO bend (see Figure 1.2) (Barth, 2002). Early infrared spectroscopy experiments on biological molecules were aimed towards secondary structure determination or the protonation/deprotonation of specific side chain groups (such as SH or COOH). More recently, many new developments in the field have come about whose examples include reaction-induced difference infrared spectroscopy, infrared dichroism and
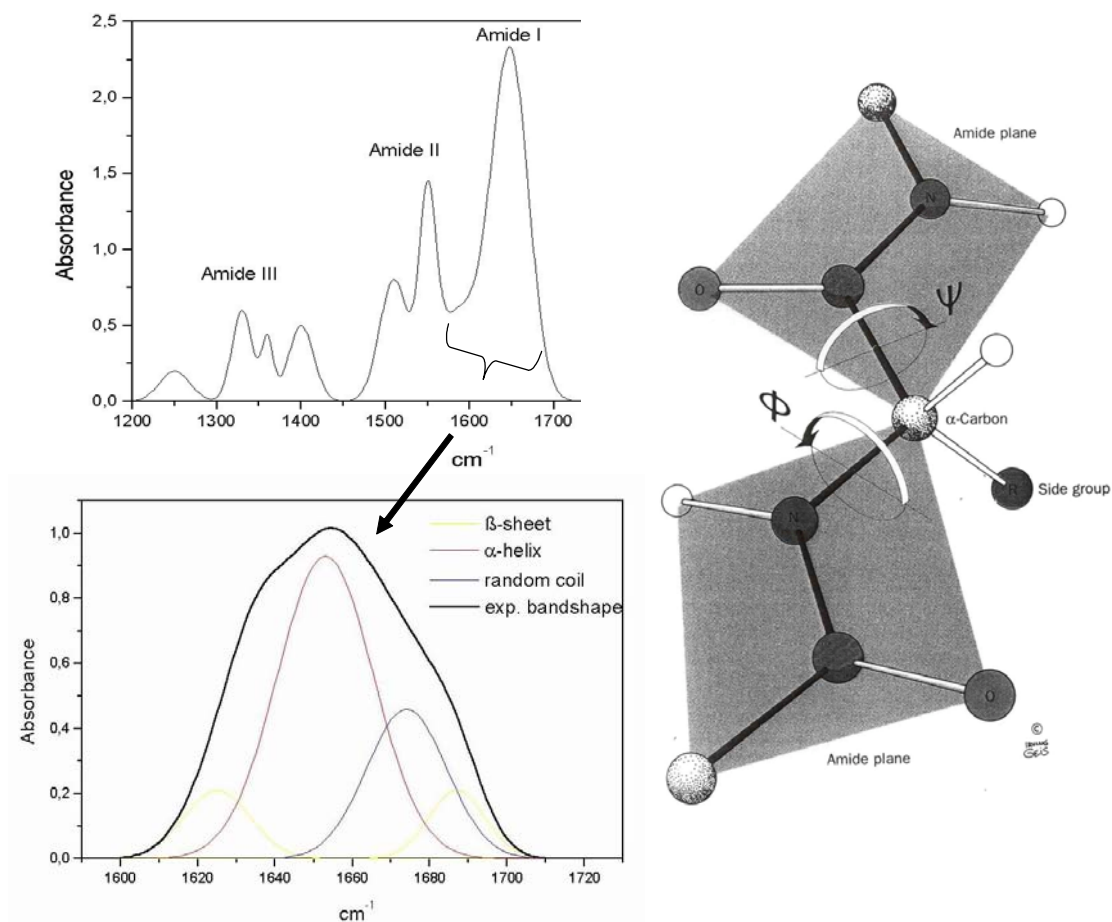
Figure 1.2. Picture depicting the amide bond in a protein showing all relevant atoms whose vibrations give rise to the amide I, II and III bands. The amide I band is most commonly used for secondary structure determination in a protein. As shown on the lower right, deconvolution of the amide I band gives rise the relative amounts of different secondary structure elements.

various time-resolved infrared techniques. With shrewd experimental planning and very careful data collection, difference infrared spectroscopy measurements can be very powerful at seeing small changes induced in a protein, even at a single side chain (Kim, 2001). The Arkin group has shown infrared dichroism to be effective at determining the orientation of proteins relative to one another or a phospholipid membrane (Torres, 2002). There are also many exciting reports of 2-dimensional time resolved infrared spectroscopy measurements which measure coupling constants between two interacting stretches, ultimately yielding information on how the groups are configured in 3-dimensional space (Zanni, 2001).

Although the infrared studies involving the amide bands have lead to tremendous gains in our knowledge of biological systems, they lack site-specific information. There are many amide bonds in a protein molecule whose frequencies overlap. Isotopically labeling a few amide bonds (with ^{13}C or ^{15}N) within the protein can sometimes allow one to discern the labeled from the unlabeled parts of the protein, but the frequency shift is small and many times requires deconvolution (Decatur, 1999). Moreover, even with modest sized proteins the amide region is so congested that many times it is not possible to distinguish labeled from unlabeled parts. Even if one is able to observe specific amide bonds within a protein, the amide bands do not consist of a single stretch but also have minor contributions from nearby stretches and bends, making local effects difficult to asses.

1.1.2 Site-Specific Incorporation of Deuterated Amino Acids into Proteins

The measurement of single carbon-deuterium bonds in a protein relies on the incorporation of a deuterated amino acid at only one site in the protein. Although it is possible to express a protein in bacteria in minimal media, supplying the growth media with the appropriate deuterated amino acid, it is impossible to get deuteration at just one site in the protein when multiple copies of a given amino acid are present. The chemical synthesis of proteins is a very useful tool in this respect, because a single deuterated amino acid can be supplied as the peptide is being made. Unfortunately, there are size limitations on the peptide that is synthesized; it is not possible to make a peptide, in good yield and purity, larger than sixty amino acids in length (Merrifield, 1995). Most proteins consist of at least one-hundred amino acids, so the use of chemo-selective ligation reactions, which link peptides together, have proven a necessity. The chemical synthesis of proteins as a field started in the 1970's when it was noticed that proteases under certain conditions could not just be made to break peptide bonds, but also form them. The classic example is the synthesis of RNase A by ligation of two peptide pieces through incubation with the protease subtilisin in glycerol (Homandberg, 1979). Since then, several other proteins have been synthesized using the technique of proteases acting in reverse, such as myoglobin and somatotropin (Graf, 1981). The semi-synthesis of cytochrome *c* through a chemical ligation reaction using a homoserine lactone is also one of the first ligation reactions performed to make a protein from peptide fragments (Wallace, 1979). This ligation reaction will be discussed in detail in Chapter 2.

In recent years, a great deal of work has been devoted to the creation of a peptide bond in order to synthesize proteins. A review of the different types of chemistry used is presented in Figure 1.3. All chemo-selective peptide ligation methods have an acyl transfer reaction that completes the peptide bond, but the initial reactions between an electrophile and nucleophile differ and can be grouped into two categories: imine and thioester. The first ligation reaction to come about containing an imine intermediate is the pseudoproline ligation (Liu, 1994). This reaction works well if the N-terminal peptide has a cystiene as the last residue, and the overall outcome produces an unnatural proline-like residue within the protein. The Staudinger ligation also involves imine chemistry and has been shown to work well if N-terminal amino acid is aliphatic (Saxon, 2000). Although the end product of this reaction is a native-like peptide bond, one must start with an unnatural N-terminal azido amino acid and the reaction is done with protected peptides under organic conditions. After the ligation reaction, the peptides containing protected side-chains must then be deprotected and the protein folded into biological aqueous solution. Ligation with a thioester intermediate is more promising in the respect that it produces a native-like cystiene residue at the site of ligation and the reaction can be performed with unprotected peptides in aqueous nondenaturing conditions (Dawson, 1997). The only drawback is that there has to be an N-terminal cystiene residue at the ligation site. Other approaches that have recently found their way into the literature are similar ligation reactions using N-terminal methionine and histidine residues, making the thioester route more versatile (Tam, 1998 and Zhang 1997).

Pseudoproline Ligation
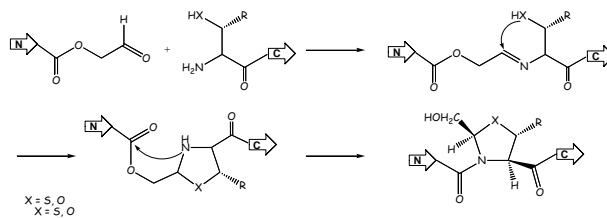
Staudinger Ligation
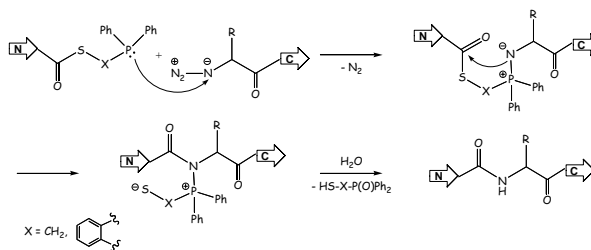
Cystiene and Selenocysteine Ligation
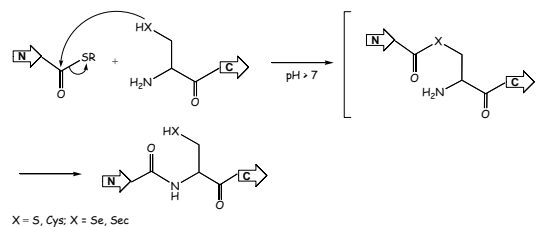
Methionine Ligation
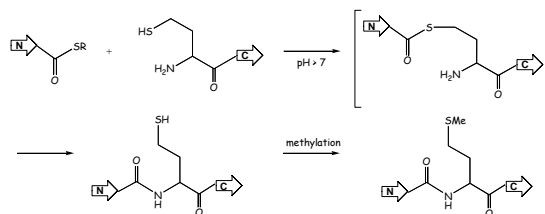
Histidine Ligation
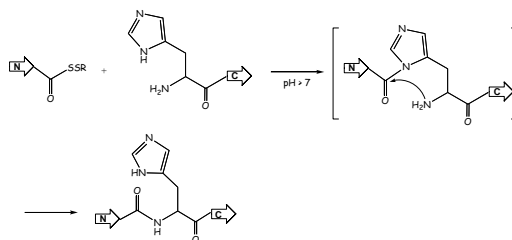
Figure 1.3. Five main types of ligation reactions for joining peptide segments together to make proteins.

Since it is only possible to synthesize relatively small peptides (60 amino acids in length), even if one ligates two large peptides together, the result is still a small protein. Unfortunately, most proteins are larger than 120 amino acids in length, and the idea of doing multiple ligation reactions in tandem is an attractive alternative for larger proteins. A few tandem ligation strategies have been recently developed that invoke protecting groups at the different sites of ligation that are photo labile or reduction-sensitive (Muir, 1997). In this manner, a ligation reaction can be performed where the other sites of ligation are protected, then the protecting groups at other sites can be removed sequentially with subsequent ligation reactions. For larger proteins, semi-synthesis is also an option, where the protein is expressed in bacteria and digested into peptide fragments with a protease, then ligated to synthetic peptides to make the whole protein. A particularly elegant approach was developed by the Muir group termed expressed protein ligation (Ayers, 1999). This new semi-synthetic technique allows for the expression of pieces of a protein with either N-terminal cystiene residues or C-terminal thioester groups. After expression, the synthetic peptide and expressed fragment can easily be joined using the ligation strategy mentioned above that uses a thioester group as the capture. Another option for the incorporation of deuterated amino acids into proteins is a technique involving no chemical ligation being developed in the Shultz laboratory (Ellman, 1991). This technique involves placing an amber nonsense codon within the gene of the protein being expressed where the amino acid of interest lies. This amber codon is not recognized by common tRNAs involved in protein synthesis. Instead a suppressor

tRNA is constructed which is chemically aminoacylated with the unnatural amino acid and recognizes the amber codon. This method could prove the most powerful at generating proteins with chemically modified amino acids in large quantity and more work is currently underway to increase efficiency (Liu, 1997). In the Romesberg group, we are currently exploring ligation through a thioester intermediate and expressed protein ligation as a way to extend this technique to the SH3 domain and dihydrofolate reductase.

1.1.3 Cytochrome *c*: A Good Model System

The major motivating factor for the development of any new biophysical technique is the promise that it will complement current studies and lead to a better understanding of biology. Through the years, many new biophysical techniques have greatly increasing our knowledge of how biological systems work and cytochrome *c* has been at the forefront of these discoveries. The function of cytochrome *c* was elucidated fairly soon after its discovery as being an electron carrier protein. As shown in Figure 1.4, cytochrome *c* is found in the intermembrane space between the inner and outer mitochondrial membrane. Its function is to retrieve an electron from the cytochrome *c* reductase complex and transfer it to the cytochrome *c* oxidase complex. This is an integral part of the electron transfer chain which ultimately reduces oxygen to water (Voet, 1995). The last protein complex located in the electron transfer chain, the ATPase, utilizes proton translocation across the membrane to drive the chemical reaction which makes water from oxygen. Thus, the ATPase

complex supplies respiratory energy in the form of ATP molecules to the cells by pumping protons across an electrochemical gradient across the membrane produced by the electron transport chain. As shown in Figure 1.4, mammalian cytochrome *c* is a protein consisting of 104 amino acids, contains only α-helical secondary structure elements, and has a heme group covalently bound to the protein through two thioester linkages at residues Cys14 and Cys17. There are two protein-based ligands, the distal ligand being His18 and on the proximal side, under standard conditions, Met80 is the sixth ligand.

Since its discovery by David Keilin in the 1920's, there have been more than 15,000 publications on cytochrome *c* in a variety of journals. Due to its robustness, cytochrome *c*, along with hemoglobin, was one of the first proteins to be discovered. The heme group covalently bound to the protein fascinated early scientists with its changing color upon oxidation and reduction. Cytochrome *c* was one of the first proteins whose visible spectra was acquired and assigned (Dixon, 1931). It was also one of the first proteins whose amino acid sequence was assigned in 1961 and the third protein for which an x-ray crystal structure was obtained (Dickerson, 1968). The amino acid sequence determination of horse heart cytochrome *c* facilitated the sequence determination of other species of cytochromes. The cytochrome family of proteins then became the model system to study the evolutionary relationship of phylogeny and primary sequence (Margoliash, 1963). The phylogenetic comparisons which identified conserved residues as well as the ease to which the cytochromes were
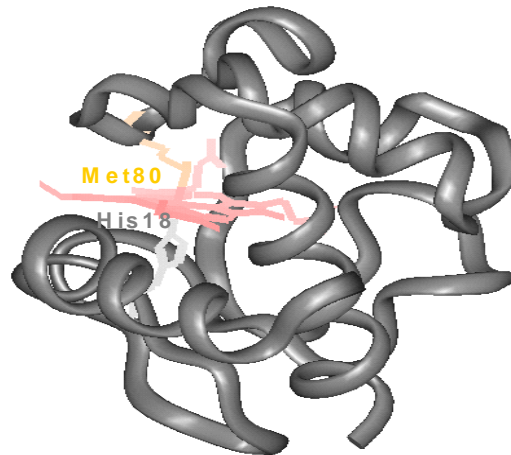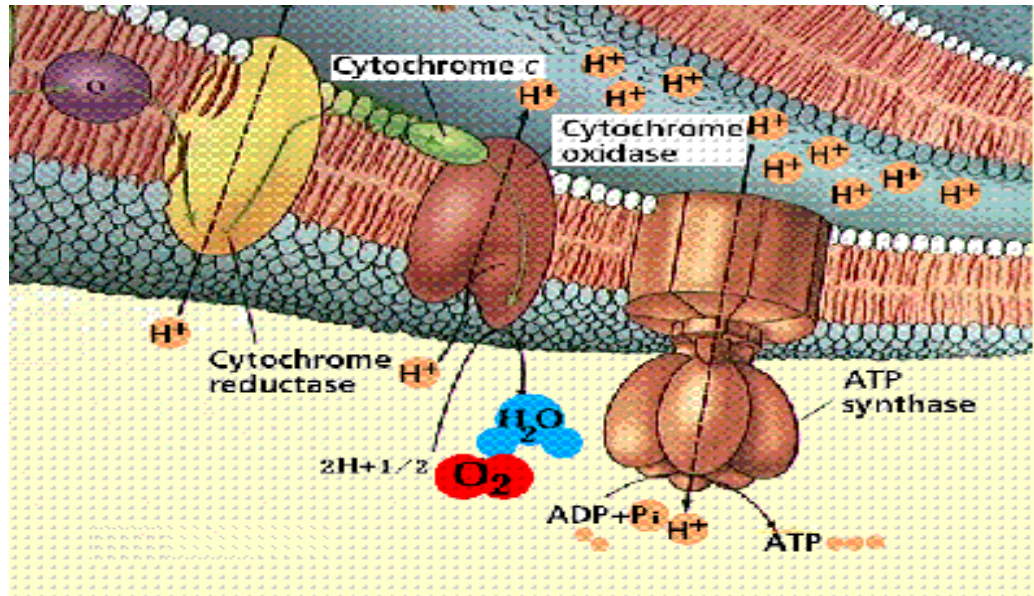
Figure 1.4. Three out of the five protein complexes that comprise the electron transport chain for mammalian mitochondria. Cytochrome *c* transfers an electron from cytochrome *c* reductase to cytochrome c oxidase where two protons are translocated across the membrane. The ATPase complex then makes use of this proton gradient to make ATP, a form of cellular energy. The lower shows the crystal structure of oxidized horse heart cytochrome *c*. The polypeptide chain is shown as a ribbon with the side chains of the proximal Met80 and distal His18 ligands. The heme cofactor is shown in red.

purified by charge using ion-exchange column chromatography paved the way for mutagenesis studies. Cytochrome *c* was one of the first proteins to be expressed in *E. coli* and used as a model system for protein engineering and mutagenesis (Montgomery, 1978). The early discovery of the semi-synthesis of the protein and the ability to carry out chemical modification also strengthened cytochrome *c*'s standing as a good system with which to probe protein structure and function (Corradin, 1974). Furthermore, primary experiments identified the redox activity of the protein and revealed that it could be easily reacted *in vitro* with non-physiological oxidants and reductants. This worked to also make cytochrome *c* a model system for studying redox proteins and testing theories of electron transfer (Ferguson-Miller, 1979).

In more recent years, cytochrome *c* is still being used as a model system for new biophysical techniques. The first electron paramagnetic resonance (EPR), magnetic circular dichroism (MCD) and resonance raman studies were carried out on cytochrome *c* (Sutherland, 1972 and Spiro, 1972). A fluorescent tryptophan residue (the only tryptophan residue in the whole protein) was found to be quenched by the heme in the folded state of the protein, leading to pioneering fluorescence studies measuring changes in protein conformation. In the last few years, an exciting technique has been developed by the Roder and Englander laboratories using cytochrome *c* which combines amide-exchange with deuterium and NMR spectroscopy (Bai, 1995). This is proving to be one of the most powerful techniques for measuring both equilibrium and time-resolved protein folding. The Gray laboratory has also recently developed a way of phototriggering the folding of

cytochrome *c* and for the first time, the folding of a protein is being measured with timescales fast enough to observe the nascent collapse of the protein (Pascher, 1996).

For several reasons, cytochrome *c* appeared to also be a good model system for the development of our technique. Its small size makes it very amenable to computational studies which have and will greatly aid in our understanding of the observed spectral features. The small number of amino acids also makes probing residues of interest not a formidable task. The fact that the protein is very soluble in aqueous buffer is also a big advantage, since the extinction coefficients of carbon-deuterium bonds are small. The heme group, its many peaks in the UV-Vis spectra, and visible color associated with it makes its purification and assessment of oxidation state, folded state and concentration a simple task. The protein is easily purchased in purified form which makes a nice experimental standard with which to compare the synthetic protein. But perhaps the biggest advantage to using cytochrome c in our initial studies is that it is very well characterized and therefore straightforward to put any new observation in the context of biological significance.

1.2 Previous Measurements Using Carbon-Deuterium Bonds in Cytochrome *c*

Two previous studies using carbon-deuterium bonds as infrared probes in cytochrome *c* have been carried out by Dr. Jodie Chin in the Romesberg group. In both papers, horse heart cytochrome *c* was expressed in *E. coli* and d_3-methionine supplied to the media. The protein contains two methionine residues, Met65 and Met80. One of the methionine residues, Met65 (a surface residue) was mutated to

leucine, so that only one deuterated residue remained as a probe; Met80, a ligand to the heme. The Met80 ligand has been shown to play a direct role in controlling the redox properties of the protein. In the reduced state of the protein, the Met80 ligand is more strongly bound to the heme (it is easily displaced in the oxidized state) and thus the protein itself stabilizes the reduced state. Both studies carried out by Dr. Chin were aimed at a better understanding of the role the Met80 ligand plays in the redox behavior of the protein.

The first study was a pioneering report showing that it is possible to incorporate carbon-deuterium bonds into proteins and directly observe these bonds (Chin, 2001). This study measured the symmetric stretch of a deuterated ε-methyl group of Met80 in the oxidized and reduced state of horse heart cytochrome c. As shown in Figure 1.5, a frequency shift of 2.5 cm^{-1} was observed when the protein changed oxidation state, with the reduced state having a lower frequency. Three possibilities were presented to explain this frequency shift. The first, proposed by Brayer, is that the Tyr67-Met80 hydrogen bond, only present in the oxidized state, increases the electron withdrawing strength of the sulfur ligand. This increases electron density in the neighboring carbon-deuterium bonds and causes a rise in frequency. The second alternative is that the reduction of the heme causes changes in the electron density at the sulfur atom, which in turn increases the electron donation into the antibonding (σ*) orbitals or takes away density from the σ orbitals of the carbon-deuterium bonds, both decreasing the frequency. The third option is that the spectral shift is a result from a change in polarity of the local environment around the carbon-deuterium bonds. To investigate

	Frequency	FWHM
Fe(III)	2128.6	6.4
Fe(II)	2126.2	6.0
Fe(III)/cyanide	2132.1	12.8
Fe(III)/nitric oxide	2131.3	10.4
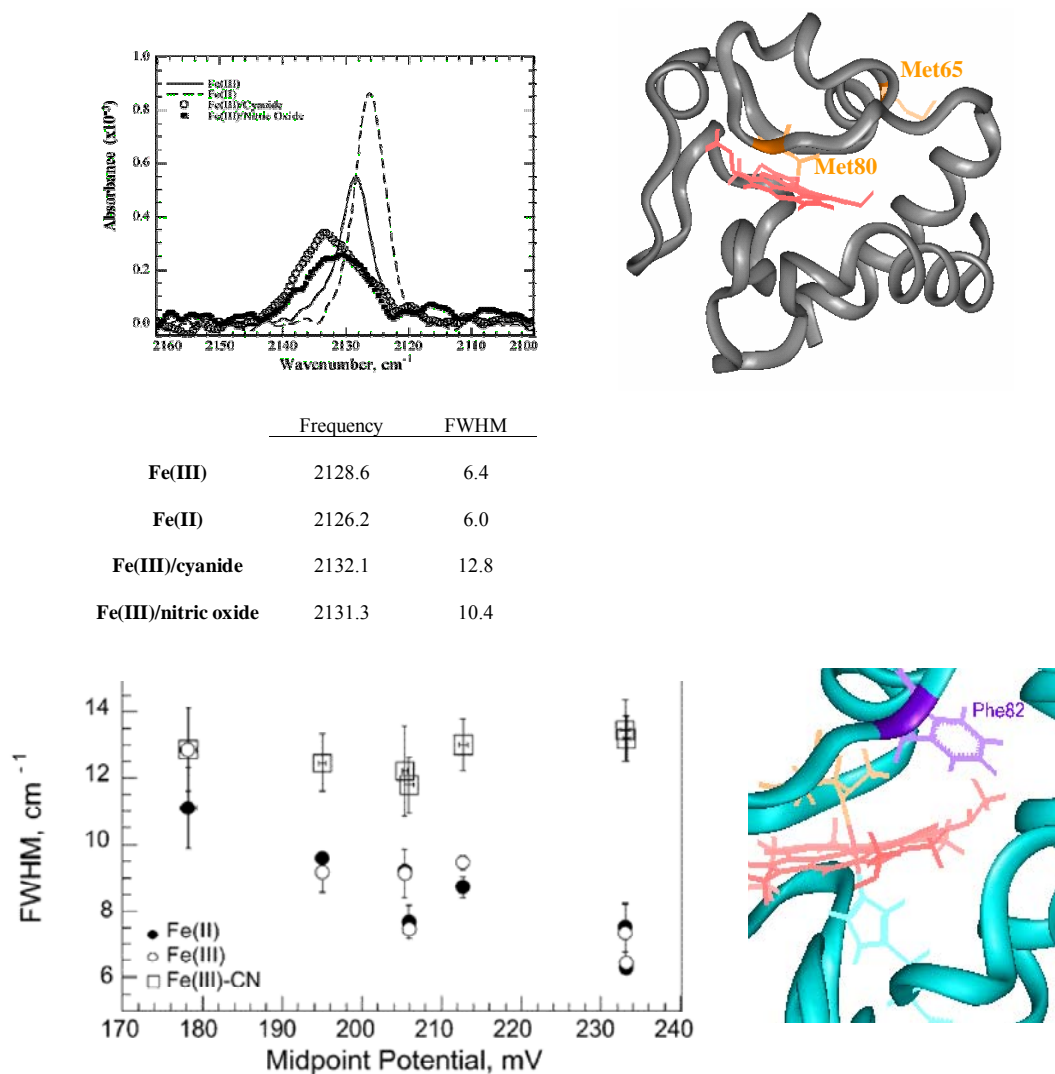
Figure 1.5. Previous data using carbon-deuterium bonds as IR probes in cytochrome *c*. The ε-methyl group of the Met80 ligand was measured as the protein changed redox state. The first study showed small frequency shift and no linewidth change accompanies oxidation of the heme. In the second study, the redox potential of the protein was changed by mutating Phe82 and a correlation was observed between the linewidth of the carbon-deuterium vibrations of Met80 ligand and the redox potential.

the role of the Tyr67-Met80 hydrogen bond, a Tyr67Phe mutant was constructed where the hydrogen bond was lost. In the oxidized form of the protein this mutation had no effect on the frequency, however in the reduced form a small 1 cm^{-1} red shift was observed. Based on structural data, the Tyr67-Met80 hydrogen bond is thought to only exist in the reduced state of the protein and be broken in the oxidized form. The fact that a frequency shift was only observed in the reduced protein suggests that possibly this hydrogen bonding could be playing a role, but still doesn't account for the magnitude of the shift in the Met80 carbon-deuterium bonds upon reduction of the protein. To further evaluate the possibility of through-space polarity giving rise to the observed frequency shift, the Met80 ligand was displaced with NO and CN ligands. Even though the two ligands have different charge the observed frequency of the Met80 carbon-deuterium bonds were very similar, indicating through-bond effects rather than through-space effects are the dominant factor. This frequency shift of the Met80 ligand upon reduction of the heme was evaluated computationally and the results will be further discussed in Chapter 2. The computational results confirm that through-bond effects are dominant and changes in electron density as well as geometry give rise to the observed frequency shift.

It is currently believed (see Chapter 2) that the frequency of the observed spectra is primarily sensitive to local environmental polarity and the strength of the bond whereas the linewidths (full-width-at-half-maximum) report on the local motions or local flexibility. Interestingly, it has been proposed by several groups that the oxidized form of the protein is more flexible, yet these results for the Met80 ligand

show no change in linewidth upon reduction of the heme. Therefore, the second report, which also looked at the deuterated ε-methyl group of the Met80 ligand, looked in detail at the correlation between linewiths of the carbon-deuterium bonds and redox potential (Chin, 2002). As shown in Figure 1.5, Phe82 has been shown to pack close to the Met80 ligand and drastically change the redox potential when mutated to other residues. Six mutants were constructed in which Phe82 was changed to Histidine, Serine, Leucine, Valine, Alanine and Glycine. In all mutants, the frequency and linewidth of the Met80 methyl group symmetric stretch was recorded. Interestingly, in each mutant the linewidths of the oxidized and reduced state were identical. As shown in Figure 1.3, there was a dependence of the observed frequencies on the redox potential, but only in the oxidized state. It is expected that the change in redox potential would also lead to changes in the polarity of the protein environment around the heme. The S-Fe^{3+} bond is expected to be more polarizable than the S-Fe^{2+} bond and may be more susceptible to changes in local polarity. There is also a trend observed in the linewidths of the Met80 methyl group at different redox potentials and again, the oxidized and reduced proteins show the same trend. This trend is lost when the sulfer-iron bond of the Met80 ligand is lost by replacement with cyanide. This change in linewidth highlights the fact that the structure of the protein is intimately connected to the heme cofactor and possibly indicates the ability of the protein to reorganize itself with the change in redox potential. Surprisingly, these results were able to reveal observations, such as the reduced and oxidized state showing no change in flexibility, that were not uncovered using other biophysical techniques on one of the

most well-studied systems. Therefore, these first two reports bring out the need for a localized direct probe within the protein. Further studies investigating the relationship of the oxidation state of the heme to the local protein environment were carried out on more residues throughout the protein and will be discussed in detail in Chapter 3.

1.3 Background to Studies on the Redox Behavior of Cytochrome *c*

Since the early 1980's when the crystal structure of cytochrome *c* was revealed in both the oxidized and reduced forms, one important and elusive question surfaced: how is it that a protein whose function is to transfer an electron can have such similar structures in the oxidized and reduced states (Takano, 1980)? How the structure of cytochrome *c* mediates its function is still widely studied and not well understood. Several factors have been measured by various techniques showing that the structural differences between the two oxidation states do exist, but are very subtle. Since the first study on the oxidized and reduced forms of tuna cytochrome *c* by Takano and Dickerson, several other species of cytochromes have been crystallized as well as the acquisition of NMR structures of various species (Berghuis, 1992). The conformation of the protein in both redox states is remarkably similar, with the exception of a few small differences in thermal factors (indicating greater flexibility in certain regions), the heme structure, internal water molecules and hydrogen bond interactions. These structural differences observed when the redox state of the protein is changed are highlighted in Figure 1.6 (Scott, 1996). Three regions of the protein showed higher thermal factors in the crystal structure in the oxidized form of the protein: residues
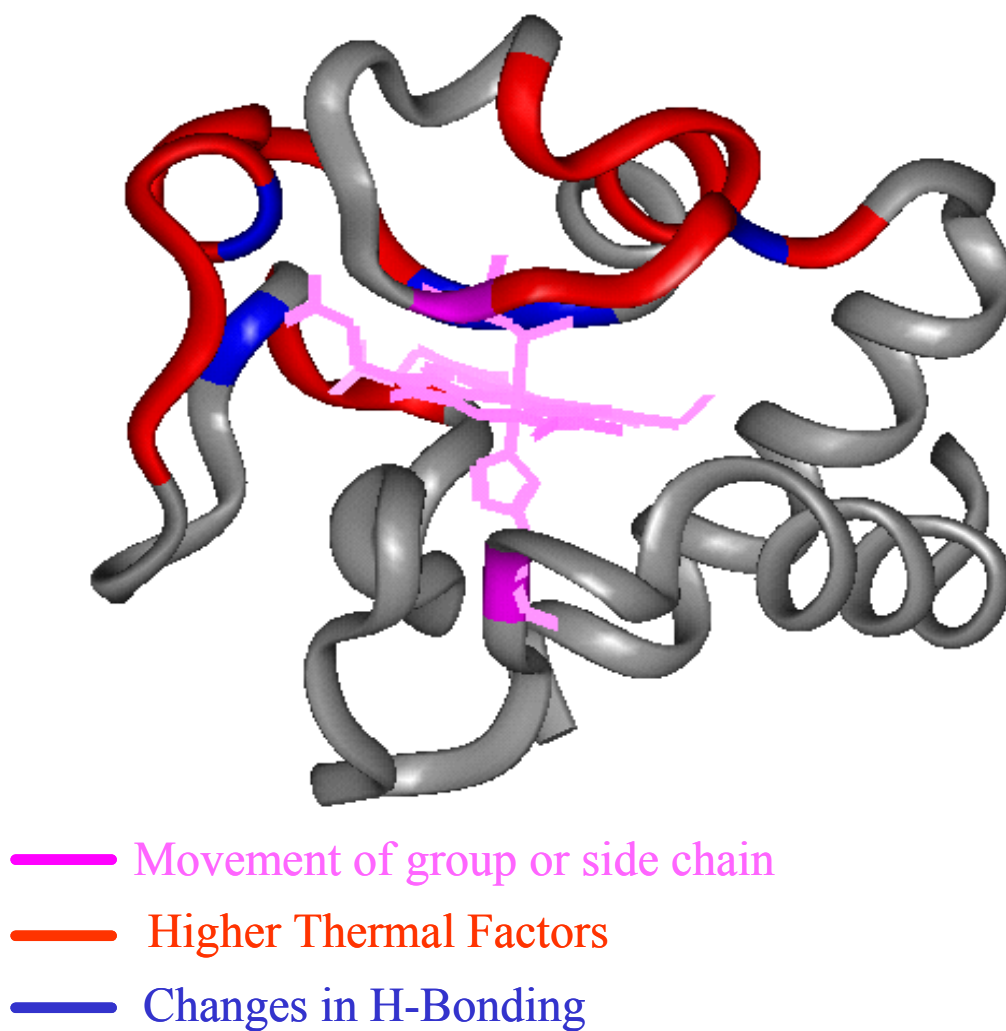
Figure 1.6. Crystal structure of oxidized horse heart cytochrome *c* showing areas of the protein where changes occur during reduction of the heme. Note that changes are most pronounced on the proximal side of the heme, where the Met80 ligand is bound.

around Asn52, residues around Tyr67 and residues around Phe82 (Louie, 1990). Interestingly, as previously discussed, Asn52 and Tyr67 are involved in a hydrogen bonding interaction only present in the reduced state and Phe82 packs close to the heme and greatly influences the redox potential of the protein when mutated. There are also changes observed in the vicinity of the heme cofactor such as increased distortion of the heme planarity and readjustment of some of the propionate groups upon oxidation. The ligands bound the heme undergo slight modification as a result, the His18 ligand rotates about the imidazole ring plane and the Met80 side chain shows higher thermal factors with oxidation. The largest displacement of internal water molecules upon oxidation occurs in Wat166, which incidentally is intimately connected to a hydrogen bond network extending from Asn52, Thr78, Tyr67 and finally Met80. This hydrogen bond network, which is lost in the oxidized state, is thought to influence the Met80 ligand and possibly explain both the greater affinity of the ligand and greater stability of the protein in the reduced state.

Along with subtle structural differences, the oxidized and reduced forms of the protein differ in their hydrodynamic properties, stability and UV-Vis spectra. The oxidized form of the protein has a larger intrinsic viscosity, increased adiabatic compressibility, and larger radius of gyration than the reduced form (Eden, 1982 and Trewhella, 1988). In addition, the oxidized form is more susceptible to proteolytic digestion (Nozaki, 1958). Several studies have also measured amide H/D exchange rates for various residues throughout the protein, and the rates are consistently faster in the oxidized form (Liu, 1989). The binding of exogenous ligands in the oxidized form

is much more favorable than the reduced form, due to the fact that the S-Fe bond from the Met80 ligand in the reduced form is much stronger and harder to displace (Horecker, 1946). The strong S-Fe bond of the Met80 ligand in the reduced form is also manifested in the enhanced intensity and splitting of its visible bands at liquid nitrogen temperatures (Estabrook, 1957). Lastly, the reduced form has been found to be much more stable upon denaturation, possibly due to the strengthening of the Met80 ligand, which has been proposed to be the first step in the unfolding of the protein (Dickerson, 1975). Although the changes observed upon oxidation of the heme appear to be subtle and localized, the majority of studies have used techniques that are general and indirect. As illustrated by previous reports, carbon-deuterium bonds yield residue-specific information and are able to report on local effects. Despite the fact that the redox behavior of cytochrome *c* has been widely studied through the years, there is a lack of site-specific information at residues surrounding the heme, which are expected to be the most sensitive to the change in redox state. Chapter 3 describes our efforts towards characterizing cytochrome *c* in the oxidized and reduced states at residues around the heme.

1.4 Background to Studies on the Folding of Cytochrome *c*

In order for a protein to perform its function in living cells, it must first be folded into the right configuration, which has been shown through structural studies to be intimately linked to its function. In 1968, Levinthal proposed a paradox that given the size and conformational space available to a protein, finding a particular conformation,

or folded state, should take on the order of 10^{10} years (Levinthal, 1968). The fact that protein folding proceeds in nature on subsecond timescales has highlighted the complexity and need for a better understanding of even the most basic process in biology. The field of protein folding has grown tremendously over the years attracting both experimental and theoretical interest. Recently, the misfolding of proteins has been implicated as the cause of several diseases such as Mad Cow and Alzheimer's. Cytochrome *c*, due to its early discovery and ease of experimental handling, has been a paradigm for folding studies.

Equilibrium folding studies of oxidized cytochrome *c* have for many years been easily fit to a two-state model, where the protein simply goes from unfolded to folded in one step, showing no intermediates along the way. Several tyrosine residues throughout the protein, a single tryptophan residue whose fluorescence is quenched by the heme in the folded state, and the heme group itself have allowed for convenient monitoring of the protein as it unfolds. Several optical techniques have been used to measure the folding of cytochrome *c;* fluorescence spectroscopy, UV-Vis spectroscopy and circular dichroism. Remarkably, every probe, under the same conditions, shows the same transition, indicating a simple two-state mechanism of folding (Tsong, 1974 and Myer, 1968). Equilibrium thermal and denaturant induced unfolding not only show a two-state mechanism but NMR and EPR studies indicate that the unfolded state still has the two protein-based ligands attached to the heme (Babul, 1971). The first sign of any deviation from two-state behavior came when the protein was unfolded under acidic conditions. Studies using NMR spectroscopy

identified an extra peak present as the protein was unfolded, which was the signature of a strong field ligand, such as histidine, bound to the heme instead of the native-state methionine (Hartshorn, 1989). The addition of sodium chloride to the acid unfolded protein also yielded unexpected results. The protein formed a more compact state, known as the molten globule state, which NMR measurements revealed as having some native-like structure in the C and N-terminal helices (Hamada, 1996). Further investigation of pH effects on the folding of cytochrome c lead to some interesting observations when the protein was unfolded at high pH. Similar to the acid unfolding, NMR studies revealed at alkaline pH values, other strong field ligands were undergoing misligation (Osterhout, 1985).

Recent equilibrium folding studies on oxidized cytochrome c have at the same time answered many questions and unveiled new areas of uncertainty. The misligation at acid and alkaline pH values is now well understood. At acidic pH, two histidine ligands are known to take the place of Met80, both located in the N-terminal half of the protein, His26 and His33 (Colon, 1997). At alkaline pH, three lysine residues have been implicated in misligated intermediates, all three located close in proximity to Met80, Lys72, Lys73 and Lys79 (Dopner, 1998). Furthermore, recent NMR studies by Bren have revealed that even at neutral pH values using chemical denaturant, misligation also occurs. Through mutagenesis studies it has been shown that at neutral pH in urea, possibly only Lys79 is involved in misligation of the heme (Russell, 2000). However, neutral pH studies using guanidine hydrochloride have shown a number of residues undergo misligation, but interestingly, the same residues

implicated at high and low pH, His26, His33, Lys72, Lys73 and Lys79 (Russell, 2002). The invention of a new technique that combines hydrogen-deuterium exchange with NMR spectroscopy has produced recent folding results on cytochrome *c* challenging the two-state cooperative model (Maity, 2005). These studies show the protein in fact unfolds in a sequential pattern, with parts of the protein that are less stable unfolding first. The loop containing the Met80 ligand was shown to be least stable, unfolding first, followed by the nearby 60's helix containing residues 60-68. The most stable unit in the protein consists of the C and N-terminal helices which were shown to unfold last. A map of cytochrome c showing the folding units and order of stability as well as the residues involved in misligation is shown in Figure 1.7. More studies have recently surfaced that show a similar trend. An equilibrium 2D-IR study has shown the same ordering of stability among folding units throughout the protein (Filosa, 2001). Further evidence of the stability of the C and N-terminal helices has come about through the observation that residual structure in the C and N-terminal helices even at high denaturant concentrations still exists. Another recent controversy has surfaced concerning the nature of the unfolded state of the protein. Until recently, it was assumed that the unfolded state measured by various techniques consisted of a fully extended random coil. Small angle x-ray scattering as well as fluorescence energy transfer were measured on equilibrium unfolded cytochrome c in 6 M guanidine hydrochloride and it was discovered that a significant amount of compact structure exists (Segel, 1998 and Lyubovitsky, 2002). The nature of this
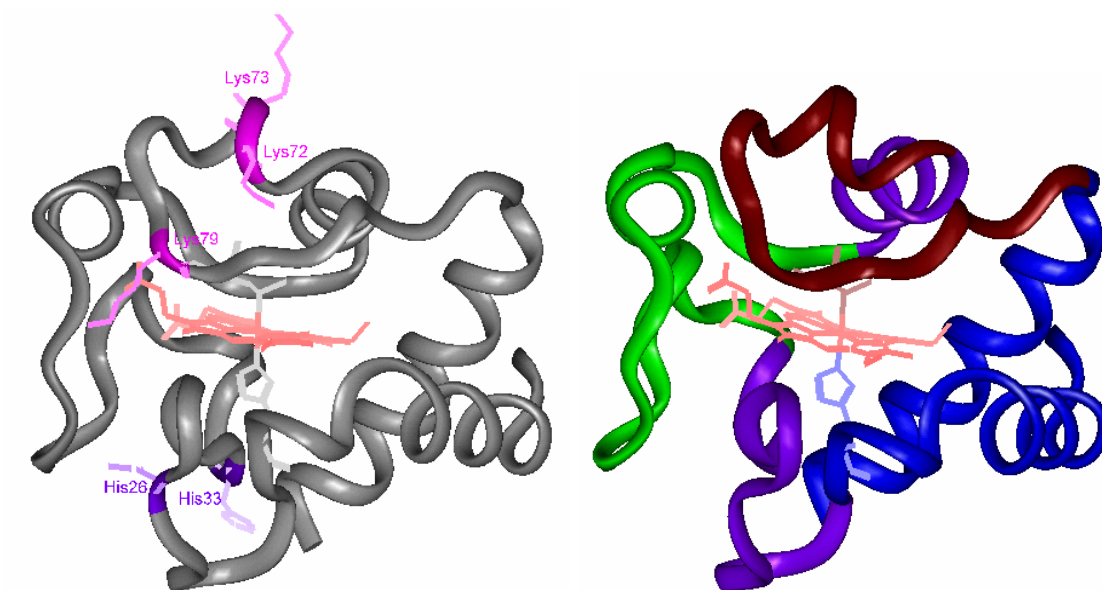
Figure 1.7. Left. Crystal structure of horse heart cytochrome c showing all residues involved in the replacement of Met80 and misligation of the heme. The lysine residues misligate at high pH, the histidines at low pH and a mixture of both with denaturants at neutral pH. Right. Amide-exchange NMR results for horse heart cytochrome c color-coded to show cooperative folding units within the protein. The Met80 loop is the least stable (brown), next is the 50's loop (green), then the 60's helix with the 20-35 loop (purple) and the most stable is the N and C-terminal helices (blue).

compact structure remains unknown, but there is evidence that the compact structure is most prevalent at the N and C-terminal helices.

Once it was determined that the equilibrium folding of oxidized cytochrome *c* does populate partially folded intermediates, the question remained as to whether these intermediates were just misfolded conformers that never folded to the correct configuration or obligatory intermediates that were necessary stops along the pathway to the native state. The best way to address the question of on- or off-pathway intermediates was kinetic folding measurements which measure the folding of the protein in real time. The invention of several new techniques to probe and initiate the folding of cytochrome *c*, such as submilisecond mixing and temperature jump studies, have greatly aided in the current understanding of how the protein folds. Kinetic studies have revealed that the protein folds in three distinct phases, one fast phase which accounts for 80 percent of the total amplitude, and two slower phases (Nall, 1983). The fast phase, or burst phase, has been attributed to the nascent collapse of the protein, where a compact intermediate is formed with very little native-like contacts. One of the slow phases has been assigned to proline isomerization. Proline residues in solution can exist in either a cis or trans configuration, but in the folded form of a protein must exist as only one of these forms. There is evidence that as proteins fold, proline residues fall into the wrong configuration, causing the protein to misfold and undergo local conformational changes to reach to right conformation (Brandts, 1975). The second slow phase is due to misligated intermediates also observed in equilibrium studies. Kinetic mutagenesis studies have also shown that the

residues involved in misligation in equilibrium studies under the same conditions are also doing so in real time (Shastry, 1998). This offers direct evidence that the misligated forms of the protein are partially folded intermediates that are populated as the protein folds and perhaps help lead to the final native state of the protein.

The reduced state of the protein, being much more stable to unfolding conditions, was shown to fold in a different manner than the oxidized form. Equilibrium studies have not revealed any partially unfolded intermediates or any signs of misligation, possibly due to the increased stability of the Met80 ligand in the reduced form (Bhuyan, 2001). Similar to the oxidized state, kinetic studies show a burst phase which accounts for a large percentage of the total amplitude and is also assigned as the nascent hydrophobic collapse of the protein. Following the burst phase is a slower monophasic phase indicative of the two-state folding of the protein (Kumar, 2005). Recent studies scrutinizing the kinetic data in the reduced state have found evidence for a second slower phase. It is now believed that the slower phase may also be due to misligation with the heme, possibly with the other methionine residue nearby, Met65 (Pabit, 2004).

The oxidized form of horse heart cytochrome *c* is well studied, experimentally easier to handle than the reduced form, and shows interesting folding patterns in both equilibrium and kinetic studies. Not only have partially unfolded intermediates been shown to exist more so in the oxidized state than the reduced, but the intermediates are well characterized. Particularly interesting are the misligated intermediates in the oxidized state, which have been shown in both equilibrium and time-resolved studies

and are believed to lie directly on the pathway to folding. For these reasons, we have chosen to focus on the folding of the oxidized form of the protein. Chapters 4 and 5 describe the first application of using carbon-deuterium bonds as a probe to measure the equilibrium folding of cytochrome c.

1.5 References

Ayers, B., Blaschke, U., Camarero, J., Cotton, G., Holford, M. and Muir, T. W., Introduction of Unnatural Amino Acids into Proteins Using Expressed Protein Ligation. *Biopolymers* **1999**, 51, 343.

Babul, J., Stallwagen, E., Existence of Heme-Protein Coordinate-Covalent Bonds in Denaturing Solvents. *Biopolymers*, **1971**, *10*, 2359.

Bai, Y., Sosnick, T. R., Mayne, L., and Englander, S. W., Protein Folding Intermediates: Native-State Hydrogen Exchange. *Science* **1995**, 269, 192.

Barth, A., and Zscherp, C., What Vibrations Tell Us About Proteins. *Quarterly Reviews of Biophysics* **2002**, 35, 369.

Becker, J., Becher, F., Hucke, O., Lebahn, A., and Koslowski, T., Theory and Simulation of Vibrational Coupling in Ceuterated Proteins: Toward a New Structural Probe? *J. Phys. Chem. B.* **2003,** 107, 12878.

Bellamy, L., *The Infrared Spectra of Complex Molecules, 2ⁿᵈ Edition*. John Wiley & Sons, Inc, New York, New York, 1958.

Berghuis, A. M., and Brayer, G. D., Oxidation State Dependent Conformational Changes in Cytochrome *c*. *J. Mol. Biol*., **1992**, *223*, 959.

Bhuyan, A., and Udgaonkar, J. B., Folding of Horse Cytochrome *c* in the Reduced State. *J. Mol. Biol*., **2001**, *312*, 1135.

Brandts, J. F., Halvorson, H. R., and Brennan, M., Consideration of the Possibility that the Slow Step in Protein Denaturation Reactions is Due to Cis-Trans Isomerism of Proline Residues. *Biochemistry*, **1975**, *14*, 4953.

Chin, J. K., Jimenez, R., and Romesberg, F. E., Direct Observation of Protein Vibrations by Selective Incorporation of Spectroscopically Observable Carbon-Deuterium Bonds in Cytochrome *c*. *J. Am. Chem. Soc.* **2001,** 123, 2426.

Chin, J. K., Jimenez, R., and Romesberg, F. E., Protein Dynamics and Cytochrome c: Correlations Between Ligand Vibrations and Redox Activity. *J. Am. Chem. Soc.* **2002,** 124, 1846.

Colon, W., Wakem, L., Sherman, F., Roder, H., Identification of the Predominant Non-Native Histidine Ligand in Unfolded Cytochrome *c*. *Biochemistry*, **1997**, *36*, 12535.

Corradin, G., and Harbury, H. A., Reconstitution of Horse Heart Cytochrome *c* – Reformation of Peptide Bond Linking Residues 65 and 66. *Biochem. Biophys. Res. Comm.*, **1974**, *61*, 1400.

Dawson, P. E., Churchill, M. J., Ghadiri, M. R., and Kent, S., Modulation of Reactivity in Native Chemical Ligation Through Use of Thiol Additives. *J. Amer. Chem. Soc.* **1997**, 119, 4325.

Decatur, S. M., and Antonic, J., Isotope-Edited Infrared Spectroscopy of Helical Peptides. *J. Amer. Chem. Soc.* **1999**, 121, 11914.

Dickerson, R. E., Kopka, M. L., Weinzierl, J. E., Varnum, J. C., Eisenberg, D., and Margoliash, E. *An Inerpretation of a Two-Derivative 4Å Resolution Electron Density Map of Horse Heart Ferricytochrome c*. Ed.; Okunuki, K., Kamen, M. D., and Sekuzu University of Tokyo Press, Tokyo, Japan, 1968.

Dickerson, R. E., and Timokovich, R. The Enzymes, 3rd Edition, Volume II. Ed.; Boyer, P. D., Academic Press, New York, New York, 1975.

Dixon, M., Hill, R., and Keilin, D., Proc. Roy. Soc. London Series B, vol 109, pg 29 1931

Dopner, S. Hildebrandt, P., Rosell, F., and Mauk, A. G., Alkaline Conformational Transitions of Ferricytochrome *c* Studied by Resonance Raman Spectroscopy. *J. Am. Chem. Soc.*, **1998**, *120*, 11246.

Eden, D., Matthew, J. B., Rosa, J. J., and Richards, F. M., Increase in Apparent Compressibility of Cytochrome *c* upon Oxidation. *Proc. Natl. Acad. Sci. USA*, **1982**, *79*, 815.

Elliot, A., and Ambrose, E. J., Structure of Synthetic Polypeptides. *Nature*, **1950**, *165*, 921.

Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C. J. and Shultz, P. G., Biosynthetic Method for Introducing Unnatural Amino Acids Site-Specifically into Proteins. *Method in Enzymology* **1991**, 202, 301.

Estabrook, R. W., The Low Temperature Spectra of Hemoproteins. I. Apparatus and its Application to a Study of Cytochrome *c*. *J. Biol. Chem.*, **1956**, *223*, 781.

Ferguson-Miller, S., Brautigan, D. L., and Margoliash, E. *The Prophyrins,Vol III*. Ed.; Dolphin, D., Academic Press, New York, New York, 1979.

Filosa, A., Wang, Y., Ismail, A. A., and English, A. M., Two-Dimensional Infrared Correlation Spectroscopy as a Probe of Sequential Events in the Thermal Unfolding of Cytochromes *c*. *Biochemistry*, **2001**, *40*, 8256.

Gaietta, G., Deerinck, T. J., Adams, S. R., Bouwer, J., Tour, O., Laird, D. W., Sosinsky, G., and Ellisman, M. H., Multicolor and Electron Microscopic Imaging of Connexin Trafficking. *Science* **2002,** 296, 503.

Goutev, N. and Matsuura, H., Hydrogen Bonding in Chloroform Solutions of Ethylenedioxy Ethers. Spectroscopic Evidence of Bifurcated Hydrogen Bonds. *J. Phys. Chem. A.* **2001,** 105, 4741.

Graf, L., and Li, C. H., Human Somatotropin: Covalent Recconstitution of Two Polypeptide Contiguous Fragments with Thrombin. *Proc. Nat. Acad. Sci. USA* **1981**, 78, 6135.

Hamada, D., Kuroda, Y., Mikio, K., Saburo, A., Yoshimura, T., Goto, Y., Role of Heme Axial Ligands in the Conformational Stability of the Native and Molten Globule States of Horse Cytochrome *c*. *J. Mol. Biol.*, **1996**, *256*, 172.

Hartshorn, K. L., and Moore, G. R., A Denaturation-Induced Proton-Uptake Study of Horse Ferricytochrome *c*. *Biochem. J.*, **1989**, *258*, 595.

Homandberg, G. A., and Laskowski, M., Enzymatic Resynthesis of the Hydrolysed Peptide Bond(s) in Ribonuclease S. *Biochemistry* **1979**, 18, 586.

Horecker, B. L., Kornberg, A., The Cytochrome *c*-Cyanide Complex. *J. Biol. Chem.*, **1946**, *165*, 11.

Kim, S., and Barry, B., Reaction-Induced FT-IR Spectroscopic Studies of Biological Energy Conversion in Oxygenic Photosynthesis and Transport. *J. Phys. Chem. B.* **2001**, 105, 4072.

Kumar, R., and Bhuyan, A., Two-State Folding of Horse Ferrocytochrome *c*: Analyses of Linear Free Energy Relationship, Chevron Curvature, and Stopped-Flow Burst Relaxation Kinetics. *Biochemistry*, **200**5, *44*, 3024.

Levinthal, C., Are There Pathways for Protein Folding? *J. Chim. Phys.*, **1968**, *65*, 44.

Liu, D., Magliery, T., Pastrnak, M. and Shultz, P. G., Engineering a tRNA and Aminoacyl-tRNA Synthetase for the Site-Specific Incorporation of Unnatural Amino Acids into Proteins *in vivo*. *Proc. Nat. Acad. Sci.* **1997**, 94, 10092.

Liu, C. F., and Tam, J. P., Peptide Segment Ligation Strategy Without Use of Protecting Groups. *Proc. Natl. Acad. Sci. USA* **1994**, 91, 6584.

Liu, G., Grygon, C. A., and Spiro, T. G., Ionic Strength Dependence of Cytochrome *c* Structure and Trp-59 H/D Exchange from Ultraviolet Resonance Raman Spectroscopy. *Biochemistry*, **1989**, *28*, 5046.

Louie, G. V., Brayer, G. D., High-Resolution Refinement of Yeast Iso-1 Cytochrome *c* and Comparisons with Other Eukaryotic Cytochromes *c*. *J. Mol. Biol.*, **1990**, *214*, 527.

Lyubovitsky, J. G., Gray, H. B., and Winkler, J. R., Mapping the Cytochrome *c* Folding Landscape. *J. Am. Chem. Soc.*, **2002**, *124*, 5481.

Maity, H., Maity, M., Krishna, M., Mayne, L., and Englander, W. S., Protein Folding: A Stepwise Assembly of Foldon Units. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 4741.

Margoliash, E., Primary Structure and Evolution of Cytochrome *c*. *Proc. Natl. Acad. Sci. USA*, **1963**, *50*, 672.

Merrifield, R. B., *Peptides: Synthesis, Structures and Applications*. Ed.; Gutte, B. Academic Press, San Diego, California, 1995.

Meyer, R. and Gunthard, H., Internal Rotation and Vibration in $CH_2=CCl-CH_2D^*$. *J. Chem. Phys.*, **1969**, *50*, 353.

Miyazawa, T., Perturbation Treatment of the Characteristic Vibrations of Polypeptide Chains in Various Configurations. *J. Chem. Phys.* **1960,** 32, 1647.

Montgomery, D. L., Hall, B. D., Gillam, S., and Smith, M., Identification and Isolation of Yeast Cytochrome *c* Gene. *Cell*, **1978**, *14*, 673.

Muir, T. W., Dawson, P. E., and Kent, S. B., Protein Synthesis by Chemical Ligation of Unprotected Peptides in Aqueous Solution. *Methods in Enzymology*, **1997**, *289*, 266.

Myer, Y. P., Conformation of Cytochromes. III. Effect of Urea, Temperature, Extrinsic Ligands, and pH Variation on the Conformation of Horse Heart Ferricytchrome *c*. *Biochemistry*, **1968**, *7*, 765.

Nall, B., Structural Intermediates in Folding of Yeast Iso-2 Cytochrome *c*. *Biochemistry*, **1983**, *22*, 1423.

Nozaki, M., Mizushima, H., Horio, T., and Okunuki, K., Studies on Cytochrome *c* .2. Further Study on Proteinase Digestion of Baker Yeast Cytochrome *c*. *J. Biochem.*, **1958**, *256*, 673.

Osterhout, J. J., and Nall, B. T., Slow Refolding Kinetics in Yeast Iso-2 Cytochrome *c*. *Biochemsity*, **1985**, *24*, 7999.

Pabit, S. A., Roder, H., and Hagen, S. J., Internal Friction Controls the Speed of Protien Folding from a Compact Configuration. *Biochemistry*, **2004**, *43*, 12532.

Pasher, T., Chesick, J. P., Winkler, J. R., and Gray, H. B., Protein Folding Triggered by Electron Transfer. *Science* **1996**, 271, 1558.

Rong, Z. and Kjaergaard, H. G., Internal Methyl Rotation in the CH Stretching Overtone Spetra of ortho-, meta-, and para-Xylene. *J. Phys. Chem. A*. **2002,** 106, 6242.

Russell, B. S., and Bren, K. L., Denaturant Dependence of Equilibrium Unfolding Intermediates and Denatured State Structure of Horse Ferricytochrome *c*. *J. Biol. Inorg. Chem.*, **2002**, *7*, 909.

Russell, B. S., Melenkivitz, R., and Bren, K. L., NMR Investigation of Ferricytochrome *c* Unfolding: Detection of an Equilibrium Unfolding Intermediate and Residual Structure in the Denatured State. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 8312.

Saxon, E., and Bertozzi, C. R., Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000**, 287, 2007.

Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Jr., Wulff, M., and Anfinrud, P. A., Watching a protein as it functions with 150 ps time-resolved X-ray crystallography. *Science* **2003,** 300, 1944.

Segel, D., Fink, A. L., Hodgson, K. O., and Doniach, S., Protein Denaturation: A Small-Angle X-ray Scattering Study of the Ensemble of Unfolded States of Cytochrome *c*. *Biochemistry*, **1998**, *37*, 12443.

Shastry, M., Sauder, J. M., and Roder, H., Kinetic and Structural Analysis of Submillisecond Folding Events in Cytochrome *c*. *Acc. Chem. Res.*, **1998**, *31*, 717.

Silbert, E. L., Reinhardt, W. P., and Hynes, J. T., Intramolecular Vibrational Relaxation and Spectra of CH and CD Overtones in Benzene and Perdeuterobenzene. *J. Phys. Chem.* **1984,** 81, 1115.

Spiro, T. G., and Strekas, T. C., Resonance Raman Spectra of Hemoglobin and Cytochrome *c*: Inverse Polarization and Vibronic Scattering. *Proc. Natl. Acad. Sci. USA* **1972**, 69, 2622

Sutherland, J. C., and Klein, M. P., Magnetic Circular Dichroism of Cytochrome *c*. *J. Chem. Phys*. **1972**, 57, 76.

Suzuki, S., Shimanouchi, T., and Masamichi, T., Normal Vibrations of Glycine and Deuterated Glycine Molecules. *Spectrochimica Acta.* **1963,** 19, 1195.

Tam, J. P., and Yu, Q., Methionine Ligation Strategy in the Biomimetic Synthesis of Parathyroid Hormones. *Biopolymers* **1998**, 46, 319.

Tokano, T., and Dickerson, R. E., Redox Conformational Changes in Refined Tuna Cytochrome *c*. *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 6371.

Torres, J., and Arkin, I. T., C-Deuterated Alanine: A New Label to Study Membrane Protein Structure Using Site-Specific Infrared Dichroism. *Biophys. J*. **2002**, 82, 1068.

Trewhella, J., Carlson, V., Curtis, E. H., and Heidorn, D. B., Differences in the Solution Structures of Oxidized and Reduced Cytochrome *c* Measured by Small-Angle X-ray Scattering. *Biochemistry*, **1988**, *27*, 1121.

Tsong, T. Y., The Trp-59 Fluorescence of Ferricytochrome *c* as a Sensitive Measure of the Over-All Protein Conformation. *J. Biol. Chem*., **1974**, *249*, 1988.

Voet, D., Voet, J., *Biochemistry, 2ⁿᵈ Edition*. John Wiley & Sons, Inc, Somerset, New Jersey, 1995.

Wallace, C. J., and Offord, R. E., The Semisynthesis of Fragments Corresponding to Residues 66-104 of Horse Heart Cytochrome *c*. *Biochem. J.* **1979**, 179, 169.

Zanni, M. T., Ge, N. H., Kim, Y. S., and Hochstrasser, R. M., 2D IR Spectroscopy Can Be Designed to Eliminate the Diagonal Peaks and Exhibit Only the Cross Peaks Needed for Structure Determination. *Proc. Nat. Acad. Sci. USA* **2001**, 98, 11265.

Zhang, L., and Tam, J. P., Orthogonal Coupling of Unprotected Peptide Segments through Histidyl Amino Terminus. *Tetrahedron Lett.* **1997**, 38, 3.

CHAPTER 2. METHODS

2.1 Methods

As with any new technique, before spectral changes can be assessed and tied to biological significance, everything involved in sample preparation, acquisition and analysis of the data had to be rigorously worked out. Sample preparation, data acquisition and data analysis all followed an evolution process closely aligned with the need to get accurate data on even the smallest of spectral peaks. A key step in the sample preparation, since the concentrations of the protein are high, was the ability to control the pH of the protein itself, as it was discovered that the folding transitions are extremely pH sensitive. Once the protein samples could be generated, the next thing to evaluate was which amino acids yielded the most desirable spectral features. After well-behaved residues were targeted, a rigorous assignment of the spectral peaks was in order. Finally, once the spectra of specifically deuterated residues in the protein were acquired, further development on the careful baseline correcting and fitting of the data was carried out.

2.1.1   Sample Preparation and Measurements

The semi-synthesis of cytochrome *c* was carried out in three steps, as shown in Figure 2.1. The first step involved cleaving the purchased protein with cyanogen bromide, which cleaves at methionine residues. Since cytochrome *c* has two methionine residues (Met65 and Met80), the cleavage product was HPLC purified to give only residues 1-65 with the heme, which is covalently attached. Cleavage with
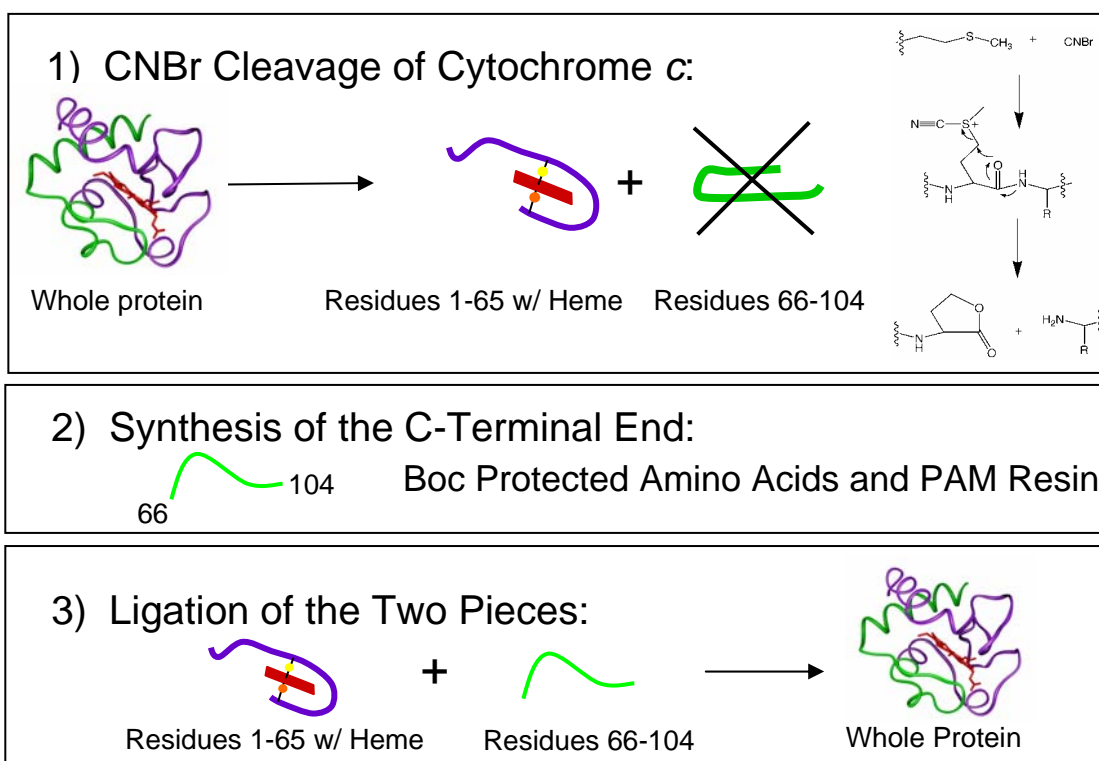
Figure 2.1. The semi-synthesis of horse heart cytochrome *c* is carried out in three steps. The first step involves cleavage of the whole protein with cyanogen bromide, which creates an 'activated' homoserine lactone at residue 65. The second step consists of synthesizing the C-terminal 39 residues, incorporating deuterated amino acids at any desired site throughout the peptide. In the last step, the synthetic peptide is ligated to the cyanogen bromide fragment in an autocatalytic ligation reaction.

cyanogen bromide creates an 'activated' homoserine lactone at residue 65, which then undergoes an autocatalytic ligation reaction in the third step. The second step is the synthesis of residues 66-104, the C-terminal half of the protein. And the last step is the ligation of the synthetic fragment to the cyanogen bromide cleavage product, which is carried out under non-denaturing conditions.

Cyanogen Bromide Cleavage. Cyanogen bromide cleavage of horse heart cytochrome *c* was carried out by dissolving 200 mg of cytochrome *c* in 8 mL of 30% (aqueous) formic acid (both purchased from Sigma/Aldrich) (Wallace, 1992). This solution was degassed for approximately 10 minutes then 400 mg of solid cyanogen bromide (purchased from Sigma/Aldrich) was added. The flask was covered in foil and allowed to stir gently under argon for three hours before gel-exclusion chromatography on Sephadox G25 (Amersham Biotech) equilibrated with 10% (aqueous) formic acid. The 1-65 fragment with heme was purified on reverse phase HPLC using a preparative C-18 Vydac 70 mm column, a Waters 4000 HPLC and a Gilson, model 116, detector. The crude mixture was loaded onto the preparative column and the retained material eluted with a 25-50% water-acetonitrille gradient in 0.1% TFA over 90 minutes at a flow rate of 30 ml/min.

Synthesis of the C-terminal Fragment. Synthesis of the 30-residue C-terminus of cytochrome *c* (residues 66-104) was carried out according to Schnölzer (Schnölzer, 1992). The synthesis was done on a 0.4 mmol scale with (*t*Boc)-Glu-PAM resin (NeoMPS) and amino acids (Midwest Biotech) were added in a six-fold excess to the resin at each step with a yield of at least 99.5% per step as monitored by quantifying

the coupling reaction using ninhydrin tests (Sarin, 1981). Deuterated amino acids (Cambridge Isotopes, Inc.) were added in a two-fold excess to the resin and coupling efficiency was generally above 99%. Triflouroacetic acid (TFA) was purchased glass distilled from Halocarbon Inc., dimethylformamide (DMF) was purchased from Burdick and Jackson, and diisopropylethylamine (DIEA) from Sigma/Aldrich. The resin was cleaved using 90% hydrogen fluoride (HF) in 10% anisole at 0 ºC for 60 minutes. The peptide and resin was then precipitated in cold diethyl ether and dissolved in a 30% acetonitrille-water mixture with 0.1% TFA. Purification was done similar to the 1-65H fragment described above.

Fragment Condensation. Fragment condensation was performed according to Wallace and Lewis in 50 mM sodium phosphate buffer at pH 7. First 1.8 mg/ml of synthetic peptide was dissolved in 50 mM monobasic sodium phosphate, then 50 mM of dibasic sodium phosphate was added to reach a pH of 7. Then 3.2 mg/ml of 1-65H fragment was added to reach a final concentration of complex of 0.4 mM. Both monobasic and dibasic sodium phosphate were thoroughly degassed prior to use and a solution of 1 mg/ml sodium dithionite (Sigma/Aldrich) was added to the final solution to reduce the complex. The solution was kept in a sealed vial (with septa) under argon for 24 hours before gel-exclusion chromatography using Sephadex G50 equilibrated with 10% aqueous formic acid. If necessary, purification of the final product was also carried out using HPLC similar to that described above for the 1-65H fragment.

All fragments and final product were characterized by HPLC, ESMS and UV-Vis spectra as shown in Figures 2.2 and 2.3. The measurements in which the redox state

of the heme was changed were characterized using the change in Q bands in the UV-Vis spectra upon reduction. Measurements involving the unfolding of the protein were also characterized using UV-Vis spectroscopy; the Soret band (408 nm), Q band (528 nm) and charge-transfer band (695 nm) were monitored as a function of GnHCl or urea (Eaton, 1967). Folding samples were tested for refolding to ensure that no aggregation was present. Folding titration curves using UV-Vis bands comparing synthetic and purchased cytochrome *c* were carried out to ensure the purity of the synthetic samples (data not shown).

Preparation of Samples for IR Measurements. Oxidized samples were prepared by dissolving the condensation product in 100 mM sodium acetate buffer, pH 5, and 2 ml of a saturated solution of bis(dipicolinato)cobaltate(III) (Mauk, 1979). The mixture was then passed through a Sephadex G15 column equilibrated with 10 mM sodium acetate buffer, pH 5, washed several times to make sure the protein was equilibrated in the buffer at pH 5 and lyophilized. Samples were stored in a -80 ºC freezer until just prior to measurement. Approximately 2 hours before measurement, samples were allowed to defrost at room temperature and oxidized samples were dissolved into 100 mM sodium acetate buffer, pH 5. Reduced samples were prepared by taking the oxidized lyophilized protein and adding a degassed mixture of 100 mM sodium acetate buffer, pH 5 and 100 mM ascorbic acid (also at pH 5). Samples for folding measurements were prepared by adding bis(dipicolinato)cobaltate(III) to the

CNBr Cleavage:

Crude HPLC   Purified HPLC   Purified ES-MS

MW = 7,755 ± 1.5 g/mole

Synthesis of Residues 66 - 104:

Crude HPLC   Purified HPLC   Purified ES - MS

MW = 4576 ± 1 g/mole

Ligation:

Purified HPLC   Purified ES-MS

MW = 12,332 ± 2 g/mole
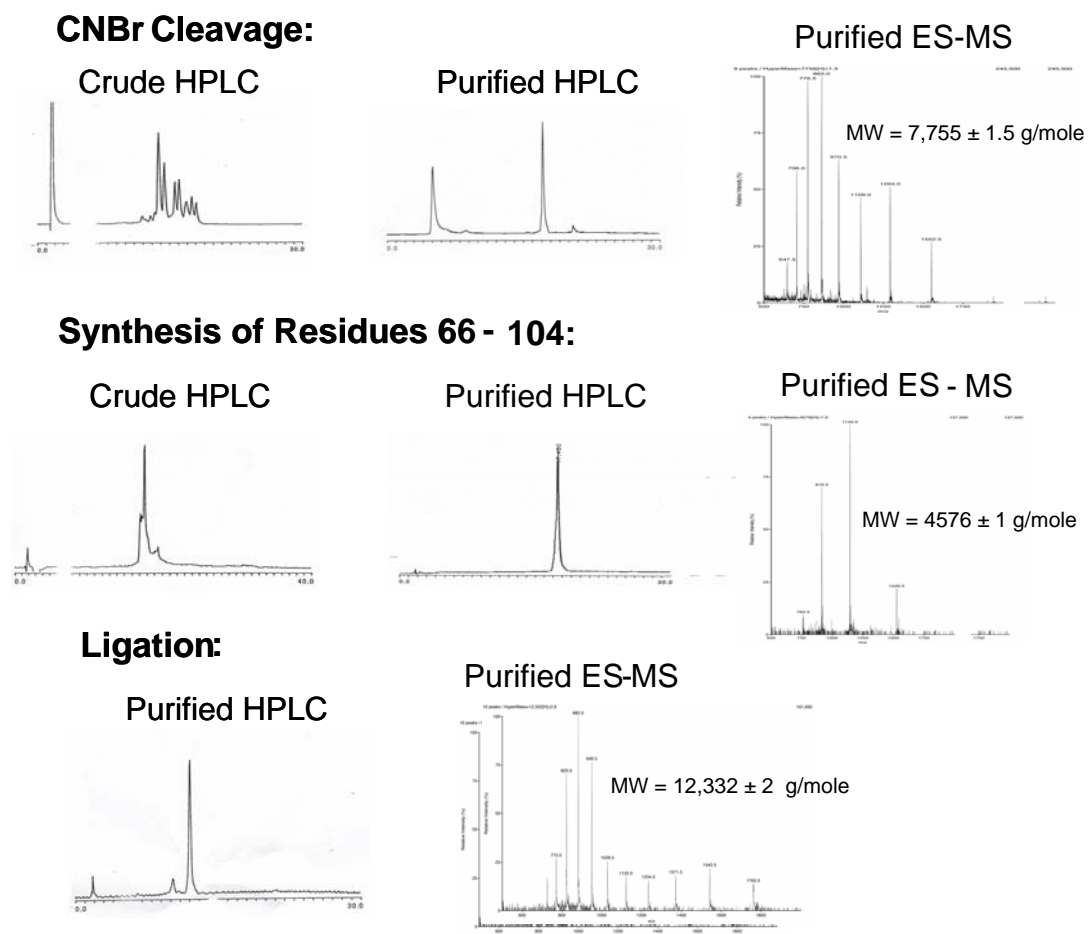
Figure 2.2. Characterization at each step of the semi-synthesis of cytochrome *c* was carried out by HPLC and ES-MS.

ES-MS Spectra of All Final Products

Proteo Cytochrome *c*

MW = 12,329 ± 2 g/mole

d_3-Cytochrome *c*

MW = 12,332 ± 2 g/mole
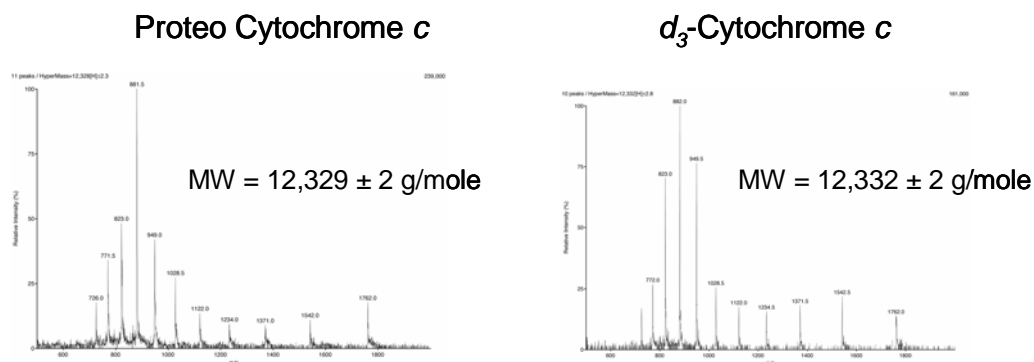
Characterization of Final Product with UV-Vis Spectroscopy

- - - Synthetic Cytochrome *c*
— Bottle Cytochrome *c*

— Synthetic Cyt *c*, Oxidized
— - — Synthetic Cyt *c*, Unfolded
- - - Synthetic Cyt *c*, Reduced
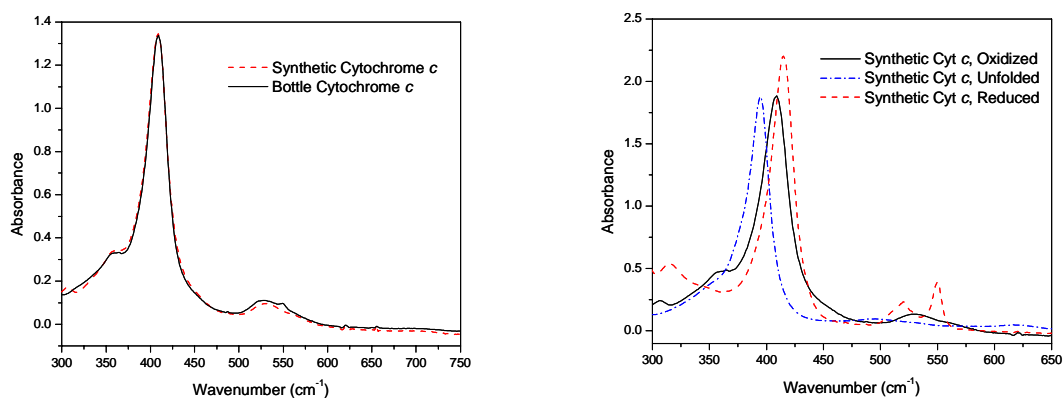
Figure 2.3. Further characterization of the final synthetic cytochrome *c* product was carried out by ES-MS and UV-vis spectroscopy. All deuterated species were checked and ensured that the final product had the appropriate mass. The UV-Vis spectra of synthetic cytochrome *c* was compared to that of the purchased protein and various forms of the synthetic protein were also checked for the appropriate spectral features.

condensation product and passing the protein over a Sephadox G15 column equilibrated with either 10 mM sodium acetate buffer, pH 6.2 (for folding measurements in GnHCl) or 5 mM potassium phosphate buffer, pH 7 (for folding measurements in Urea). Samples were then washed thoroughly, equilibrated at the appropriate pH and then lyophilized. Careful attention was paid to the final volume of protein lyophilized, which was never over 40 µL per sample, to keep the salt concentration in the final samples low. Guanidine hydrochloride (GnHCl) solutions were made by dissolving the proper amount of solid GnHCl into 5 mL of 200 mM sodium acetate buffer, pH 6.2. After dissolving the GnHCl, water was then added to approximately the 10 mL mark and the pH of each solution thoroughly checked. Solutions of higher amounts of GnHCl had a lower pH than 6.2 and a small amount of 0.1 M NaOH was added to these solutions until the pH was 6.2. Urea solutions were made by adding solid urea to 100 mM of 5 ml of potassium phosphate buffer, pH 7 and 0.4 M KCl. After the solid urea was dissolved, water was added to the 10 ml mark. Since urea increases the pH of the solution, different amounts of monobasic and dibasic potassium phosphate buffer (along with 0.4 M KCl) were mixed to account for the change in pH caused by urea. The final pH of all solutions were thoroughly checked and it was insured that the pH did not change upon addition of the lyophilized protein. The concentration of GnHCl and urea solutions were all individually checked using index of refraction measurements and the concentration calculated from these measurements (Creighton, 1989). After approximately 16 µl of

buffer solution was added to the lyophilized protein, samples were allowed to equilibrate at least 10 minutes prior to measurement.

FTIR Measurements. All protein samples were measured using a Bruker Equinox 55 FTIR spectrometer equipped with a liquid nitrogen cooled MCT detector, KBr beam splitter and continuously flushed with dry nitrogen. Spectra were collected at both 2 and 4 cm^{-1} resolution and constructed from 8192 scans. The Blackman-Harris 3-term apodization function was used, as well as a 16 kHz low pass filter and zero-filling of 16. The aperture setting was adjusted to have the maximum light with no sample in the sample holder and was usually a value close to 2000 μm. The FTIR transmission cell consisted of CaF_2 disks 32 mm in diameter with a Teflon spacer 50 μm thick.

2.1.2   Assessment of Deuterated Amino Acids in Cytochrome *c*

Before the site-selective incorporation of deuterated residues in the C-terminal half of cytochrome *c* was carried out, an assessment of which deuterated amino acids were best suited for these experiments was carried out. Several criteria concerning the amino acids had to be met. First, the deuterated amino acid had to be accessible either commercially or easily synthesized on site. Second, less rather than more deuterated bonds per amino acid were preferable so that spectral peaks can be easily assigned. Third, non-polar amino acids were more desirable than polar ones, since non-polar residues are more likely to point toward the inner core of the protein, thus reporting on the protein environment. Not only do polar residues often point into solution and not
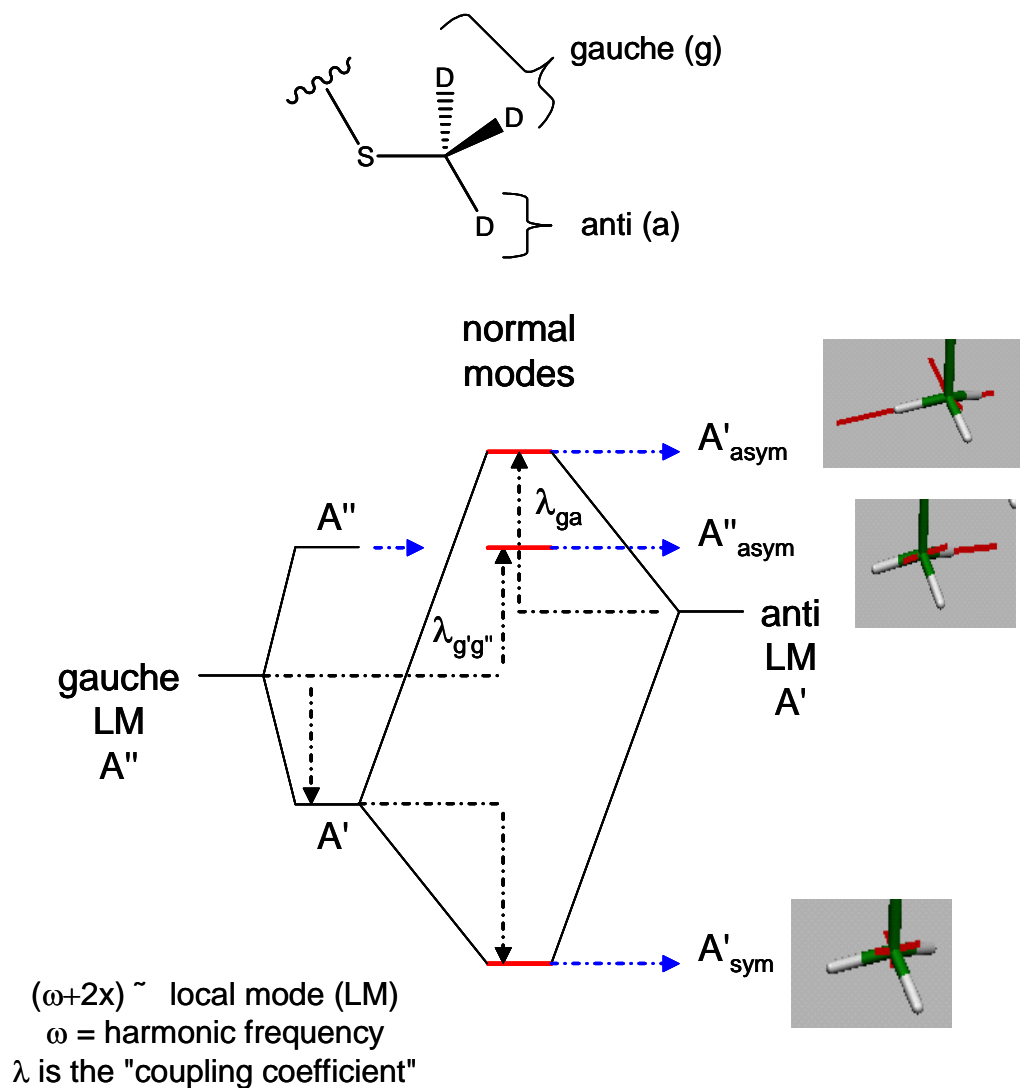
Figure 2.4. Depiction of how one carbon-deuterium bond in the anti position is coupled to one carbon-deuterium bond in the gauche position for a deuterated methyl group. This coupling gives rise to two asymmetric stretches of A′ and A″ symmetry and one symmetric stretch of A′ symmetry (Halonen, 1997).

report on the protein of interest, but they also present the problem of interacting more with the solvent molecules, such as H-bonding or protonation/deprotonation by solvent molecules. Fourth, the amino acid of choice has to have desirable spectral features, a high absorption coefficient and preferably a small number of overlapping peaks. Lastly, more desirable amino acids will be in biologically significant locations in the protein, as demonstrated through other studies, and therefore more likely to yield interesting information about the protein. The deuterated amino acids with the best properties aforementioned are shown in Figure 2.4. The high signal to noise ratios for methionine, leucine, and alanine compared to glycine and glutamic acid resulted in those residues being the focus in Chapters 3, 4, and 5. Leucine and alanine are aliphatic residues likely to be pointing toward the inner core of the protein and were expected to be good reporter molecules for redox and folding studies. Their abundance throughout the C-terminal half of the protein also made them attractive residues of study. Methionine was the original residue of choice and is of particular biological significance since it is one of the protein-based heme ligands.

2.1.3 Spectral Assignments

The infrared spectra of deuterated amino acids were assigned using three different criteria: a literature review of similar compounds and their assignments, the spectra of the free amino acid in buffer, and *ab initio* deuterium frequency calculations of the free amino acid of interest. Figure 2.6 shows the free amino acid spectra, a comparison between calculated and observed spectral frequencies, and a figure
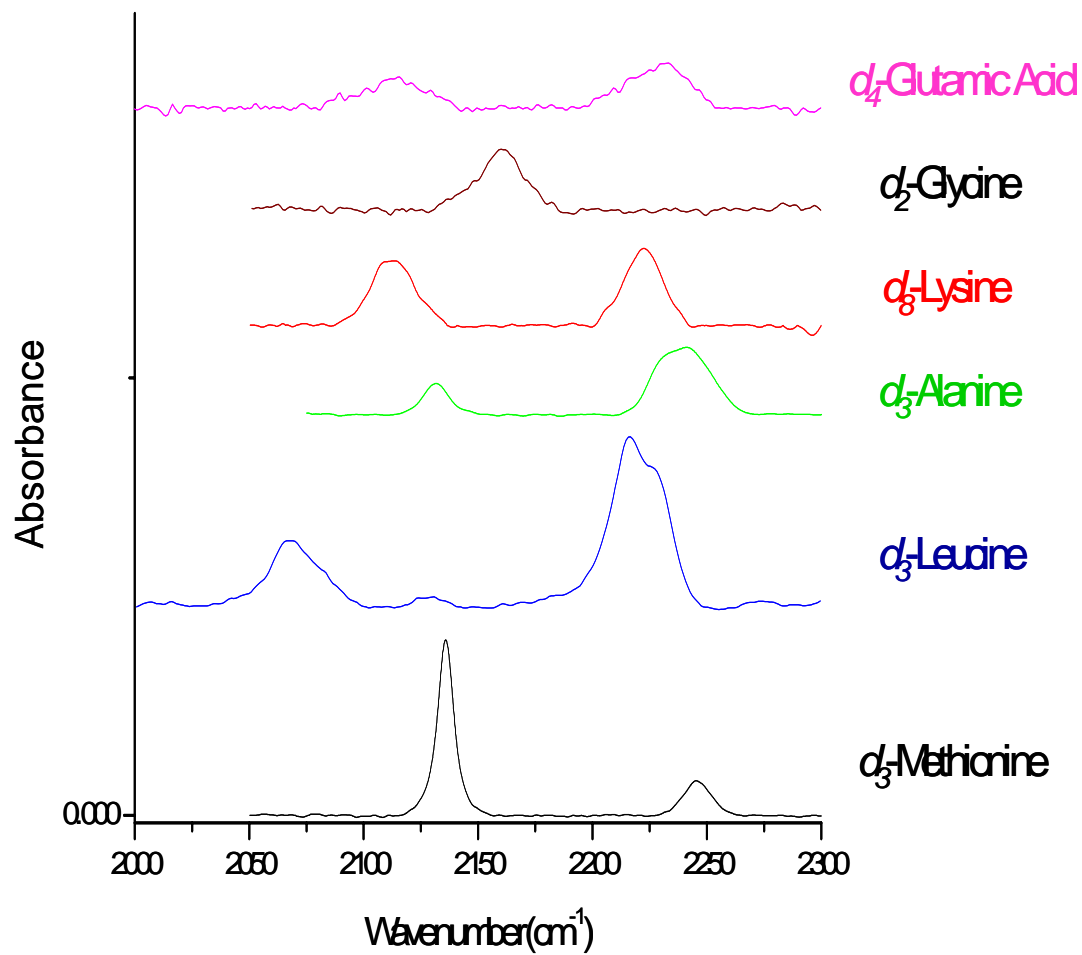
Figure 2.5. FTIR Spectra of 50 mM deuterated free amino acids in 100 mM sodium acetate buffer, pH 5. These six residues have the most suitable spectral features and four of the six residues were used in redox and folding measurements discussed in Chapters 3, 4, and 5.

d₃-Methionine

	Calculated	Observed
Symmetric Stretch	2084 cm⁻¹	2136 cm⁻¹
Asymmetric Stretch	2234 cm⁻¹	2246 cm⁻¹
Asymmetric Stretch	2239 cm⁻¹	2262 cm⁻¹

d₃-Leucine

	Calculated	Observed
Symmetric Stretch	2062 cm⁻¹	2070 cm⁻¹
Asymmetric Stretch	2174 cm⁻¹	2218 cm⁻¹
Asymmetric Stretch	2182 cm⁻¹	2233 cm⁻¹

d₃-Alanine

	Calculated	Observed
Symmetric Stretch	2072 cm⁻¹	2079 cm⁻¹
Asymmetric Stretch	2190 cm⁻¹	2243 cm⁻¹
Asymmetric Stretch	2207 cm⁻¹	2258 cm⁻¹

*Scaled, harmonic frequencies at the HF/LANL2DZ Level of Theory Using Gaussian03

Figure 2.6. Spectral assignments of all deuterated residues used in subsequent measurements. A table of calculated (using the LANL2DZ basis set) and observed values is on the far left. The middle panel contains the FTIR spectra of the deuterated free amino acid in 100 mM sodium acetate buffer, pH 5. And on the right side is a diagram showing the sites of deuteration for each residue. For d_3-leucine, alanine and methionine, there are two high frequency asymmetric stretches and one low frequency symmetric stretch. Although the spectral assignments for the symmetric stretches are similar in the protein to that of buffered solution, the assignments for the asymmetric stretches differ.

highlighting the sites of deuteration for each amino acid. A literature survey of similar compounds, such as acetone, demonstrates that a deuterated methyl groups will show two asymmetric stretches of almost degenerate frequency and one symmetric stretch of lower frequency (Dellepiane, 1966) (see Figure 2.5). Scaled harmonic frequencies were calculated at the HF/LANL2DZ level of theory (scaling factor 0.9) for C^{ε}-d_3-methionine and two asymmetric stretches at 2278 and 2290 cm^{-1} and a symmetric stretch at 2144 cm^{-1} were predicted. The spectra of the free amino acid contained one broad peak at 2245 cm^{-1} and a narrower peak at 2136 cm^{-1}. The C^{ε}-d_3-Met80 residue in cytochrome c had similar spectral features as the free amino acid, with one broad peak at 2254 cm^{-1} and a narrower peak at 2128 cm^{-1}. The high frequency broad peak is therefore assigned as the overlapping asymmetric stretches and the lower frequency narrow peak as the symmetric stretch. For C^{β}-d_3-alanine, the two asymmetric stretches were predicted to be 2270 and 2293 cm^{-1} and the symmetric stretch 2155 cm^{-1}. The free amino acid spectra showed two high frequency peaks at 2243 and 2258 cm^{-1} and one lower frequency peak at 2079 cm^{-1}. The spectra of alanine in the protein were similar to that of the free amino acid and depend on the position in the protein. Therefore, similar to methionine, the two high frequency peaks are assigned as the asymmetric stretches and the low frequency peak as the symmetric stretch. Unfortunately, due to the low absorption coefficient of alanine at 10 mM, the symmetric stretch could not be observed in the protein.

The spectral assignments for leucine are similar to alanine and methionine with the exception that C^{δ}-d_3-leucine was only available as a 1:1 mixture of $C^{\delta 1}$- and $C^{\delta 2}$-

isotopomers. It was then predicted that there should be four high frequency asymmetric stretches, 2174 and 2175, 2178 and 2182, and two low frequency symmetric stretches, 2061 and 2062. The spectra of the free amino acid could be accurately fit to only three Gaussians, two at high frequency and one at lower frequency with values of 2233, 2218 and 2070 cm^{-1} respectively. Since the observed spectra of the free amino acid matches the calculated values as closely as the other methyl-deuterated residues, and it is conceivable that the two methyl groups in solvent would not be differentiated, the two high frequency peaks were assigned as the two asymmetric stretches of overlapping methyl groups. However, d_3-leucine incorporated in the protein yielded different spectral features. Two clearly resolved high frequency peaks are present in all incorporated leucine residues and, similar to solvent, only one low frequency peak was observed. A few key observations have led to different spectral assignments for the leucine residues in the protein versus the free amino acid in buffer. The spectral features for the high frequency region for both d_3-Met80 and d_3-Ala101 in the protein show only one peak resulting the severe overlapping of the asymmetric stretches. This suggests that, in general, for deuterated methyl groups within the protein, the two asymmetric stretches cannot be differentiated. Second, the two high frequency bands of a leucine residue within the protein seem to behave independently, both in the redox and folding studies, i.e. one band will show changes while the other does not. Lastly, the ratio between the two peaks varies with the batch of d_3-leucine purchased. These observations led to the assignment of the two high frequency peaks of d_3-Leu68, Leu94 and Leu98 as the overlapping asymmetric

absorptions of the two diasteriotopic methyl groups. The single low frequency peak is assigned as the two overlapping symmetric stretches.

2.1.4   Spectral Manipulation: Baseline Correction and Fitting

The semi-synthesis of the site-selectively deuterated cytochrome *c* was carried out alongside that of the proteo, containing no deuterated species. Care is taken that the proteo species is treated identical to the deuterated one, the pH, lyophilized volume, concentration and final buffer solution are all monitored to be identical. Although it is possible to use purchased cytochrome *c* for background purposes, to alleviate any concerns caused by the mutation at residue 65 with the synthetic protein, proteo cytochrome *c* was always used as a background. With each FTIR measurement, a careful background subtraction is performed with the OPUS program where the proteo species is subtracted from the deuterated one. The subtraction is done using a subtraction factor close to unity that ensured flatness in non-saturating regions of the spectra.

The resulting peaks after proteo subtraction typically have absorbance values of 0.001 and a background absorbance of 0.5. Not only is the background absorbance high compared to the peaks of interest, but the baseline is usually not flat. To correct for the baseline and produce a spectrum with the peak of interest and a flat baseline, a smooth polynomial is subtracted from the proteo-subtracted data. This process is illustrated in Figure 2.7. First, a spectral window of at least three times the full-width-at-half-maximum on either side of the peak of interest is selected. Next, a sufficiently
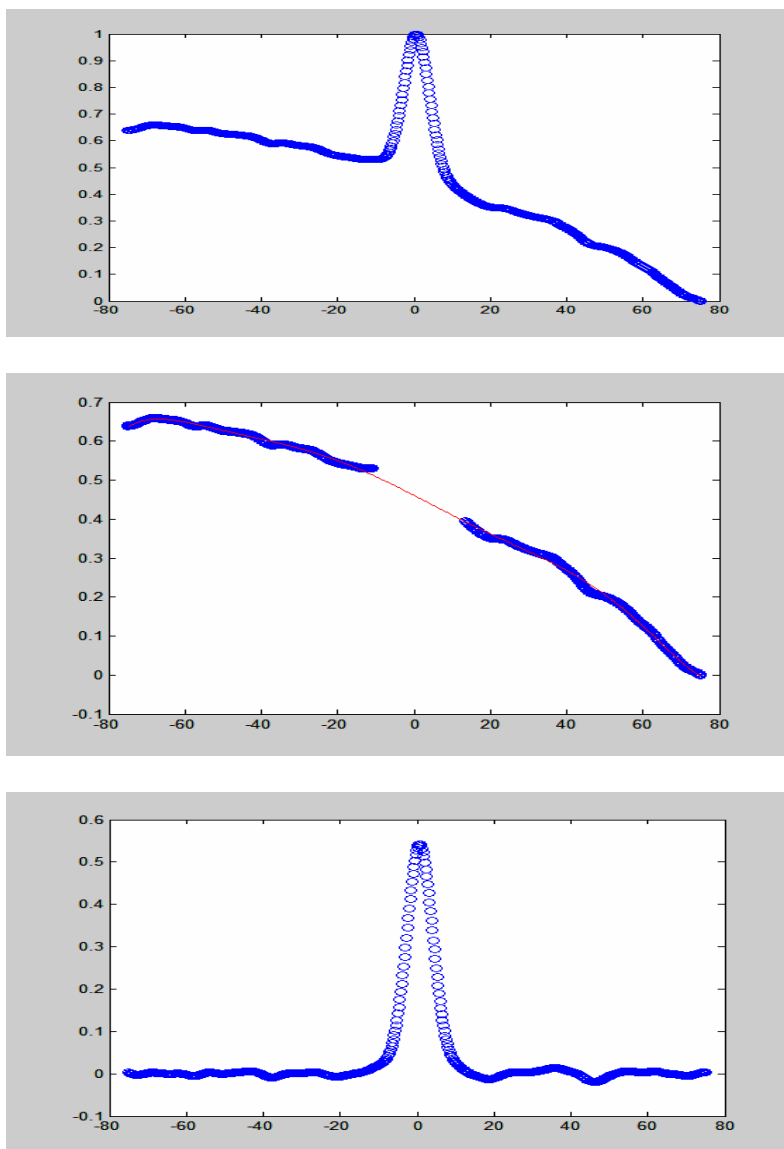
Figure 2.7. Baseline correction was carried out using Matlab by first taking a window of data (at least 100 cm^{-1}) and fitting the baseline (without the peaks of interest) to a 5th order polynomial. The part of the spectral window underneath the peaks of interest were approximated with the smooth polynomial as shown. The polynomial was then subtracted out, resulting in a relatively flat baseline around the peaks of interest. The peaks are then fit to Gaussians signifying the chemically meaningful species present.

narrow window is cut out so that the peak itself is not included in the polynomial fit. Then, a polynomial of either 5th or 7th order is fit to the baseline. This polynomial is then subtracted from the spectra yielding the peak of interest and a flat baseline. Care must be taken to subtract a smooth polynomial since any sudden changes in the fit polynomial may influence the spectral features of the baseline-corrected spectrum. The lowest order polynomial is used to best fit the baseline, to avoid changing the spectral features of the peak of interest. With a given set of data (such as a titration in urea or multiple samples under the same set of conditions), the same spectral window and polynomial order are used to ensure that a change in the baseline correction does not change the spectral features.

Once the proteo subtracted, baseline-corrected spectrum is obtained, the peaks of interest are fit to a sum of Gaussians representing chemically meaningful species. For example, based on the spectral assignments, if two asymmetric stretches are predicted, the peaks of interest will be fit to two Gaussians. Many times the expected number of peaks overlap and in this case, statistical analysis was used to distinguish how many Gaussians are needed to fit the data. Statistically significant Gaussians were determined using the 99% confidence limit of an F-test (see Figure 2.8). For example, the d_3-Ala101 data revealed two severely overlapping peaks and although two asymmetric stretches should be present, the question arose whether only one or two Gaussians were needed to fit the data. An average of three or four data files, taken under the same conditions was fit to both one and two Gaussians and an F-test number calculated. The F-test number is obtained by taking the ratio of the least squared sum

d₃-Met80, F-Tests for Asymmetric Stretches

F = 1.06
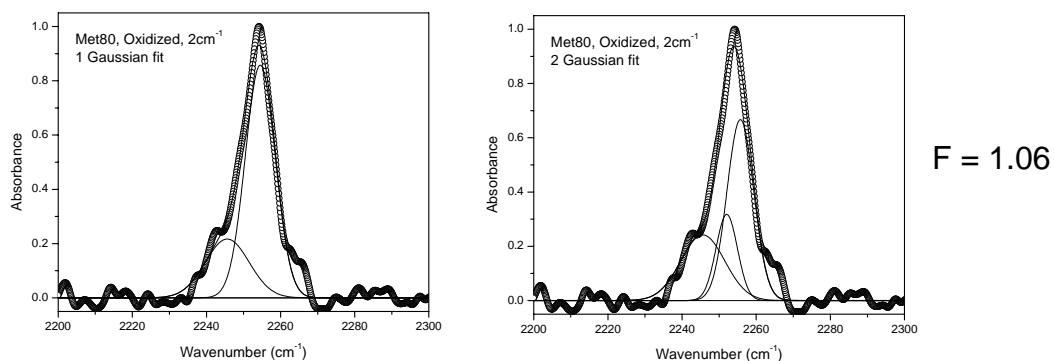
d₃-Ala101, F-Tests for Asymmetric Stretches

F = 1.03
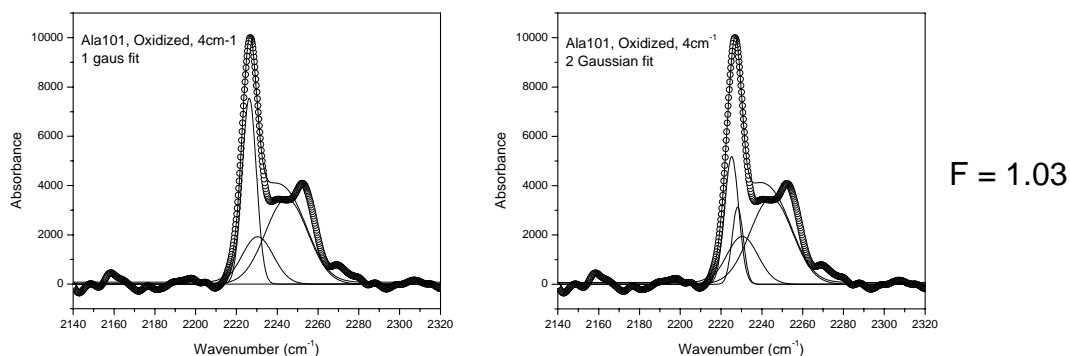
Figure 2.8. F-tests carried out on the asymmetric stretches of d_3-Ala101 to determine the number of Gaussians needed to fit the folded spectra. Since these spectra were taken at pH 5, where there is a mixture of folded and unfolded states, the F-tests were done using three and four Gaussian fits, two unfolded and one or two folded. The bold numbers to the right are the calculated F-test values. The cut-off value is 1.03 for 99% confidence limit. Although the calculated values are above the cut-off value for the d_3-Met80 data, the second Gaussian is small in amplitude and asymmetric stretches for the free amino acid in buffer is fit best by just one peak. The d_3-Ala101 data is closer to the cut-off value, with the oxidized data showing no significant improvement to the fit. In addition, the smaller linewidth and greater intensity of the peak relative to the free amino acid in buffer implies that the asymmetric stretches are significantly overlapped. It was therefore concluded that for both d_3-Met80 and d_3-Ala101, only one Gaussian was needed to fit the asymmetric stretches.

with more parameters (larger number of Gaussians) divided by the least squared sum with less parameters (smaller number of Gaussians). This number is then compared to a cut-off number for the 99% confidence limit (Snedecor, 1989).

The final step in analyzing the spectra is the actual fitting of the data. As described above, for a given residue in the protein in a given state, the number of Gaussians used to fit the data is determined from the *ab-initio* calculations and F-test results. For the data acquired in different redox states of the protein, all that was desired from the spectral data were the features of the *folded* spectra and the amount of unfolded present. Based on folding data presented in Chapters 3, 4 and 5, the spectral features of the free amino acid in solution are identical to those of the same amino acid in the protein when the protein is unfolded. Therefore, the unfolded spectral features were fixed to those of the free amino acid in solution and the parameters corresponding to the folded features not fixed.

Fitting the folding data was usually more involved. For a simple two-state model, Gaussians are fit to the data representing the folded and unfolded states. As many parameters in the Gaussians representing the folded and unfolded states are fixed as possible, such as relative amplitudes of multiple peaks representing a given state and frequencies of the folded state. The frequencies and linewidths of the unfolded state were generally not fixed due to salt effects. If inspection of contour plots of a series of unfolding spectra reveal the possibility of an intermediate, then another Gaussian was added to represent the intermediate. The placement of the third Gaussian is determined in a couple different ways. One way was to calculate the first and second

derivatives (first and second moments) of the raw spectra and plot these values versus denaturant concentration. These plots are carefully inspected for any abnormalities. Another option often invoked is to fit the data set in question with a two-state model. The spectral features of the folded and unfolded states are then plotted against denaturant and carefully inspected. Once a general spectral window is assigned to the intermediate Gaussian, the amplitude of that Gaussian is plotted versus denaturant concentration to verify that the species grows in and out as would be expected for an intermediate species. After spectral fitting is carried out, normalization steps ensure that the midpoints reported are accurate. The first normalization involves dividing the spectra in the folding titrations by the area underneath all peaks of interest and plotting the normalized spectra together (see Figure 2.9). If the spectra in the folding titrations indicate a two-state transition, an isosbestic point was defined for that window. The intensities from the Gaussians in the fit were normalized using the isosbestic point, which ensured that the amplitude at the isosbestic point was the same in all spectra. The individual peak amplitudes were then divided by the total amplitude of all species present. Further normalization was sometimes required in cases where the extinction coefficient of the folded and unfolded states differed and a correction factor (consisting of the ratio of extinction coefficients) was added to equalize the spectral intensities. As shown in Figure 2.9, fitting the raw spectra to Gaussians and carrying out various normalization steps result in a sigmoidal titration curve for two-state transitions. If more than two states were evident in the raw spectra a third Gaussian representing an intermediate species was added. Unfortunately, in a three-state model
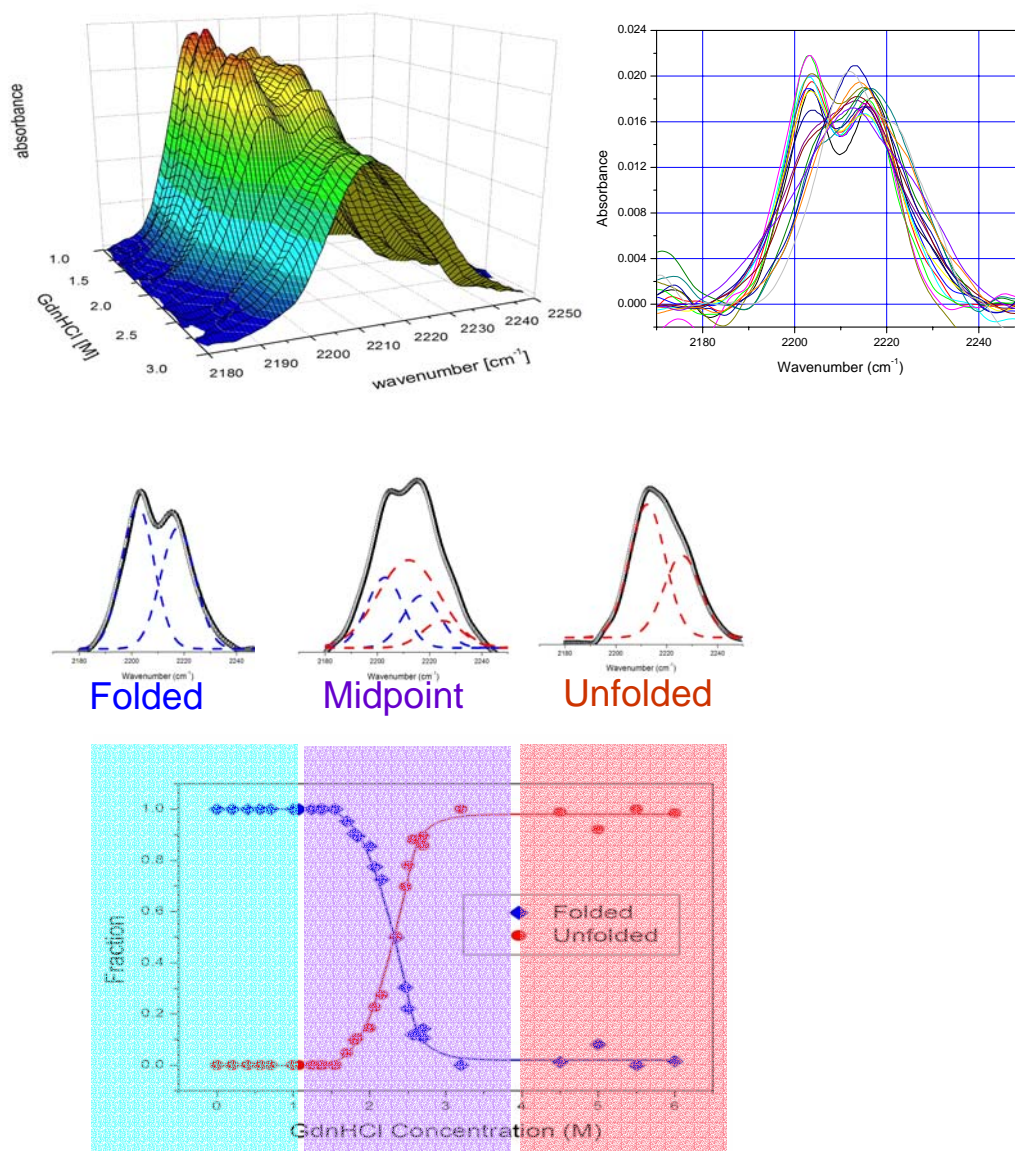
Figure 2.9. Spectral changes in d_3-Leu98 with increasing guanidine hydrochloride concentration. Upper left shows a 3-D contour plot of spectral changes with denaturant. Upper right is a plot of normalized spectra for all denaturant concentrations, revealing an isosbestic point at 2206 cm^{-1}. The bottom plots show how the data is fit to a two-state model, two Gaussians representing the folded state and the other two the unfolded state. As denaturant is added to the sample the amplitudes of the folded and unfolded states change, at low denaturant there is no unfolded protein and at high denaturant no folded protein present. The region in the middle contain both folded and unfolded states to varying degrees.

an isosbestic point often does not exist, so the spectra in this case could not be normalized so that the relative intensities were the same. After fitting the raw data to Gaussians, the intensity of each Gaussian was divided by the sum of all three species. If the extinction coefficients of the folded and unfolded states were different a further normalization step was carried out to equilibrate the two states.

2.2 Spectral Observables

The first step to realizing the goal of using carbon-deuterium bonds as probes for the local environment within a protein is to understand the factors giving rise to the spectral observables and how they change with local environment. Although several observables exist for a steady-state FTIR measurement, we chose to focus on the three that are most easily distinguished and interpreted, the frequency, lineshape and linewidth of the carbon-deuterium infrared bands. Changes in frequency reflect both specific through-space and through-bond interactions. Frequency changes can be instrumental in determining the local dielectric constant within the protein, pKa values of a specific residue, H-bonding characteristics, and the conformational state of the protein (Barth, 2002). Linewidths include both homogenous and inhomogeneous broadening, both of which are interesting for different reasons. Homogenous broadening may have implications for coupling mechanisms and energy pathways in proteins. Inhomogeneous broadening yields direct information on the local flexibility of the protein, which is currently difficult to acquire through other biophysical techniques.

2.2.1 Carbon-Deuterium Frequencies

Solvent studies on the free deuterated amino acids were the first to reveal that carbon-deuterium absorptions were not only possible to measure, but also sensitive to the local environment. As shown in Figure 2.10, the symmetric stretch of methionine and asymmetric stretches of leucine and alanine blue-shift as the polarity of the solvent is increased. This effect was also observed in the symmetric stretches of the same residues, but to a lesser degree, which is in agreement with literature studies showing the asymmetric stretch to be more sensitive to the environment than the symmetric stretch (Bellamy, 1958). Interestingly, when both hydrogen-bonding solvents and nonhydrogen-bonding solvents are used, the same trend is observed in the carbon-deuterium bonds of the free amino acids. A blue shift with increasing polarity is also well documented for various organic compounds that focused on more polarizable bonds such as S=O, C=O or C≡N groups. Although a hydrogen-bonding argument is often invoked to explain this effect in more polarizable bonds, the rare study involving non-hydrogen bonding solvents maintains that the effect is due to through-space, dipole interactions with the solvent (Bellamy, 1958). The solvent studies are an important reference to asses the local dielectric constant in a protein environment. Furthermore, these studies revealed that indeed carbon-deuterium bonds might be suitable for measuring protein folding, since the dielectric constant of the protein is different than that of buffered solution and it is expected that a residue would experience a difference in local polarity as the protein is unfolded.
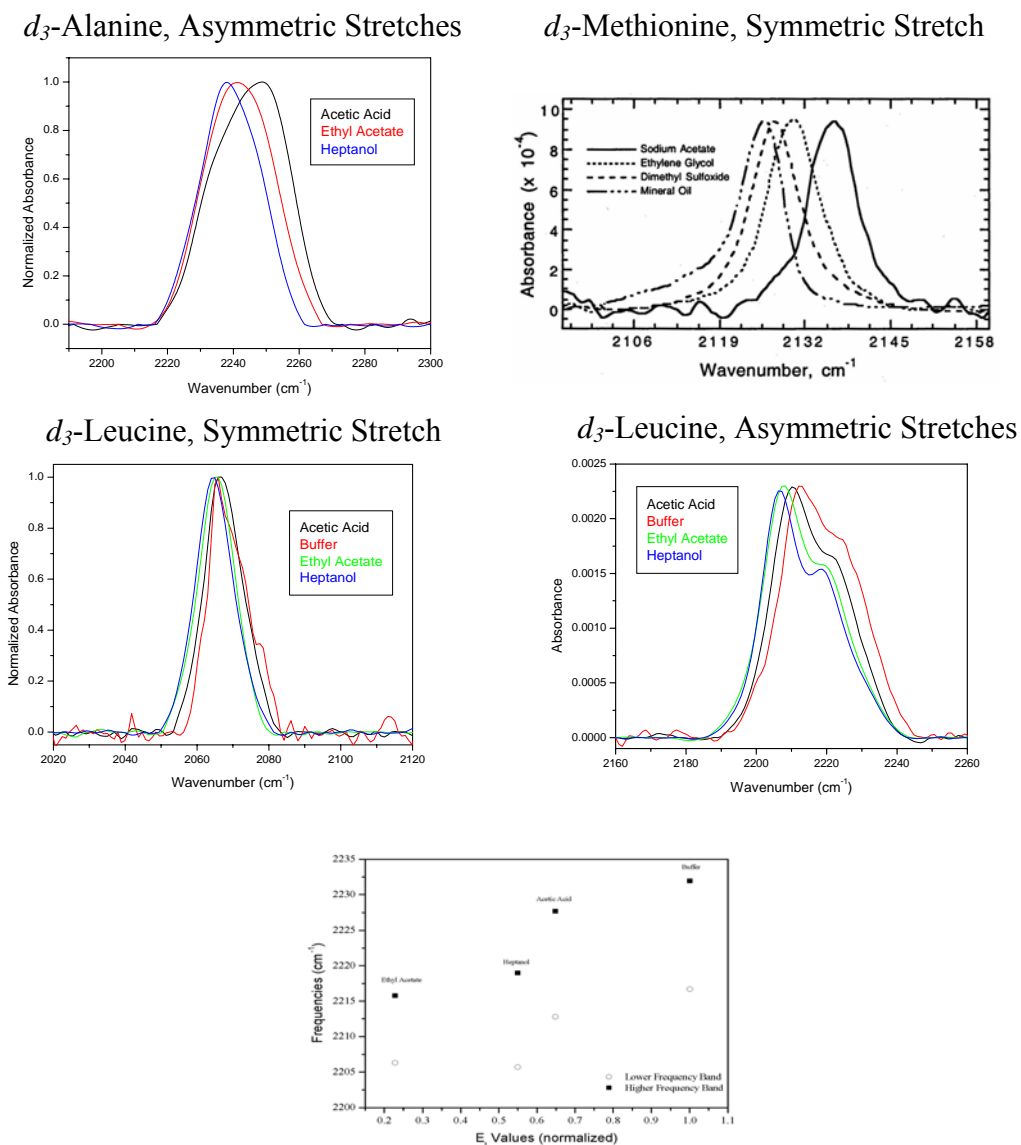
Figure 2.10. Top panel shows the symmetric stretch of d_3-methionine and asymmetric stretches of d_3-alanine in solvents of different polarity. The bottom panel shows the asymmetric and symmetric stretches of d_3-leucine in similar solvents. All residues shown shift to higher frequency as the dieletric constant of the solvent is raised. The largest observed shift is for d_3-methionine, most likely due to inductive effects from the nearby electronegative sulfer atom. For every residue examined, the asymmetric stretches have also been shown to be much more sensitive to the environment than the symmetric stretches, showing larger shifts in magnitude.

A full investigation of the frequency shift observed in the Met80 ligand of cytochrome *c* upon reduction of the heme was carried out in collaboration with Professor Kim Baldridge (unpublished results). The initial study reporting a frequency shift to the red of 2.5 cm^{-1} upon reduction of the heme cofactor suggested that it was a through-bond effect and thus amenable to quantum mechanical studies (Chin, 2001). The model system used in these calculations (Figure 2.11) consists of the porphyrin ring system, a dimethylsulfide group (representing the Met80 ligand) and an imidazole group (representing the His18 ligand). Geometry optimization and deuterium frequency calculations using the B3LYP/LANL2DZ basis set revealed that both changes in geometry and electron density give rise to the reduction in frequency of the reduced state of the protein. Figure 2.12 shows the changes in geometry induced upon reduction of the heme. The angle of the C-S-Fe bond is changed slightly, resulting in a more tetrahedral sp^3 character bond in the reduced state. The oxidized state has more sp^2 character to the bonds attached to the carbon atom and thus more s character, which equates to a slightly stronger bond, as reflected in the carbon-deuterium frequency. A dramatic difference in the occupation of antibonding orbitals of the carbon-deuterium bonds in two oxidation states is shown in the electron density plots of Figure 2.13. The reduced state has a significant amount of electron density in antibonding orbitals, whereas the oxidized state has very little, also implying a stronger bond in the oxidized state. And indeed, the deuterium frequency calculations carried out on the d_6-dimethyl sulfide species show a higher frequency in the oxidized state by 9 cm^{-1} when averaging all the C-D bonds on the dimethyl sulfide.
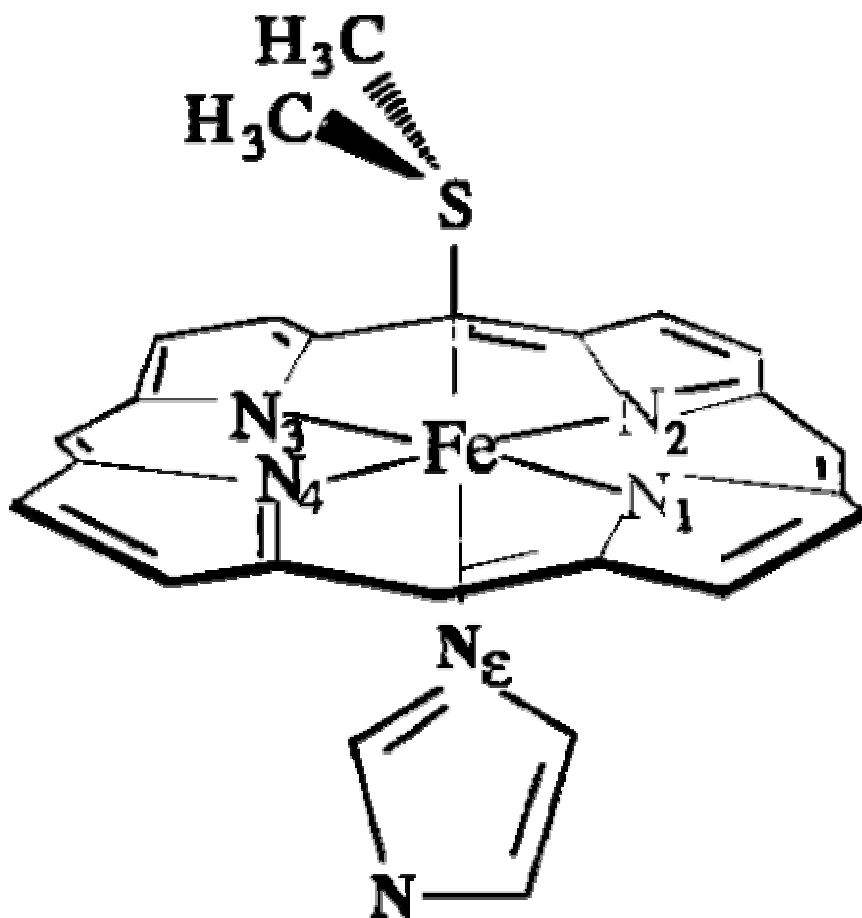
Figure 2.11. Model system used in computational studies aimed at understanding the observed frequency shift upon reduction of the heme. The porphyrin ring system is included, along with a dimethyl sulfide group on top to represent the Met80 ligand and an imidazole group on the bottom to represent the His18 ligand. Geometry optimizations and deuterium frequency calculations were performed on this model system using density functional theory.

Reduced   **Oxidized**
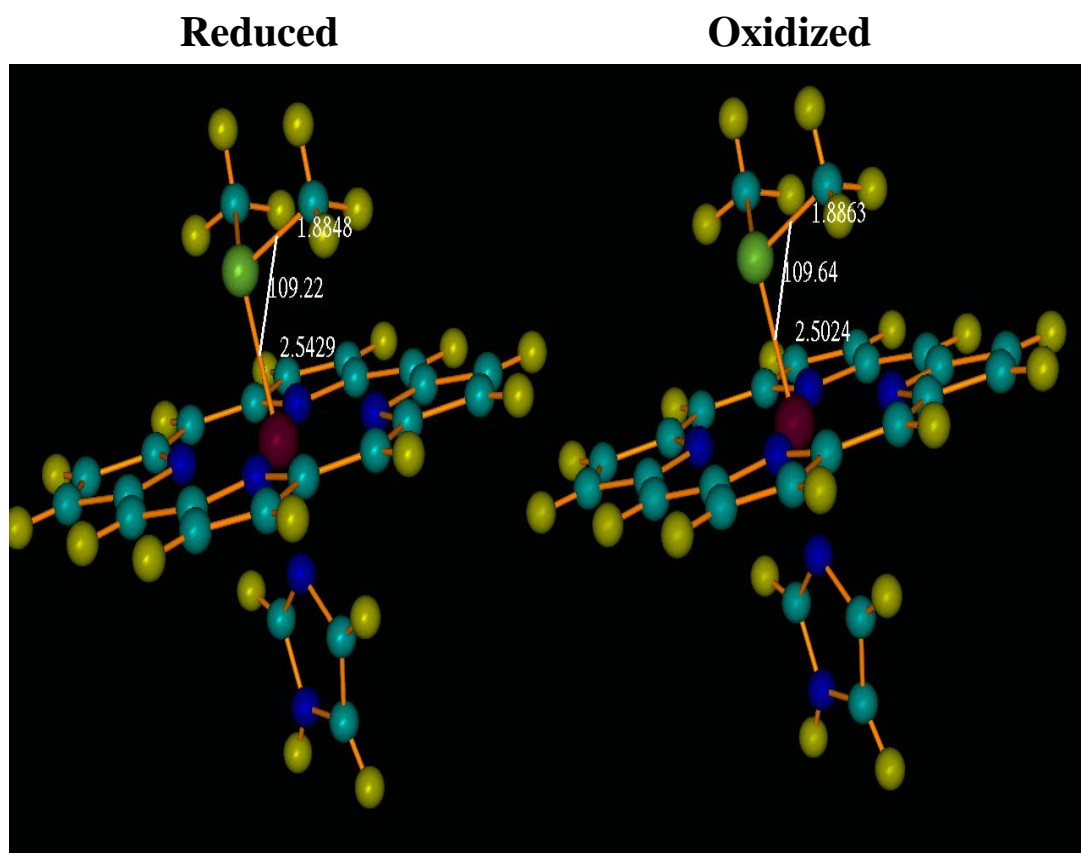
Figure 2.12. Geometry optimization performed on the model system in Figure 2.11 at the B3LYP/LANL2DZ level of theory. The change in bond angle of the Fe-S-C bond indicates more sp^2 character in the oxidized state, indicating stronger bonds.
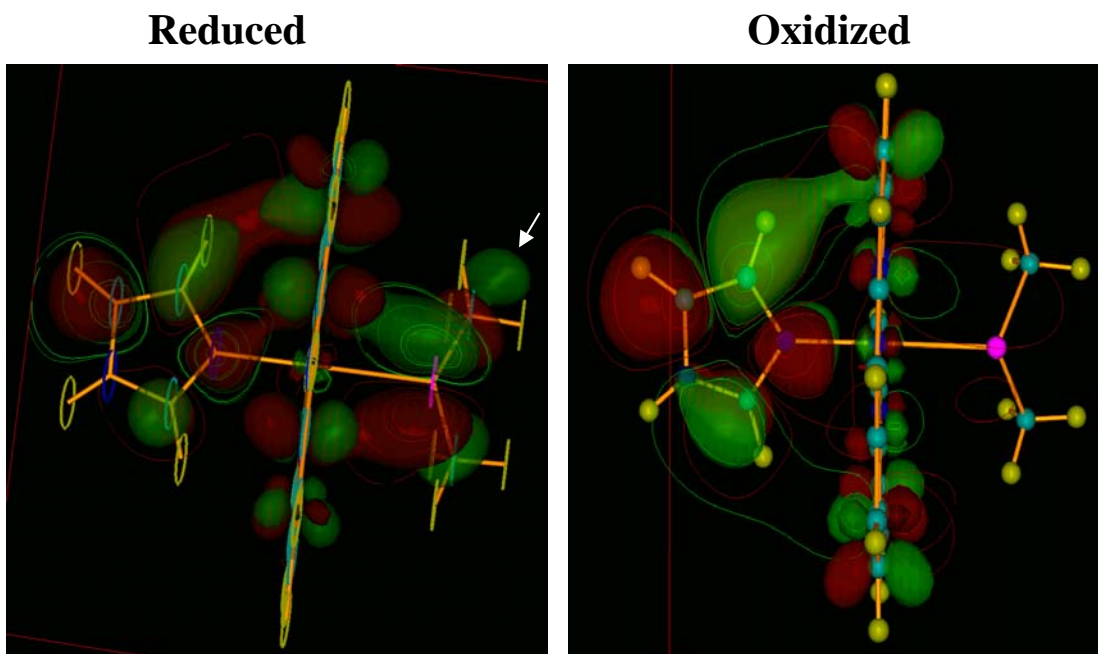
Reduced   **Oxidized**

Figure 2.13. Electron density maps of the reduced and oxidized model system shown in Figure 2.11 using the B3LYP/LANL2DZ basis set. Electron density shown in red has a negative charge and green a net positive charge. The most pronounced difference in electron density at the carbon-deuterium bonds on the dimethyl sulfide is the presence of density in the antibonding orbitals of the reduced species. This is another indication of a weaker bond in the reduced state.
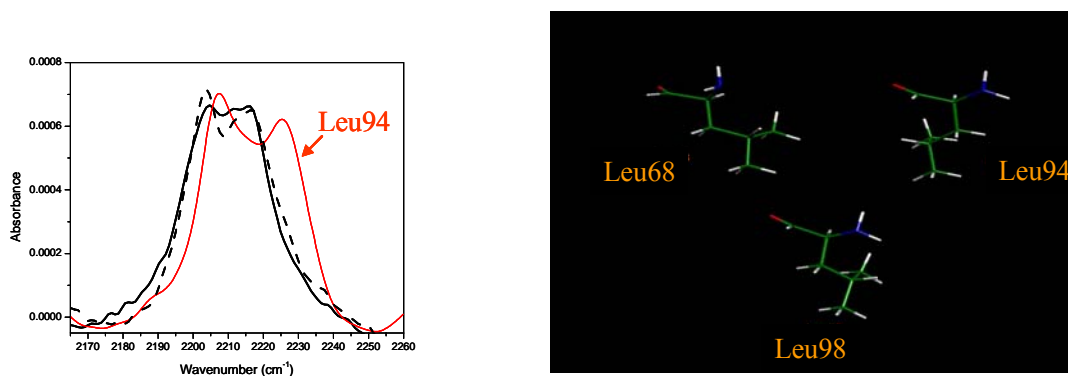
	Calculated Frequencies	Observed Frequencies
Leu68	2262.2	2204
	2271	2217
Leu98	2266.8	2203
	2274	2217
Leu94	2265.4	2208
	2280.2	2226

Figure 2.14. Computational study directed at understanding why the spectral features of d_3-Leu94 in cytochrome c are different from d_3-Leu68 and Leu98. As shown, upper left, d_3-Leu94 has two peaks of higher frequency with a larger gap than the other two leucines. Geometry optimizations, starting from the x-ray crystal structure, upper right, reveal Leu94 has a different geometry than the other two leucines. Deuterium frequency calculations can quantitatively reproduce the different spectral features of d_3-Leu94 from the difference in geometry alone.

Another computational study revealed that geometric effects are also contributing to the observed frequencies at residue Leu94 in cytochrome *c*. The two residues Leu68 and Leu98 have similar spectral features, with two resolved peaks at ~2203 and ~2216 cm^{-1}. However, Leu94 has two peaks further apart and at higher frequency, 2208 and 2226 cm^{-1} (see figure 2.14). A geometry optimization was preformed, using the B3LYP/DZV(2df,pd) basis set, on the three residues starting from the crystal structure geometry, and resulted in a very different orientation for Leu94. The other two leucine residues, which show similar spectral features in the protein, Leu68 and Leu98, had a similar geometry. Deuterium frequency calculations on the geometry optimized leucines showed that the spectrum of Leu94 does have higher frequency values with a larger gap between them, and appear to be a direct result of the different geometry.

Many factors affect the observed frequency and the direction to which the frequency shifts upon a given change to the local environment. Carbon-deuterium bonds have been shown to be sensitive to both through-space and through-bond interactions. The solvent studies show that a higher frequency results from a more polar environment. Through-bond interactions lend themselves to complementary computational studies and have proven useful in understanding the factors contributing to a given effect, such as changes in electron density or geometry. However, there are still unanswered questions reflecting the biological significance of the observed changes. It can be argued that the electron density changes at the Met80 ligand can be important for stabilizing the reduced state of the protein, a result of the

stronger Fe-S bond in the reduced state. Likewise, is it conceivable that Leu94 has more significant tertiary interactions since it is at the apex of where the C and N-terminal helices come together and that this tight packing might give rise to a different side-chain geometry. Through-space interactions are difficult to assess without further computational studies, especially within the rigid, inhomogeneous environment of a protein. Without further studies, it is difficult to disentangle a frequency shift arising from the deprotonation of a nearby residue or a more global effect from changing the state of the protein. Although an understanding of spectral frequencies and changes associated with them is an ongoing goal that has not fully been realized, simply observing changes of any kind with high molecular resolution is often times a sufficiently sensitive to assess biological significance. Interestingly, folding studies have shown that the observed frequencies within a protein are very sensitive to local tertiary packing. The incorporation of deuterated residues into an α-helical peptide yielded spectral features similar to those observed for the free amino acid in solution. Likewise, if the x-ray crystal structure shows the side chain of a residue to be pointing into solution, the observed carbon-deuterium frequencies are also close to those of the free amino acid in buffer.

2.2.2 Carbon-Deuterium Lineshapes and Linewidths

Carbon-deuterium bonds offer a unique view into the coupling and dynamics of a single bond within a protein. Since it is possible to view a single carbon-deuterium vibration, the inhomogeneous linewidth should provide localized dynamic information

with unprecedented time resolution. In addition, although the deuteration of a C-H bond results in protein vibrations that are adiabatically decoupled from the majority of other protein vibrations, they often possess specific low order resonances that facilitate intramolecular vibrational energy redistribution (Swofford, 1976). Thus, the homogenous linewidth component should yield information relating to intramolecular vibrational energy redistribution within the protein.

An effort to understand factors contributing to the observed linewidth and lineshape was carried out by Dr. Mathew Cremeens in the Romesberg group in collaboration with the Straub group. In this study, Boc-d_1-methionine was synthesized (by Dr. Shigeo Matsuda) and incorporated at the Met80 position in cytochrome c. Both d_1- and d_3-methionine were compared as free amino acids in solution and the Met80 position in cytochrome c. Three experimental observations resulted: the spectra of both the d_1 and d_3-methionine have smaller linewidths in the protein than in solution, the linewidths of the d_1-methionine spectra both in the protein and solution are larger than those of d_3-methionine, and the symmetric stretch of d_3-methionine in solution has a significant homogenous contribution to its linewidth. Figure 2.15 illustrates these observations, showing a comparison of the d_1 and d_3-methionine spectra in solution and the protein.

Lineshape analysis was carried out by fitting the spectra to a Voigt profile, which consists of a convolution of Lorenztian and Gaussian distributions. This analysis revealed that the majority of d_1 and d_3-methionine spectra both in the protein and solution are dominated by the Gaussian component, which suggests that
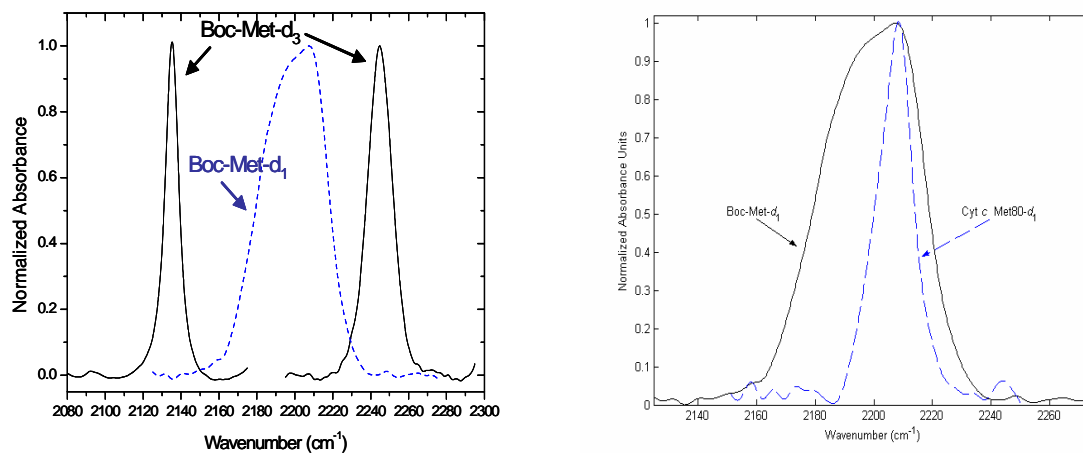
Figure 2.15. A study directed at understanding the dominant contribution to the observed linewidth. It was observed that all spectra have a smaller linewidth when going from solution to the protein. Also, the singly deuterated methionine has a much greater linewidth than the d_3-methionine in both solution and the protein. It was expected that the triply deuterated species would have a greater linewidth due to an increased homogenous contribution from coupling of adjacent carbon-deuterium bonds. Thus, the spectral observations suggest the inhomogeneous contribution is dominant.

inhomogeneous broadening is the major line-broadening mechanism. The dominance of inhomogeneous line-broadening is further supported by the fact that the d_1-methionine spectra have a greater linewidth, and not a smaller linewidth, than the spectra of d_3-methioinine. In the d_3 species, it is expected that homogenous factors would be much greater since it is likely that adjacent carbon-deuterium bonds would couple. Likewise, the greater linewidths in solution compared to the protein also suggests inhomogeneous broadening is the prevailing factor since the protein environment is expected to be much more distinct, and therefore yield different spectral features for each possible conformation. Interestingly, the symmetric stretch of d_3-methionine in solution shows a distinctly larger homogenous contribution than any other absorption, indicating the symmetric stretch is more sensitive to homogenous factors. While a detailed analysis of the lineshapes and linewidths of d_3 and d_1-methionine in solution and cytochrome c is being carried out, time-resolved studies are also underway. The homogenous contribution to the linewidth consists of two factors, the natural lifetime and pure dephasing times. A time-resolved pump-probe measurement has the capability to directly measure lifetimes. A photon echo measurement can alternatively measure dephasing times. The direct assessment of the homogenous contribution through time-resolved studies will allow for a better understanding of factors giving rise to the observed linewidth and lineshape, and lead to the quantification of the inhomogeneous effects. With knowledge of the homogenous contribution to the linewidth, one can tailor the incorporation of carbon-deuterium bonds in proteins to gain insight into energy pathways of the protein.

Overall, the dominance of the inhomogeneous contribution to the observed linewidths implies that a change in linewidth results from a change in local flexibility within in the protein, which will be further discussed in Chapter 3.

2.3 References

Barth, A., and Zscherp, C., What Vibrations Tell Us About Proteins. *Quarterly Reviews of Biophysics* **2002**, 35, 369.

Bellamy, L., *The Infrared Spectra of Complex Molecules, 2nd Edition*. John Wiley & Sons, Inc, New York, New York, 1958.

Creighton, T. E. *Protien Structure: A Practical Approach*. Ed.; Creighton, T. E., IRL Press, Oxford, England, 1989.

Dellepiane, G., and Overend, J., Vibrational Spectra and Assignment of Acetone Alphaalphaalpha Acetone-d3 and Acetone-d6. *Spectrochimica Acta*, **1966**, *22*, 593.

Eaton, W. A., and Hochstrasser, R. M., Electronic Spectrum of Single Crystals of Ferricytochrome c. *J. Chem. Phys*., **1967**, *46*, 2533.

Halonen, L., Theoretical Study of Vibrational Overtone Spectroscopy and Dynamics of Methanol. *J. Chem. Phys*., **1997**, *106*, 7931.

Mauk, A. G., Coyle, C. L., Bordignon, E., Gray H. B. Bis(Dipicolinate) Complexes of Cobalt(III) and Iron(II) as New Probes of Metalloprotien Electron Transfer Reactivity – Analysis of Reactions Involving Cytochrome c and Cytochrome C551. *J. Am. Chem. Soc*., **1979**, *101*, 5054.

Sarin, V. K., Kent, S. B., Tam, J. P., and Merrifield, R. B., Quantitative Monitoring of Solid-Phase Peptide Synthesis by the Ninhydrin Reaction. *Analytical Biochemistry*, **1981**, *117*, 147.

Schnölzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B., In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *International Journal of Peptide and Protien Research*, **1992**, *40*, 180.

Snedecor, G. W., Cochran, W. G. *Statistical Methods, Eighth Edition*. Iowa State University Press, Ammes, Iowa, 1989.

Swofford, R. L., Long, M. E., and Albrecht, A. C., C-H Vibrational States of Benzene, Naphthalene, and Anthracene in Visible Region by Thermal Lensing Spectroscopy and Local Mode Model. *J. Chem. Phys*., **1976**, *65*, 179.

Wallace, C. J. A., and Clark-Lewis, I., Functional Role of Heme Ligation in Cytochrome *c*. *J. Biol. Chem.,* **1992,** *267*, 3852.

CHAPTER 3. AN INVESTIGATION OF REDOX-COUPLED DYNAMICS IN

CYTOCHROME C

3.1 Introduction

A central question in all biological sciences is how the structure of a protein mediates its function. For many proteins, conformational changes within the subunits of a protein, changing the overall structure, work to activate or inactivate the protein (Voet, 1995). Structural changes of a protein are often brought about through binding another protein or chemical agent and are an important link between biophysical research and the pharmaceutical industry. Understanding the structural changes, how they occur, how they transcend through different parts of the protein and the effect they have is often times a difficult task. One example is the hemoglobin protein, which despite extensive study, the conformational changes that regulate oxygen binding are still not well understood (Eaton, 1999). Likewise, the conformational changes associated with the oxidation or reduction of the heme in cytochrome c also remain an enigma.

Structural changes that might accompany a change in redox state have been sought after since the discovery of cytochrome c itself. The protein is presumed to selectively stabilize the reduced state of the cofactor, possibly to optimize its ability to accept an electron from the cytochrome c reductase complex (Tezcan, 1998). Surprisingly, both the x-ray crystal structures and the NMR structures of the oxidized and reduced states of the protein are almost identical (Qi, 1996 and Berghuis, 1992). However, careful inspection of the structural studies indicate a few subtle differences between the two redox states of the protein. Higher thermal factors, associated with

increased mobility, were observed in the crystal structure of the oxidized protein for residues 47-59, 65-72 and 81-85. Oxidation of the protein also induces increased distortion of the heme plane and a rotation of the imidazole ring of the His18 ligand. Lastly, oxidation of the heme causes a displacement of an internal water molecule, which in turn disrupts a hydrogen bonded network that includes residues Met80, Try67, Thr78, and Asn52 (Brayer, 1995). The disruption of this hydrogen bonded network has been suggested to play a role in the decreased stability of the Met80 ligand in the oxidized state of cytochrome *c*.

Before the x-ray crystal structure of cytochrome *c* was solved, early studies indicated that the main difference between the two redox states of the protein is a dynamical one. The oxidized form has a larger radius of gyration, larger adiabatic compressibility, and lower viscosity than the reduced form. In addition, the oxidized form of the protein was shown to react with proteases and chemical reagents much faster as well as undergo hydrogen-deuterium exchange at increased rates compared to the reduced protein (Brayer, 1995). A recent NMR study measuring the relaxation rates of the oxidized and reduced forms of cytochrome *c* also concluded that the oxidized form of the protein is more flexible (Barker, 2001). Through use of a diverse array of techniques and many years of study, the general consensus remains that the main difference between the two redox states of cytochrome *c* lies in the increased flexibility in the oxidized state. However, a recent amide-exchange NMR study investigating the oxidized and reduced forms of cytochrome *c* showed that some regions of the protein exhibit a redox-dependent reduction of protection factors and

exchange at faster rates (Baxter, 1999). Interestingly, the residues implicated in the amide exchange studies involve the same regions shown to change upon oxidation in structural studies. These same structural motifs, the Met80 loop and the 60's helix, have also been shown to be more prone to partial unfolding and have decreased stability in the oxidized state (Bai, 1995).

In principle, the most direct way to measure flexibility in a protein would be to employ vibrational spectroscopy to characterize atomic motions. However, the spectral complexity inherent to proteins has thus far limited the use of vibrational spectroscopy in the context of protein dynamics. The site-selective deuteration of residues throughout the protein offers an attractive alternative. Unlike most of the previously discussed techniques used to measure changes in flexibility of cytochrome c, this technique offers a direct measurement with the time resolution to distinguish flexibility from structural diversity. To further investigate the structure-function relationship of cytochrome c, several residues throughout the protein were deuterated and the vibrational spectra of the carbon-deuterium bonds examined in both redox states. The sites of deuteration are shown in Figure 3.1. Three residues are located in areas shown by other techniques to be effected by the change in redox state, Leu68, Met80 and Ala83 (Baxter, 1999 and Barker, 2001). The other three residues are all located in the C-terminal helix and have not been shown to be directly involved in the redox behavior of the protein.
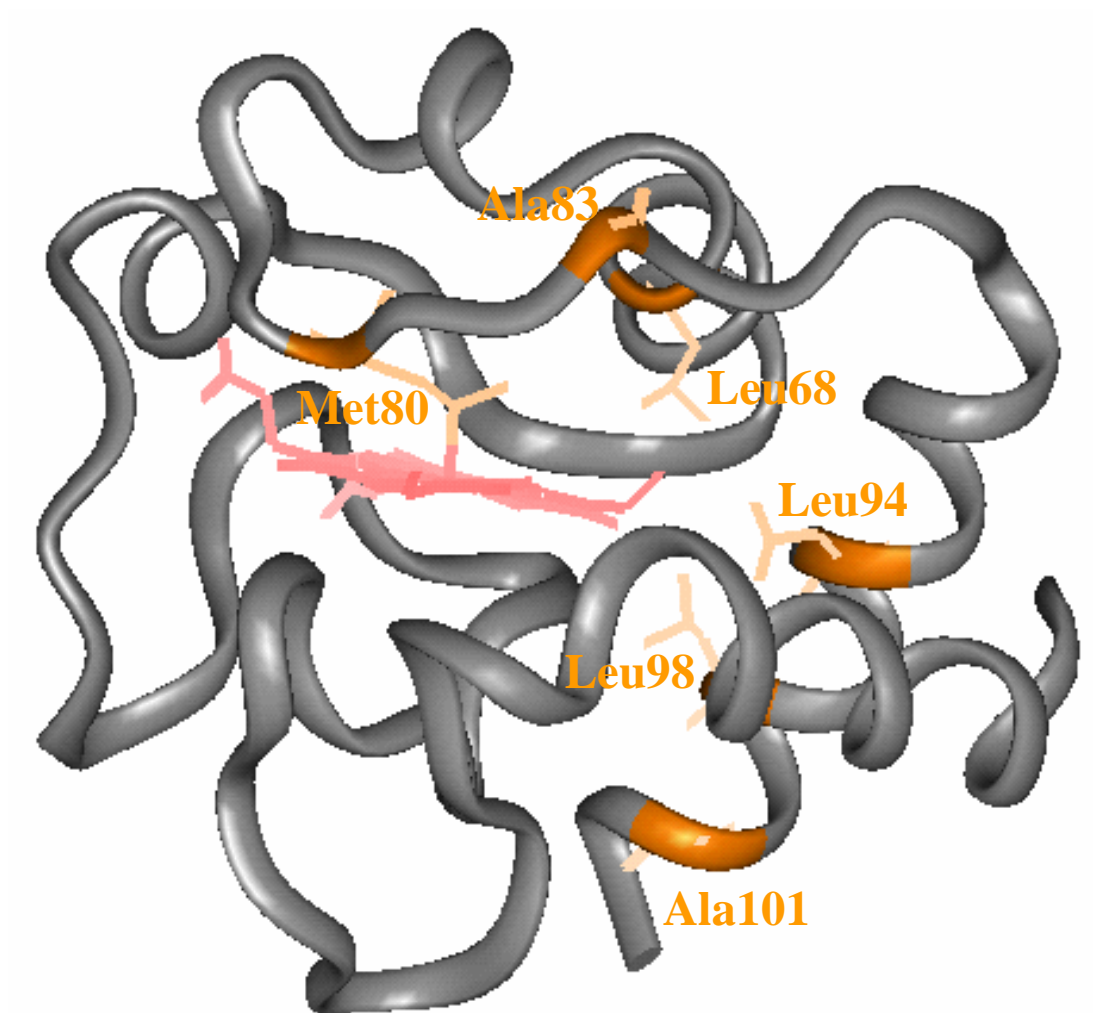
Figure 3.1. Crystal structure of horse heart cytochrome *c* showing the sites of deuteration. Six residues throughout the protein, including three different structural motifs were probed. In every case, the residue examined contained a deuterated terminal methyl group. Sites of deuteration include: two residues in the Met80 loop, d_3-Met80 and d_3-Ala83, one residue in the 60's helix, d_3-Leu68, and three residues in the C-terminal helix, d_3-Leu94, d_3-Leu98 and d_3-Ala101.

3.2 Materials and Methods

Sample Preparation. As described in Chapter 2, the site-selective deuterated cytochrome *c* was prepared through its semi-synthesis. Briefly, a cytochrome *c* 1-65 homoserine lactone peptide fragment containing the covalently bound heme was generated using CNBr cleavage at the Met65 residue of horse heart cytochrome *c*. This peptide was purified and then refolded at pH 7 in the presence of one equivalent of chemically synthesized residues 66-104 of cytochrome *c* containing a specifically deuterated leucine or methionine residue (Schnölzer, 1992). Fragment association and aminolysis of the homoserine lactone resulted in high yields of semi-synthetic cytochrome *c* with the slight modification of a homoserine at position 65 instead of a methionine (Wallace, 1992). This mutation has been shown to have no effect the redox potential, function and common folding markers of the protein. The final condensation product was purified from the fragments, and oxidized by adding a saturated solution of Bis(dipicolinato)cobaltateIII followed by a sephadox column (Mauk, 1979). The samples were then washed with 10 mM sodium acetate buffer at pH 5, divided on a per sample basis and lyophilized. Samples were stored at -20 ºC until ready for use. Oxidized lyophilized protein samples were dissolved in a 100 mM sodium acetate buffer solution, pH 5, and allowed to stand approximately ten minutes prior to measurement. Reduced samples were prepared by taking the oxidized lyophilized protein and adding a degassed mixture of 100 mM sodium acetate buffer, pH 5, and 100 mM ascorbic acid.

FTIR Measurements. Samples were measured using a Bruker Equinox 55 FTIR spectrometer equipped with a liquid nitrogen cooled MCT detector, KBr beamsplitter and the sample compartment continuously flushed with dry nitrogen. Spectra were collected at both 2 cm^{-1} and 4 cm^{-1} resolution and constructed from 8192 scans. The Blackman-Harris 3-term apodization function was used, a 16 kHz low pass filter and zero filling of 16. The aperature setting was 2000 μm and the liquid transmission cell consisted of CaF$_2$ disks 32 mm in diameter with a Teflon spacer 50 μm thick. For each deuterated residue at least three deuterated and proteo samples were measured and standard deviations were calculated from the independent measurements. After the measurement of every deuterated sample, a proteo sample was taken and subtracted using a subtraction factor close to unity. Once proteo subtraction was carried out, the baseline was corrected by fitting a window of at least 100 cm^{-1} to a polynomial and subtracting the polynomial to result in a relatively flat baseline with the peaks of interest.

Spectral Assignments. Spectra of the free amino acid in buffer for each residue examined were compared with quantum mechanical calculations to assign the spectral peaks, as described in Chapter 2. In every case, the deuterated methyl groups resulted in three C-D absorptions, two asymmetric stretches at high frequency and one symmetric stretch at low frequency. The unfolded spectra for all the deuterated amino acids in the protein were assigned based on the spectral features of the free amino acid in buffer in agreement with folding results in Chapter 5 (see Figure 3.2). The folded spectra within the protein for *d$_3$*-Ala83 contained identical features to the free amino
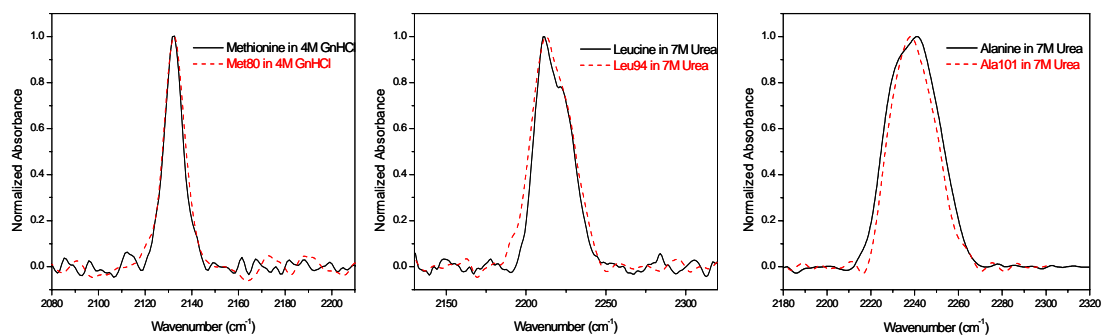
Figure 3.2. Folding data showing that any residue in the protein, when unfolded, has identical spectral features to that amino acid in solution. Thus, residues within the protein containing unfolded protein were fit by fixing the parameters to those of the free amino acid in solution.

acid in solvent (unfolded spectra) and both reported absorptions are assigned as the asymmetric stretches. Unfortunately, for both d_3-Ala83 and d_3-Ala101, the symmetric stretch was too low in intensity to yield reproducible results and is not reported. For d_3-Ala101 only one higher frequency absorption appeared to be present, and since two asymmetric stretches were predicted and observed in the unfolded protein, an F-test was carried out to distinguish if one or two peaks were present, as described below (Snedecor, 1989). Similarly, the high frequency absorptions of d_3-Met80 were assigned as the asymmetric stretches and also did not clearly show two peaks so an F-test was used to determine the number of Gaussians required to fit the spectra. In accord with the calculations, the lower frequency, intense band of d_3-Met80 was assigned as the symmetric stretch. The three leucine residues within the protein showed two absorptions similar to the spectra of the free amino acid in buffer. The free amino acid spectra of leucine in buffer is similar to that of alanine in buffer and therefore assigned as the two asymmetric stretches. However, unlike alanine, leucine was purchased in a 1:1 ratio of $C\delta^1$ and $C\delta^2$ isotopomers. Although the absorptions of the two isotopomers appear to be degenerate in the free amino acid, the leucine residues within the protein may show differentiation of the two methyl groups. F-tests were carried out to determine whether two or four absorptions were present for the folded spectra of the leucine residues within the protein and it was found the data are best fit by two Gaussians. As discussed in Chapter 2, the two folded absorptions for the leucine residues in the protein are assigned as the overlapping asymmetric stretches of the $C\delta^1$ and $C\delta^2$ methyl groups. In contrast, only one symmetric stretch

absorption for the leucine residues was observed, indicating the symmetric stretches are not sensitive to the differentiation of the methyl groups.

Spectral Fitting. The data were fit to Gaussians representing chemically meaningful species as described in Chapter 2. F-tests were carried out for the asymmetric stretches of d_3-Leu68, d_3-Leu94 and d_3-Leu98 in order to determine the number of Gaussians needed to fit the folded spectra, acquired in the absence of guanidine hydrochloride. At pH 5, the asymmetric stretches of the leucines in the protein could theoretically be fit to six Gaussians, two asymmetric stretches for each isotopomer and two Gaussians for the two unfolded asymmetric stretches. To eliminate some of the parameters needed in the deconvolution process, the F-tests were carried out with the leucine residues for spectra recorded at pH 7, where it can be assumed no unfolded protein is present. For the folded spectra (pH 7) it was then assessed whether two or four Gaussians were needed. The ratio of the least-squared sum for more parameters versus less yielded the F-test number, which was then compared to an F-table for the 99% confidence limit. As shown in Figure 3.3, the d_3-Leu98 spectra fit to four Gaussians is not a significant improvement compared to the two Gaussian fit. Although fitting the spectra for d_3-Leu68 and d_3-Leu94 to four Gaussians does show a slight improvement when compared to the two Gaussian fits, the extra Gaussians were either small in amplitude or severely overlapping. Thus, the folded spectra of all leucine residues were fit to two Gaussians representing the two isomers. Likewise the unfolded spectra, acquired at 6 M guanidine hydrochloride as

d$_3$-Leucines, F-Tests for Asymmetric Stretches
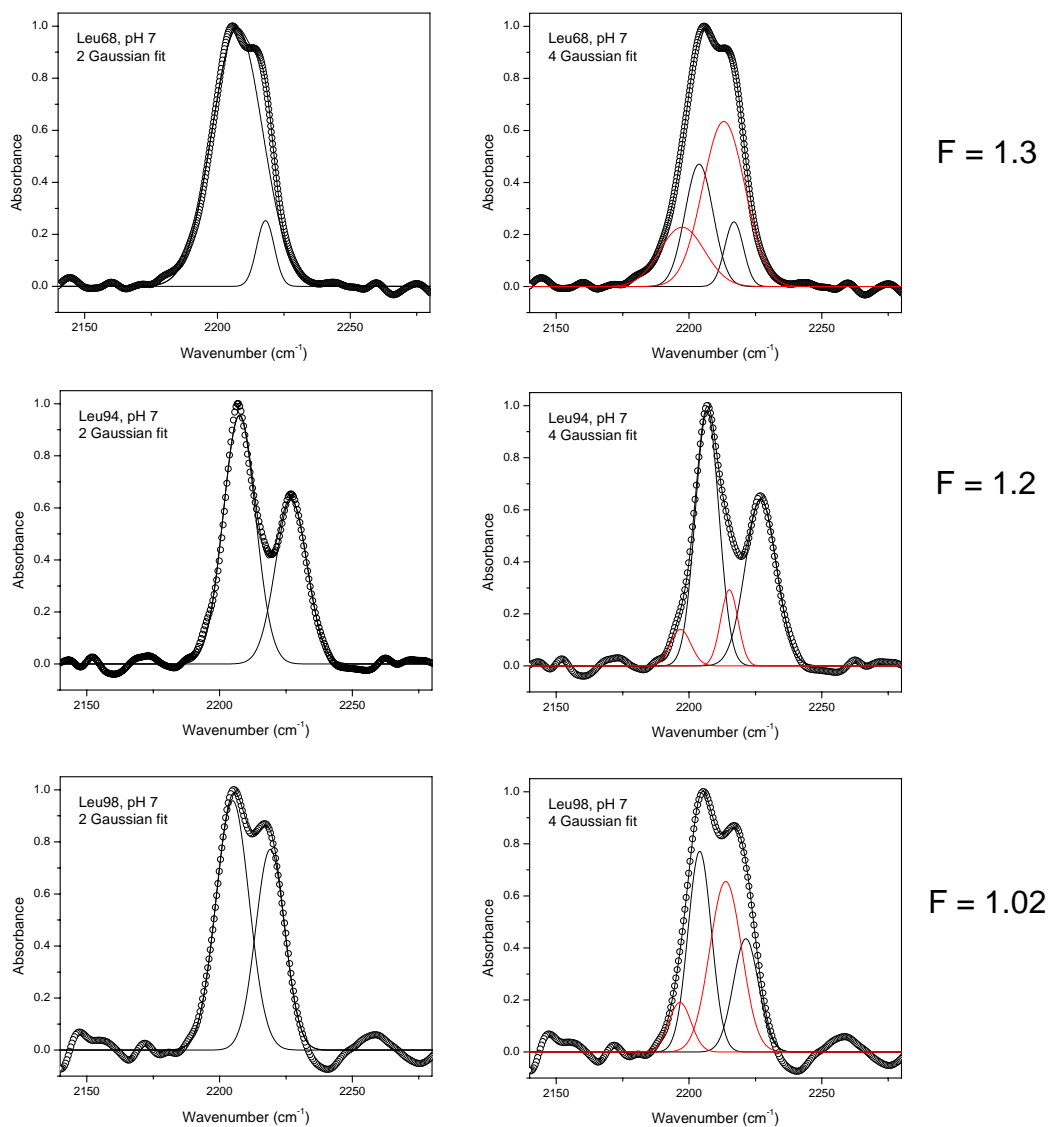
Figure 3.3. F-tests for d_3-Leu68, d_3-Leu94 and d_3-Leu98 to determine whether to fit the folded spectra to two or four Gaussians. Although d_3-Leu68 and d_3-Leu94 have F values slightly above the cut-off value of 1.03, d_3-Leu98 is below the cut-off. Since the extra Gaussians present in Leu94 and Leu68 were small in amplitude and in order to have consistency among the leucine residues, the folded spectra of all three leucines were fit to just two Gaussians.
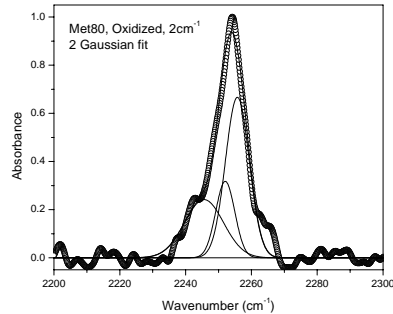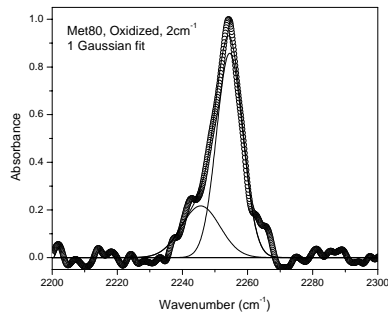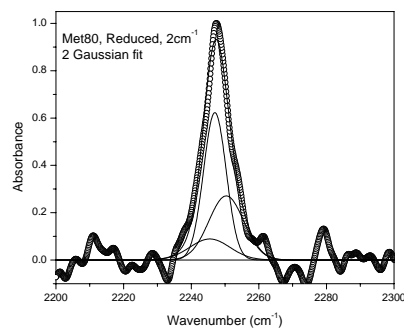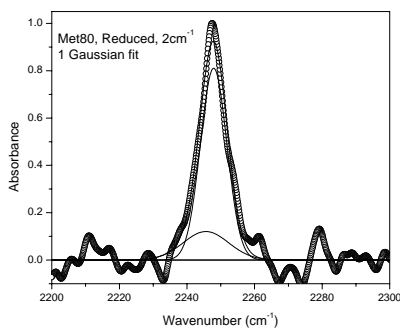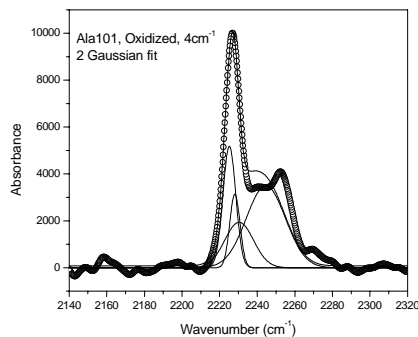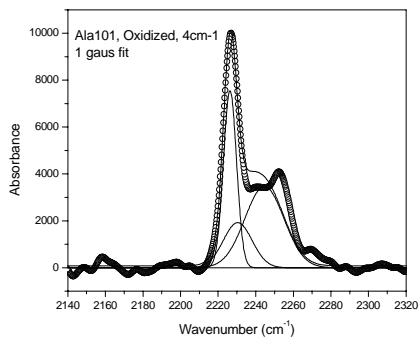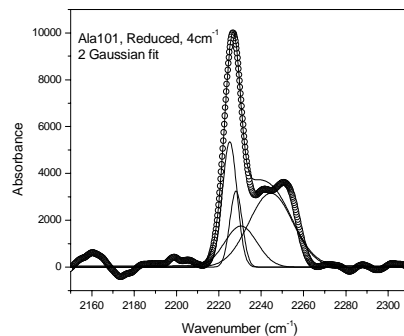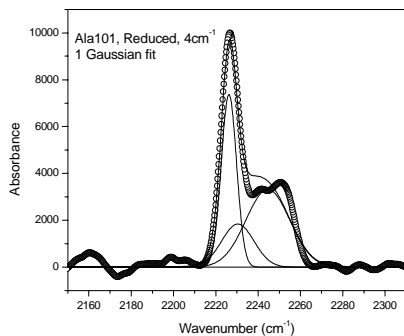
Figure 3.4. F-tests for d_3-Met80 and d_3-Ala101 to determine whether to fit the folded spectra to one or two Gaussians. All of the F numbers shown above are either below or close to the cut-off value of 1.03 for the 99% confidence limit, so only one Gaussian was used to fit the folded spectra of d_3-Met80 and d_3-Ala101.

well as the free amino acid in buffer, could also be represented by two Gaussians. Similarly, F-tests were carried out on the asymmetric stretches of d_3-Ala101 and d_3-Met80 to determine whether one or two Gaussians were needed to fit the spectra. As shown in Figure 3.4, fitting the oxidized and reduced spectra of both residues to two Gaussians did not show a significant improvement compared to the single Gaussian fit.   Therefore, the asymmetric stretches of both d_3-Ala101 and d_3-Met80 were determined to be severely overlapping and best fit with one Gaussian. In order to determine the spectral features of the folded spectra as well as the percent unfolded for each residue, only the parameters corresponding to the unfolded spectra were fixed to those of the free amino acid in buffered solution. The percent unfolded values shown in Table 3.1 are calculated from the relative intensities of an unfolded Gaussian to a folded one. In cases where the chosen unfolded absorption has a different extinction coefficient than the selected folded absorption, the percent folded values are weighted by the ratio of the extinction coefficients (see Table 3.2).

3.2 Results

The first specifically labeled cytochrome c to undergo investigation was d_3-Met80, which provides one of two protein-based heme ligands. The infrared spectra of both symmetric and asymmetric absorptions in the oxidized and reduced states of the protein are shown in Figure 3.5. Both symmetric and asymmetric stretches shift to lower frequency upon reduction. The asymmetric absorptions show greater
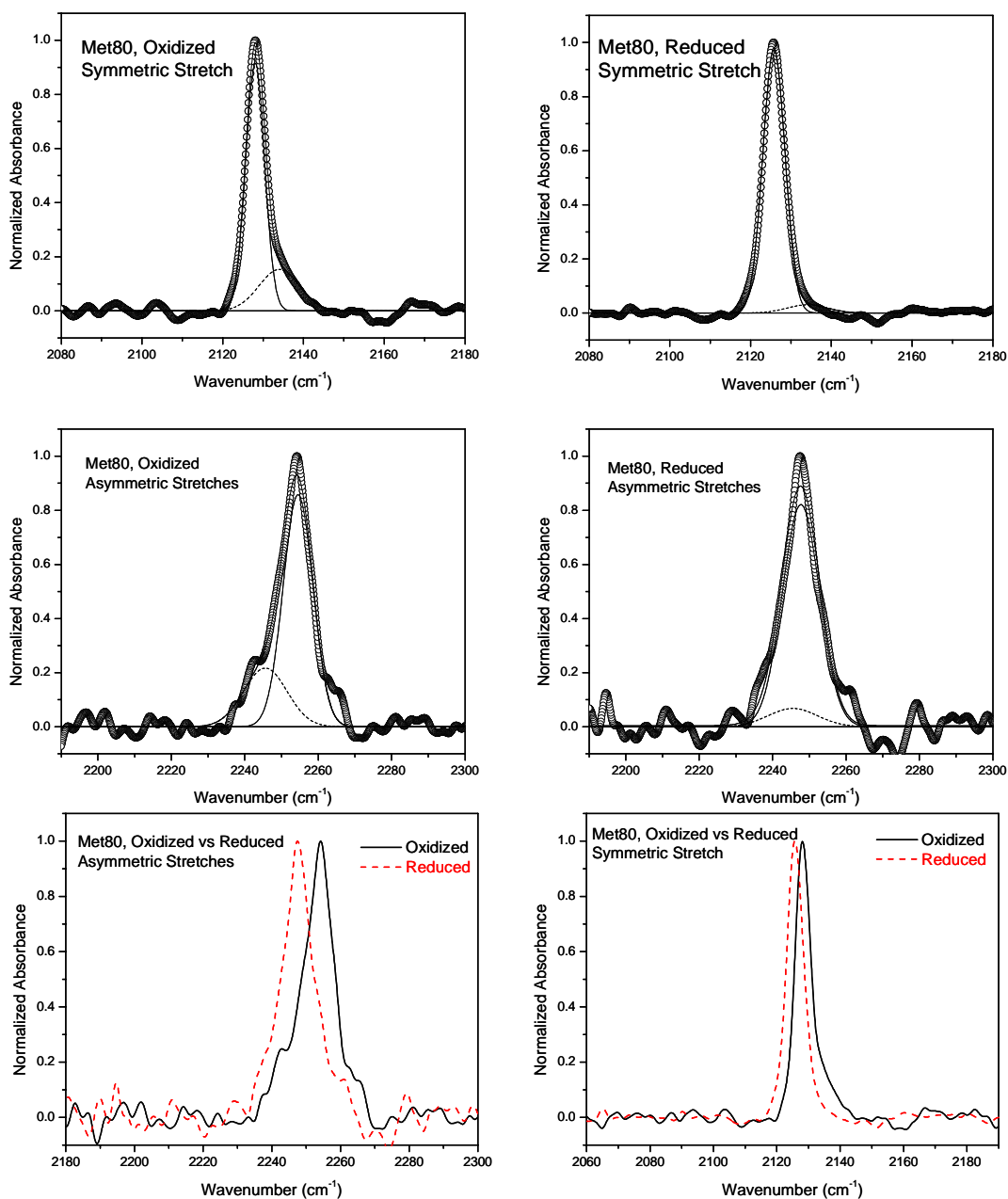
Figure 3.5. The asymmetric and symmetric stretches of d_3-Met80. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). Met80 shows the most significant redox-dependent frequency shift and percent unfolded.

sensitivity, shifting approximately 7 cm^{-1}, whereas the symmetric absorptions shift only 2 cm^{-1} upon reduction. In contrast, none of the absorptions change significantly in linewidth, indicating that, at least at this residue, there are no changes in flexibility that accompany the change in redox state. As most easily observed in the symmetric stretch absorption, the oxidized state has a considerable amount of unfolded protein present, whose spectra when fit yields approximately 10% unfolded protein. A similar amount of unfolded protein is deduced from the fits of the asymmetric absorptions, as shown in Figure 3.4. Surprisingly, the reduced state of the protein contains significantly less unfolded protein and fits to the asymmetric and symmetric absorptions only show 4%.

In order to probe the local environment and dynamics of a loop proximal to the Met80 ligand, d_3-Ala83 was examined. Unlike d_3-Met80, Ala83 shows no changes in frequency upon reduction. Likewise, no significant changes in linewidth are also observed. The asymmetric absorptions observed in the protein were similar to those of the free amino acid in solution, and therefore the folded and unfolded spectra were indistinguishable. Inspection of the crystal structure of oxidized horse heart cytochrome c reveals that the β methyl group of Ala83 is solvent exposed and most likely not sensitive to the protein environment (Bushnell, 1990). Folding measurements carried out with specifically deuterated small peptides and cytochrome c also indicate that residues with solvent exposed side chains often do not show changes in spectral features as the protein is unfolded (see Chapter 5). Thus, the amount of unfolded protein present could not be assessed for this residue.
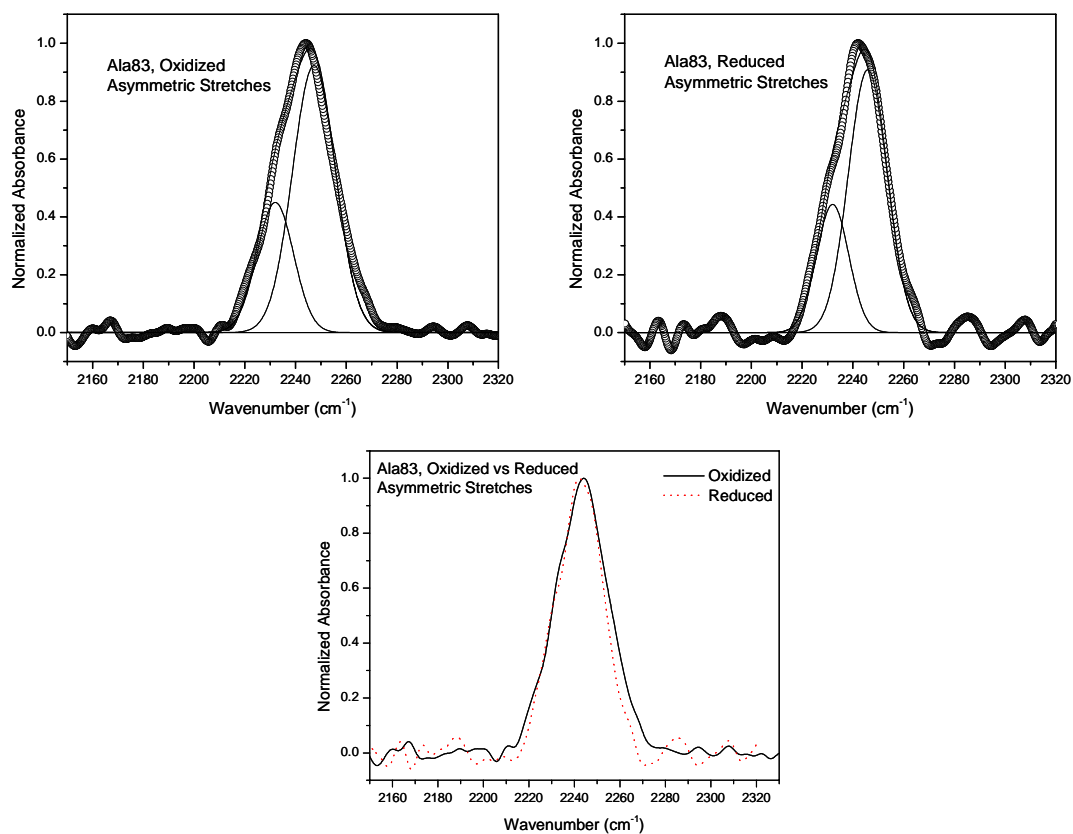
Figure 3.6. The asymmetric stretches of d_3-Ala83. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). No significant redox-dependent changes are observed at Ala83.

The next residue probed in cytochrome *c*, Leu68, resides in the 60's helix which is proximal to the Met80 loop. Interestingly, only one asymmetric absorption shifts to higher frequency upon reduction. This suggests that only one methyl group experiences an increase in local polarity (see solvent data in Chapter 2), while the environment of the other methyl group is not significantly changed. The crystal structure shows that the $C^{\delta 2}$ methyl group packs on the heme and would thus be expected to be sensitive to changes in the cofactor (Bushnell, 1990). The environment of the $C^{\delta 1}$ is not expected to change since this methyl group is oriented away from the heme. Therefore, the low frequency asymmetric absorption is assigned as the overlapping asymmetric stretches of the $C^{\delta 2}$ methyl group and the high frequency band, the overlapping asymmetric stretches of the $C^{\delta 1}$ methyl group. As shown in Figure 3.7, the symmetric absorptions are less sensitive to the change of environment and do not show frequency changes upon reduction of the heme cofactor. All absorptions show an absence of any consistent linewidth change which suggests that, similar to Met80 and Ala83, there are no significant redox-dependent changes in protein flexibility at Leu68. However, like at Met80, for both the symmetric and asymmetric absorptions there was a consistent and significant change in the amount of unfolded protein present in the two redox states. The oxidized state contains 7% while the reduced state only shows 3% unfolded protein present.

In order to probe the local environment of another structural motif further away from the Met80 loop, residues Leu94, Leu98 and Ala101, all located in the C-terminal helix, were examined. Surprizingly, Leu94 showed no redox-dependent changes in

Figure 3.7. The asymmetric and symmetric stretches of d_3-Leu68. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). Similar to Met80, significant redox-dependent changes in frequency and percent unfolded protein are observed at Leu68.

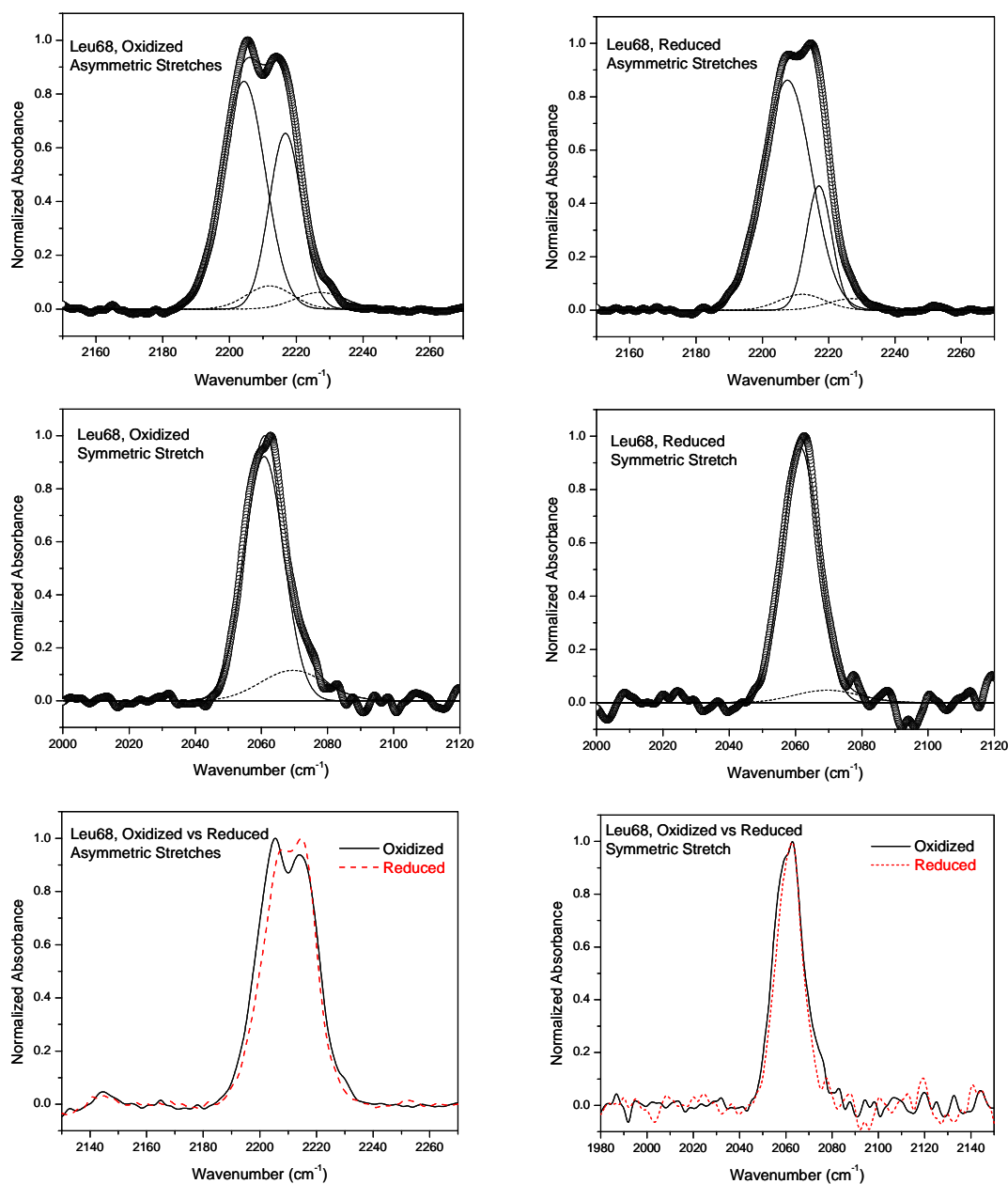**Figure 3.8**. The asymmetric and symmetric stretches of d_3-Leu94. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). No significant redox-dependent changes in frequency and percent unfolded protein are observed at Leu94.
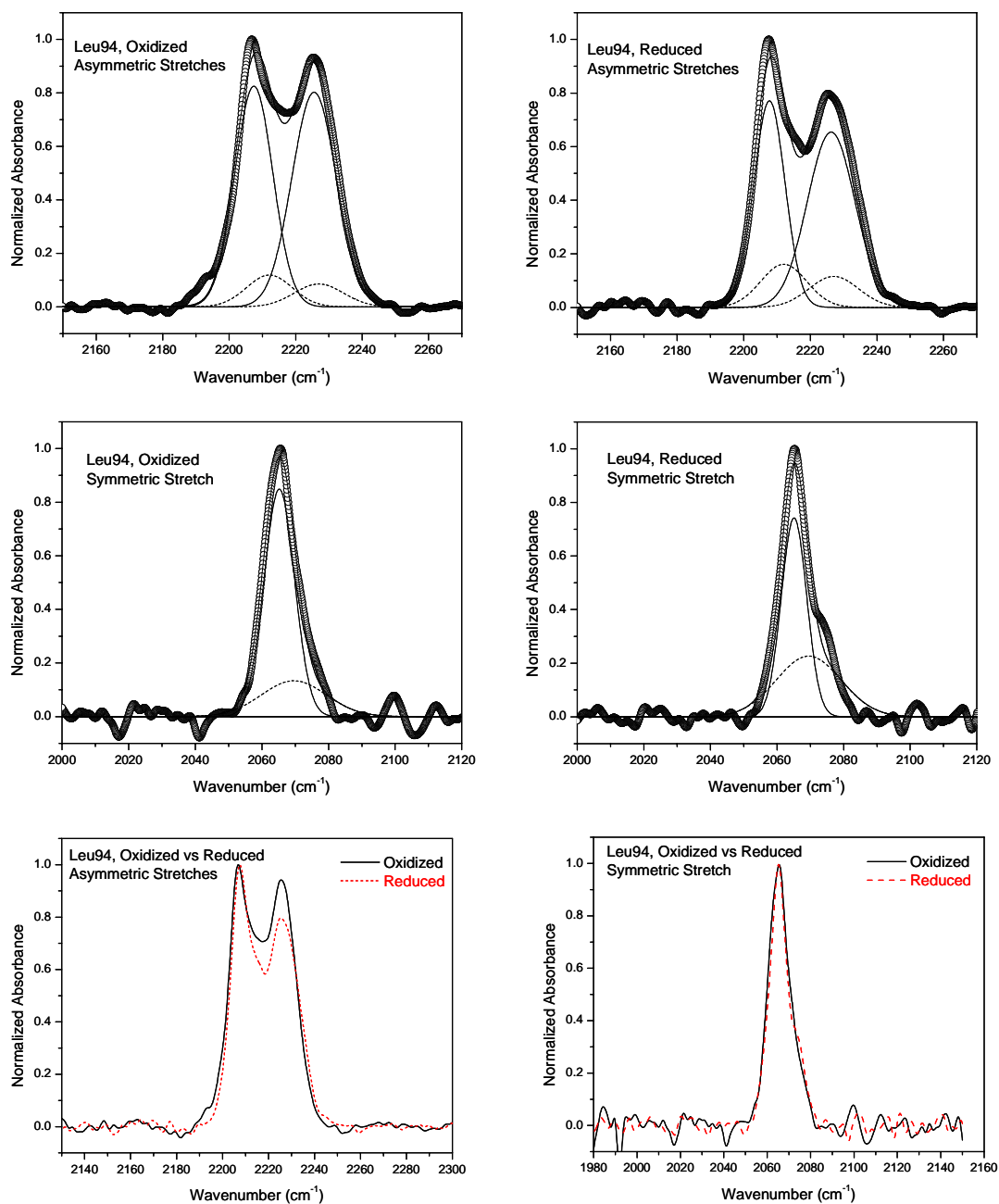
frequency, linewidth or unfolded protein present (see Figure 3.8). The crystal structure indicates this residue is not solvent exposed and the side chain is packed tightly against the N-terminal helix in the core of the protein, thus it is expected to be an accurate reporter of the local protein environment. The lack of a significant change in linewidth also indicates that there are no changes in flexibility at this residue upon reduction of the heme cofactor. Despite Leu94 being in the center of what is thought to one of the most stable structures in the protein, both the oxidized and reduced spectra indicate approximately 12% unfolded protein present.

Four residues from Leu94, residue Leu98 was also probed for redox-dependent changes. As shown in Figure 3.9, one asymmetric absorption of Leu98 does have a small frequency shift to higher frequency upon reduction. Interestingly, the direction of this frequency shift is the same as that observed at Leu68, but opposite that of Met80. Similar to Leu68, the symmetric absorptions do not show a frequency shift and like all other residues examined so far, Leu98 also shows no consistent changes in linewidth accompanying the change in redox state. The asymmetric and symmetric absorptions of Leu98 consistently show a trend similar to Leu94, with the oxidized and reduced protein containing ~12% unfolded protein.

Lastly, Ala101, close to the C-terminus of the protein was probed for redox-dependent changes. Only one asymmetric absorption is observed at 2226 cm^{-1} and like Leu94, it was insensitive to the change in redox state of the heme cofactor. No frequency, linewidth or percent unfolded differences were shown as the oxidation state of the heme cofactor is changed (see Figure 3.10). Spectral fits reveal a remarkable

Figure 3.9. The asymmetric and symmetric stretches of d_3-Leu98. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). No significant redox-dependent changes in frequency and percent unfolded protein are observed at Leu98.
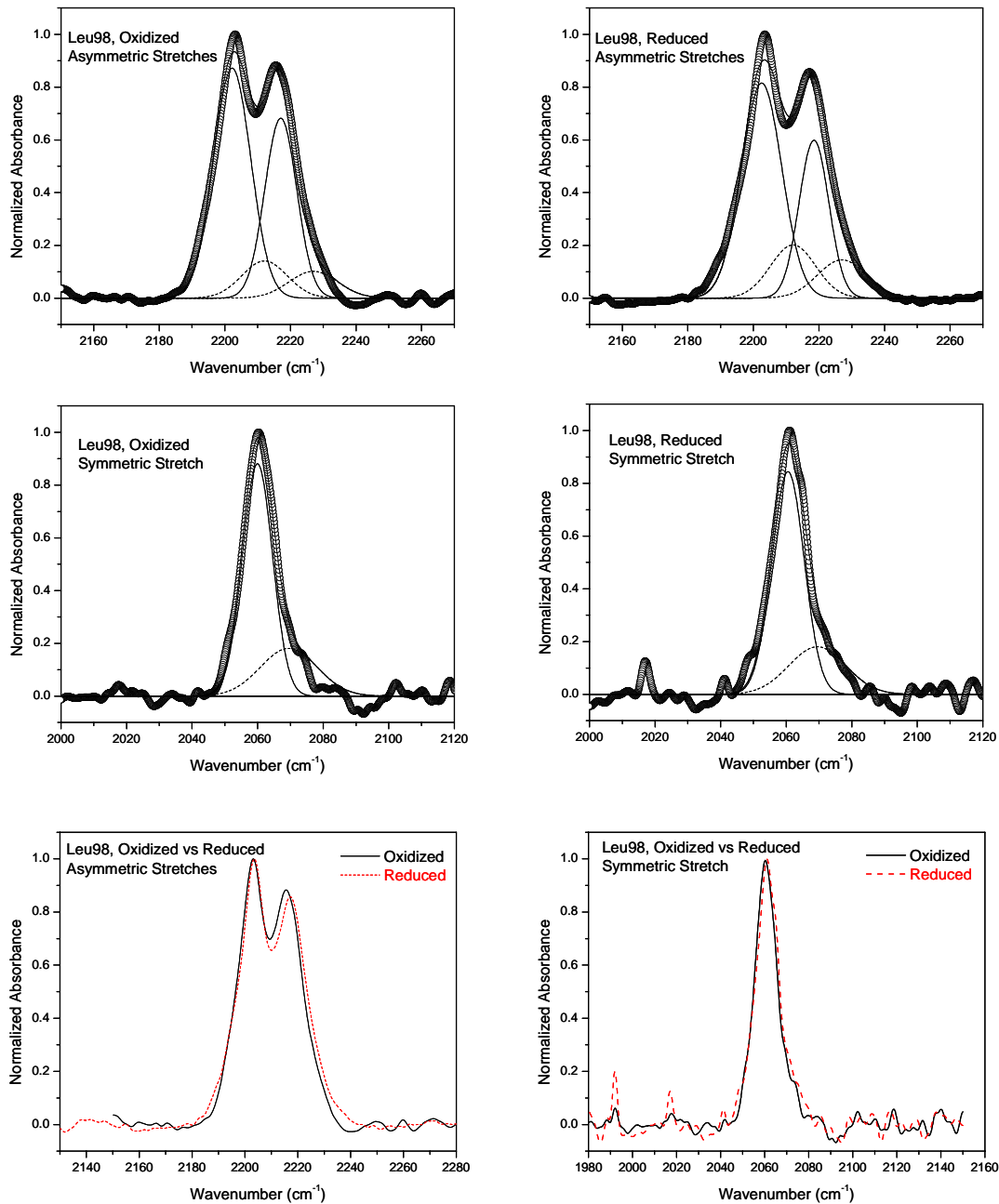
Figure 3.10. The asymmetric and symmetric stretches of d_3-Ala101. Gaussian fits to quantify the amount of unfolded protein present (top) and overlaid oxidized and reduced spectra (bottom). No significant redox-dependent changes in frequency and percent unfolded protein are observed at Ala101.
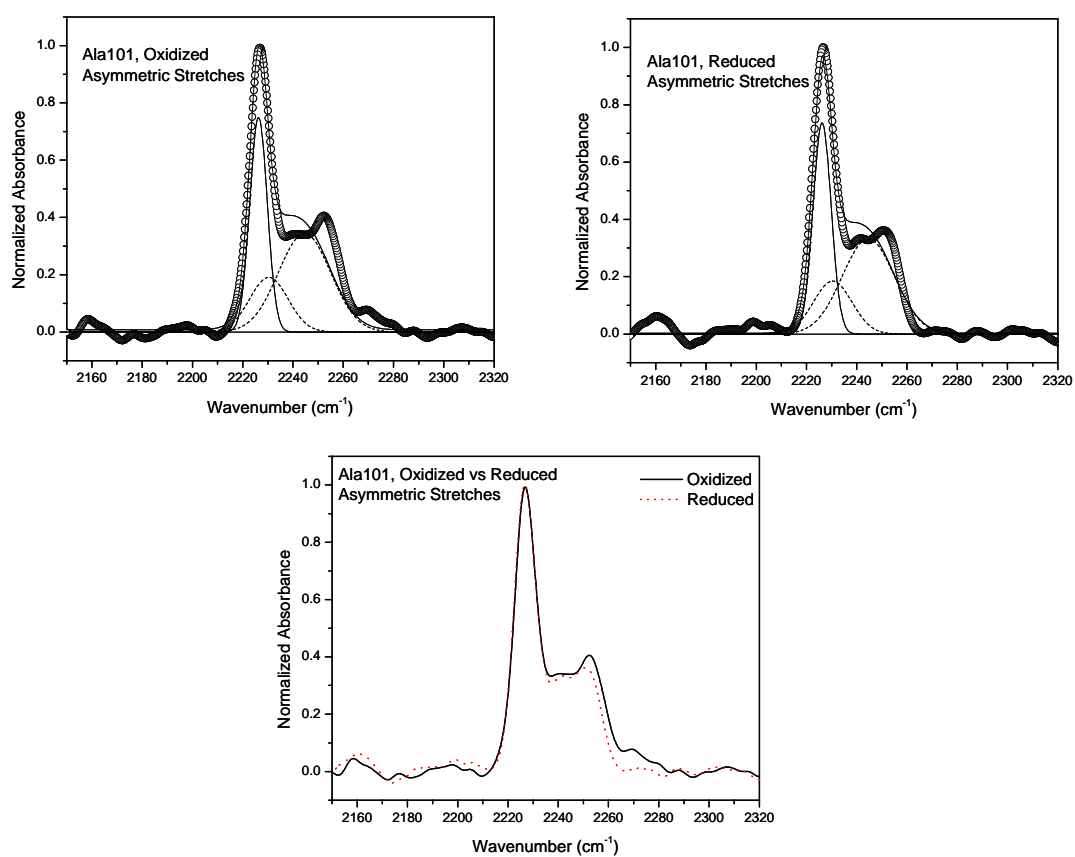
extent of unfolding at Ala101, approximately 50% in both the oxidized and reduced states, which is much larger than any other residue examined.

3.3 Discussion and Conclusions

Spectral Frequencies. As shown in Table 3.1, the only residues to exhibit a redox-dependent change in spectra frequencies are Met80, Leu68 and Leu98. Met80 shows the most pronounced change in both the asymmetric and symmetric stretches, with a frequency shift of 7 cm^{-1} and 2.5 cm^{-1} respectively. As discussed in Chapter 2, the frequency shift upon reduction of the heme has been shown to be dominated by through-bond effects (Chin, 2001). However, residues Leu68 and Leu98 are not directly connected to the heme and their spectral frequencies would therefore be expected to not be dominated by through-bond effects. Particularly noteworthy is that the sign of the frequency shift of both Leu68 and Leu98 is opposite that of Met80, increasing instead of decreasing as the heme is reduced. Based on the solvent studies of the free amino acids (Chapter 2), the frequency shift at Leu68 and Leu98 implies an increase in local polarity around these residues. As shown in Figure 3.1, both Leu68 and Leu98 pack on the heme with the deuterated methyl groups perpendicular to the heme plane. This geometry suggests a favorable interaction with the p orbitals of the heme cofactor and the possibility that the change in frequency is due to a change in orbital overlap. It is also interesting to note that the redox-dependent frequency change at Leu68 is larger in magnitude than that of Leu98. Leu68 is in a region of the protein proximal to other residues, such as Phe82 and Tyr67, whose mutation greatly

Table 3.1: Frequency, Linewidth and Percent Unfolded Values Measured for Six Deuterated Residues Throughout Cytochrome *c*

	Frequency (cm^{-1})		Linewidth (cm^{-1})		Δv [a]	ΔFWHM [a]	Percent Unfolded	
	Oxidized	Reduced	Oxidized	Reduced			Ox	Red
Leu68								
asym Cδ$_2$-A′/A″	2205 ± 0.2	2209 ± 0.3	14.9 ± 1	18.9 ± 2	**-4.4 ± 0.4**	**-4.0 ± 2.1**	**7% ± 1%**	**3% ± 3%**
asym Cδ$_1$-A′/A″	2217 ± 0.1	2217 ± 1	10.7 ± 0.3	7.0 ± 1	-0.2 ± 0.5	**+3.7 ± 1.0**		
sym Cδ$_1$/Cδ$_2$-A′	2061 ± 0.3	2061 ± 0.4	14.1 ± 1	13.0 ± 1	-0.5 ± 0.5	+1.1 ± 1.1	**7% ± 2%**	**3% ± 2%**
Ala83							--[b]	--[b]
asym A′	2231 ± 0.4	2231 ± 0.4	16.8 ± 1	14.3 ± 3	-0.2 ± 0.6	+2.5 ± 3.1		
asym A″	2247 ± 1	2246 ± 1	18.8 ± 1	17.6 ± 3	+1.1 ± 1.3	-1.2 ± 2.8		
Met80								
asym A′/A″	2255 ± 0.3	2248 ± 0.2	9.6 ± 1	11.9 ± 3	**+6.9 ± 0.4**	-2.3 ± 2.8	**15% ± 4%**	**4% ± 3%**
sym A′	2128 ± 0.1	2126 ± 0.1	5.5 ± 0.2	6.4 ± 0.2	**+2.2 ± 0.1**	-0.9 ± 0.3	**9% ± 4%**	**4% ± 1%**
Leu94								
asym Cδ$_1$-A′/A″	2207 ± 0.1	2208 ± 0.3	12.3 ± 1	10.6 ± 1	-0.3 ± 0.3	+1.7 ± 1.6	14% ± 3%	11% ± 3%
asym Cδ$_2$-A′/A″	2226 ± 0.2	2226 ± 0.3	14.9 ± 1	16.8 ± 1	-0.2 ± 0.4	-1.9 ± 1.2		
sym Cδ$_1$/Cδ$_2$-A′	2065 ± 0.3	2065 ± 0.1	10.5 ± 0.4	9.2 ± 0.4	+0.1 ± 0.3	+1.3 ± 0.6	12% ± 2%	14% ± 5%
Leu98								
asym Cδ$_2$-A′/A″	2202 ± 0.2	2203 ± 0.1	13.5 ± 1	14.5 ± 1	-0.3 ± 0.2	-1.0 ± 1.6	11% ± 1%	13% ± 1%
asym Cδ$_1$-A′/A″	2217.0 ± 0.1	2218.5 ± 0.1	11.2 ± 0.3	10.7 ± 0.4	**-1.5 ± 0.1**	+0.5 ± 0.5		
sym Cδ$_1$/Cδ$_2$-A′	2060.0 ± 0.2	2060.6 ± 0.2	10.9 ± 1.1	11.7 ± 0.8	-0.6 ± 0.3	-0.8 ± 1.4	10% ± 1%	10% ± 1%
Ala101								
asym A′/A″	2226.3 ± 0.2	2226.3 ± 0.1	8.4 ± 1.4	8.6 ± 1.2	0.0 ± 0.2	+0.2 ± 1.8	58% ± 6%	54% ± 4%

(a) Signs in front of Δv and ΔFWHM: positive values indicate oxidized state has a greater number, negative values indicate reduced state has a greater number

(b) Not determined due to spectral overlap, see text.

Changes in frequency, linewidth and percent unfolded values upon oxidation of the heme are shown above. The frequency and linewidth values are based on the Gaussian fits shown in Figures 5-10. The percent unfolded values are calculated from the intensities of the Gaussian fits weighted by the extinction coefficients shown in Table 2. The largest changes occur at residues Leu68 and Met80 on the proximal side of the heme.

Table 3.2: Extinction coefficients of amino acids in buffer and in protein.

amino acid (buffer)	symmetric stretch			asymmetric stretches	
	ν^a	ε^b		ν^a	ε^b
d_3-methionine	2136	14		2245	2.8
d_3-alanine	2081	2.5		2230.5	5.0
				2244.5	7.4
d_3-leucine	2070	7.3		2212	20
				2227	13.4

deuterated residue (protein)	symmetric stretch				asymmetric stretches			
	oxidized		reduced		oxidized		reduced	
	ν^a	ε^b	ν^a	ε^b	ν^a	ε^b	ν^a	ε^b
d_3-Met80	2128	8.0	2125.6	16	2254.5	2.2	2247.6	2.2
d_3-Ala83	*	*	*	*	2246.8	4.0	2245.5	3.0
d_3-Ala101	*	*	*	*	2226.3	6.0	2226.3	6.0
d_3-Leu68	2060.8	5.0	2061.3	4.5	2204.5	15	2208.9	19
					2216.9	12	2217.1	5.0
d_3-Leu94	2065.1	5.0	2065	5.0	2207.3	12	2207.6	10
					2225.6	13	2226.4	10
d_3-Leu98	2060	4.0	2060.6	4.0	2202.2	13	2202.5	12
					2217	12	2218.5	11

^{a} in cm^{-1}; ^{b} in $M^{-1}cm^{-1}$ * not determined

Extinction coefficient values for all relevant absorbing species. All values reported for residues within the protein were taken at pH 7 while assuming no unfolded protein to be present. The free amino acid in buffer has similar spectral features to the observed unfolded residue within the protein (unpublished results) and this spectrum was assumed to represent the unfolded species. The amounts of unfolded present thus reported are corrected for using the above extinction coefficient values.

affects the redox function of the protein (Hilgen, 1991 and Berghuis, 1994). Mutational studies have suggested that the local dielectric in this region of the protein may be important for the function of the protein. Perhaps a change in the local dielectric around Leu68 is also a factor in the observed redox-dependent frequency shift.

Despite the ample evidence to the contrary, the lack of any consistent changes in linewidth throughout the protein upon oxidation implies there are no redox-dependent changes in flexibility for cytochrome c. However, the data does offer an explanation for the results of previous studies that concluded the oxidized state was more flexible – at least some parts of the protein are more prone to unfold in the oxidized state. Increased flexibility of cytochrome c should cause increased amplitude motion about the folded structure and result in increased linewidths, as the carbon-deuterium bonds experience a greater diversity of environments. In contrast what was observed are distinct spectral absorptions corresponding to that of the unfolded protein. This locally unfolded protein under similar conditions was not detected in many previous reports, which is likely due to either low structural resolution, inherent to techniques such as tryptophan fluorescence or Soret absorption, or low time resolution relative to the chemical shift difference inherent to NMR spectroscopy (Goto, 1993 and Wooten, 1981). Although recent amide-exchange studies comparing the oxidized and reduced states of the protein under similar conditions also show a redox-dependent change in protection factors, due to the low temporal resolution of chemical exchange, the authors could not distinguish changes in flexibility from changes in local unfolding

(Baxter, 1999). On the contrary, this infrared-based method offers both high structural and temporal resolution; the folded and locally unfolded proteins would have to interconvert on the picosecond timescale for their signals not to be resolved.

The data reported also reveal a remarkable extent of unfolding in the C-terminal helix, especially at Ala101, where more than half the protein is unfolded. Although previous studies have shown C-terminal fraying to exist, the large extent of unfolding at Ala101, which is still three residues from the C-terminus, at both pH 5 and pH 7 is surprising (see Figure 11) (Baxter, 1999). While the present study was conducted at pH 5 to connect with previous work, the data indicate that the C-terminal helix is less stable at pH 5 than pH 7 (Chin, 2001). Despite the fact that the data suggest the polypeptide chain may be less stable at pH 5, studies have shown that there is significantly less misligated species present, due to the protonation of histidine side chains which are directly involved in the misligation events (Colon, 1997). Although it is possible the spectral changes observed at residues Met80 and Leu68 might result in increased misligation events in the oxidized protein, we believe that the spectral features refer to unfolded protein and not misligated protein for two reasons. First, at the Met80 residue, the misligated conformer is resolved and spectrally distinct from the unfolded species (see Chapter 4). Second, the amount of misligated conformer present at pH 5 is substantially less than that at pH 7 and kinetic studies estimate it to be less than 10% (Colon, 1996).

In addition, the central part of the C-terminal helix shows the same level of unfolding (~10%) as the Met80 loop and 60's helix in the oxidized protein. This is in
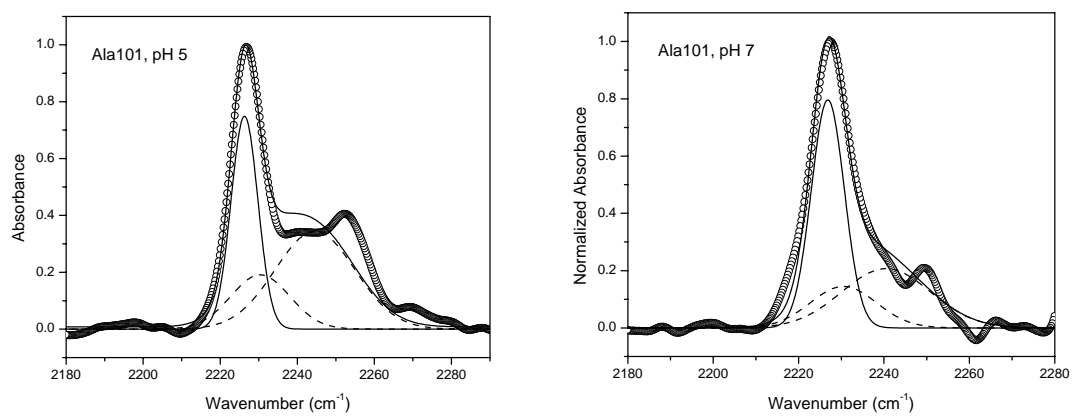
Figure 3.11. Ala101 data at pH 5 and pH 7, showing unfolded spectra (dashed) present at both pH values. At pH 5, it is estimated that 58% unfolded protein is present while at pH 7 only 35% is unfolded.

contrast with other studies carried out under similar conditions that suggested the C-terminal helix was generally one of the most stable parts of the protein (Filosa, 2001). Unexpectedly, reduction of the protein selectively stabilizes both the Met80 loop and the 60's helix, resulting in the C-terminal helix actually being less stable than other parts of the protein. This marked destabilization of the C-terminal helix in the reduced protein has not been previously observed by other techniques.

Overall, the data suggest that there are no redox-dependent changes in flexibility but rather changes in the amount of locally unfolded protein present. As previous studies have suggested, the folded structure of cytochrome c is shown to exist in dynamic equilibrium with partially unfolded states. But particularly noteworthy, is the extent of C-terminal fraying and reduced stability of the C-terminal helix in the reduced state which indicates that even under standard conditions, the structure of the protein is less well characterized than previously thought. Perhaps the local folding of cytochrome c is linked to its biological function, as has been suggested for other proteins (Demarest, 2002).

This chapter, in part, has been submitted for publication as a communication to the Journal of the American Chemical Society. Permission to include the results in this thesis has been granted by the journal and the co-authors of the paper.

3.5 References

Bai, Y. W., Sosnick, T. R., Mayne, L. and Englander, S. W., Protein Folding Intermediates – Native State Hydrogen Exchange. *Science,* **1995**, *269*, 192.

Barker, P. D., Bertini, I., Del Conte, R., Ferguson, S.J., Hajieva, P., Tomlinson, E., Turano, P., and Viezzoli, M. S. A further Clue to Understanding the Mobility of Mitrochondrial Yeast Cytochrome c – A N-15 T-1 rho Investigation of the Oxidized and Reduced Species. *European Journal of Biochemistry*, **2001**, *268*, 4468.

Baxter, S. M., and Fetrow, J. S., Hydrogen Exchange Behavior of [U-^{15}N]- Labeled Oxidized and Reduced Iso-1-cytochrome *c*. *Biochemistry*, **1999**, *38*, 4493.

Berghuis, A. M., Brayer, G. D., Oxidation State-Dependent Conformational Changes in Cytochrome *c*. *J. Mol. Biol*., **1992**, *223*, 259.

Berghuis, A. M., Guillemette, J. G., Smith, M., Brayer, G. D. Mutation of Tyrosine-67 to Phenylalanine in Cytochrome c Significantly Alters the Local Heme Emvironment. *J. Mol. Biol*., **1994**, *235*, 1326.

Brayer, G. D., Murphy, M. E. P., Cytochrome c: A Multidisciplinary Approach; Scott, R. A., Mauk, A. G., eds, University Science Books, Sausalito, CA, 1995.

Bushnell, G. W., Louie, G. V., Brayer, G. D., High-Resolution 3-Dimensional Structure of Horse Heart Cytochrome *c*. *J. Mol. Biol*., **1990**, *214*, 585.

Chin, J. K., Jimenez, R., and Romesberg, F. E., Direct Observation of Protein Vibrations by Selective Incorporation of Spectroscopically Observable Carbon-Deuterium Bonds in Cytochrome *c*. *J. Am. Chem. Soc.* **2001,** 123, 2426.

Colon, W. Wakem, L. P., Sherman, F., and Roder, H. Identification of the Predominant Non-Native Histidine Ligand in Unfolded Cytochrome c. *Biochemistry*, **1997**, *36*, 12535.

Colon, W., Elove, G. A., Wakem, L. P., Sherman, F. and Roder, H. Side Chain Packing of the N- and C- Terminal Helices Plays a Critical Role in the Kinetics of Cytochrome *c* Folding. *Biochemistry*, **1996**, *35*, 5538.

Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, J., Evans, R., and Wright, P. Mutual Synergistic Folding in Recruitment of CBP/p300 by p160 Nuclar Receptor Coactivators. *Nature*, **2002**, *415*, 549.

Eaton, W. A., Henry, E. R., Hofrichter, J., and Mozzarelli, A., Is Cooperative Oxygen Binding by Hemoglobin Really Understood? *Nature Structural Biology*, **1999**, *6*, 351.

Filosa, A., Wang, Y., Ismail, A. A., and English, A. M., Two-Dimensional Infrared Correlation Spectroscopy as a Probe of Sequential Events in the Thermal Unfolding of Cytochromes *c*. *Biochemistry*, **2001**, *40*, 8256.

Goto, Y., Hagihara, Y., Daizo, H., Hoshino, M., Nishii, I., Acid-Induced Unfolding and Refolding Transitions of Cytochrome *c* – A 3-State Mechanism in H2O and D2O. *Biochemistry*, **1993**, *32*, 11878.

Hilgen, S. E., and Pielak, G. J., The Function of the Sacharomyces-Cerevisiae Iso-1-Cytochrome *c* Gene is Independent of the Codon at Invariant Residue Phe82 When the Gee is Present on a Low Copy Number Vector. *J. Prot. Engineer*., **1991**, *4*, 575.

Mauk, A. G., Coyle, C. L., Bordignon, E., Gray H. B. Bis(Dipicolinate) Complexes of Cobalt(III) and Iron(II) as New Probes of Metalloprotien Electron Transfer Reactivity – Analysis of Reactions Involving Cytochrome c and Cytochrome C551. *J. Am. Chem. Soc*., **1979**, *101*, 5054.

Qi, P. X., Beckmn, R. A., Wand, A. J. Solution Structure of Horse Heart Ferricytochrome *c* and Detection of Redox-Related Structural Changes by High-Resolution H-1 NMR. *Biochemistry*, **1996**, *35*, 12275.

Schnölzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B., In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *International Journal of Peptide and Protien Research*, **1992**, *40*, 180.

Snedecor, G. W., Cochran, W. G. *Statistical Methods, Eighth Edition*. Iowa State University Press, Ammes, Iowa, 1989.

Tezcan, F. A., Winkler, J. R., and Gray, H. B., Effects of Ligation and Folding on Reduction Potentials of Heme Proteins. *J. Am. Chem. Soc.*, **1998**, *120*, 13383.

Voet, D., Voet, J., *Biochemistry, 2nd Edition*. John Wiley & Sons, Inc, Somerset, New Jersey, 1995.

Wallace, C. J. A., and Clark-Lewis, I., Functional Role of Heme Ligation in Cytochrome *c*. *J. Biol. Chem.,* **1992,** *267*, 3852.

Wooten, J., Cohen, J., Vig, I. and Schejter, A., PH-Induced Conformational Transitions of Ferricytochrome *c* – A C-13 and Deuterium Nuclear Magnetic Resonance Study. *Biochemistry*, **1981**, *20*, 5394.

CHAPTER 4. A HIGH RESOLUTION PROBE OF PROTEIN FOLDING:

MEASURING PROTEIN FOLDING AT TWO RESIDUES

IN CYTOCHROME *C*

4.1 Introduction

Protein folding is a key step in the manifestation of biological function from genetic information, since the fold of a protein directly impacts its function. Understanding the folding process, even for a relatively small protein, has emerged as a significant, but challenging, problem. For instance, it is still not well understood how concerted the folding process is and to what extent the folding of certain regions of the protein are affected by nearby residues. A few small proteins, such as myoglobin and cytochrome c, have shown themselves to be amenable to folding studies, either through ease of sample handling or by containing tractable chromophores. Thus, cytochrome *c,* being one of the most well studied proteins in the context of protein folding, has been chosen for the initial application of the technique.

The reversible equilibrium folding of cytochrome *c* has been characterized using several indigenous chromophores, and unfolding induced by guanidine hydrochloride has been characterized intensively. Plots of various spectroscopic signals versus denaturant concentration typically show an apparent two-state transition to the unfolded state, with a midpoint at ~2.5 M guanidine hydrochloride (Tsong, 1975, Segel, 1998, Hagihara, 1994). However, resonance Raman signals from the heme and tryptophan residue, when unfolded in guanidine hydrochloride, show different transitions, indicating the presence of an intermediate (Rush, 1992). Equilibrium amide exchange NMR studies also indicate the presence of at least two partially folded states during titration with guanidine hydrochloride (Bai, 1995). At least one

equilibrium unfolding intermediate has been detected by small angle x-ray scattering and based on magnetic circular dichroism data, this intermediate is thought to involve bis-histidinyl coordination by His18 and His26 or His33 (Thomas, 2000). Thus, a consensus regarding the presence and nature of equilibrium partially unfolded intermediates has not yet been reached due, at least in part, to the differences in techniques used. The lack of a technique that can offer site-specific detail within the protein and the time resolution to detect rapidly converting intermediates has made characterization of partially folded equilibrium intermediates difficult.

In principle, the folding pathway of a protein could be fully characterized by infrared spectroscopy. However, the spectral congestion inherent to proteins and denaturants has prevented the direct use of infrared spectroscopy for this task. Interestingly, protein and denaturants have a 'transparent window' in their infrared spectrum that is free of absorptions, between ~1800 and 2700 cm^{-1}. Any sufficiently strong absorption in this region may be directly observed, despite high protein or denaturant concentrations. Thus, to develop a nonperturbative and residue-specific probe, the Romesberg group is currently examining the use of carbon-deuterium bonds, which absorb at ~2100 cm^{-1}. In addition to spectral resolution and high molecular resolution, these probes provide inherently high time resolution to detect rapidly converting intermediates.

A site-specific investigation of the equilibrium unfolding of cytochrome c was carried out by monitoring two deuterated residues, d_3-Leu98 and d_3-Met80. Figure 4.1 shows the location of these two residues in cytochrome c. Leu98 is located in the C-
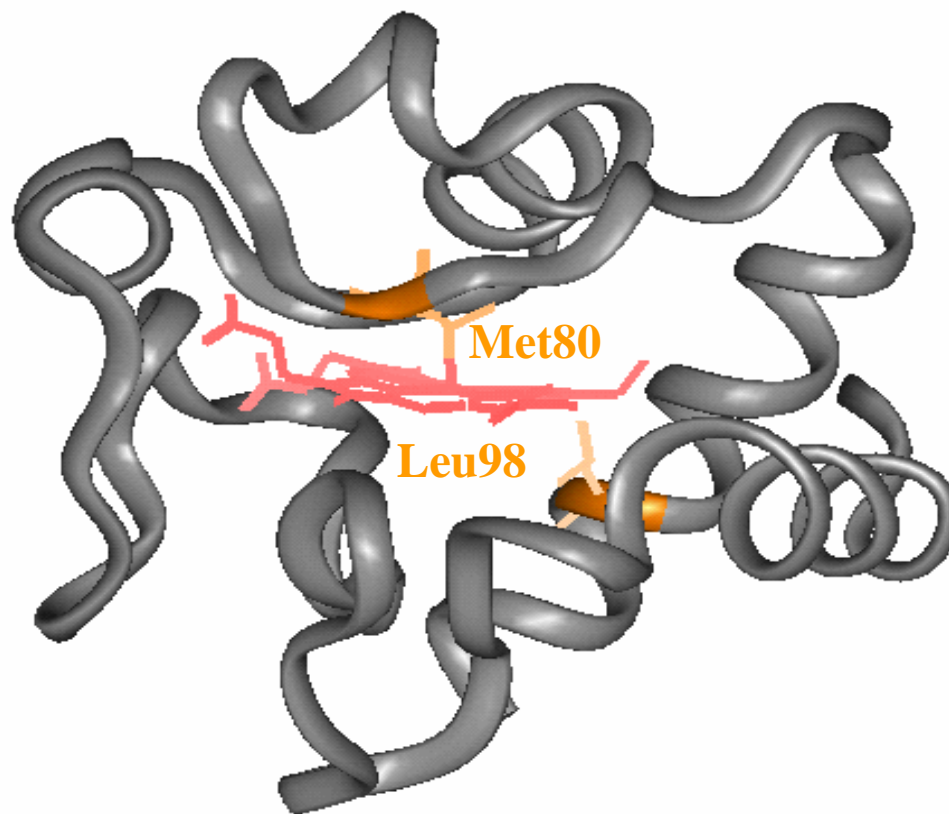
Figure 4.1. The x-ray crystal structure of horse heart cytochrome c showing the sites of deuteration. The Met80 residue is the sixth ligand to the heme and Leu98 is located in the C-terminal helix.

terminal helix, which NMR studies have shown to be one of the most stable motifs in the protein, unfolding only at high denaturant (Bai, 1995). Met80 is one of the protein-based heme ligands, whose loop (the D loop) is thought to be the least stable motif in the protein, unfolding at low denaturant concentration. The Met80 ligand has also been shown to be involved in misligated partially unfolded intermediates, which studies report can exist both at low and high denaturant (Russell, 2002).

In addition, we report the comparison of infrared data on the d_3-Met80 residue, a local probe for the Met80 ligand, with that of the 695 nm band in the UV-Vis spectrum. The 695 nm band is assigned as a charge transfer band between the sulfur atom in the Met80 ligand and the iron atom in the heme. Since the earliest folding studies of cytochrome c, the 695 nm band has been used as a reporter of whether the Met80 ligand is bound to the heme (Kaminsky, 1973). Interestingly, although the 695 nm band is believed to be a site-selective probe of Met80 ligation, it has seldom been invoked in more recent studies implicating misligation events. Moreover, early studies have indicated that the 695 nm band is sensitive to more than just Met80 ligation, with contributions from the heme pocket and geometry of the heme (Schejter, 1964). Thus, a comparison of the guanidine hydrochloride titrations of two probes for the Met80 ligand should help to assess the sensitivity of the 695 nm band.

4.2 Materials and Methods

Sample Preparation. As described in Chapter 2, the site-selective deuterated cytochrome *c* was prepared through its semi-synthesis. Briefly, a cytochrome *c* 1-65

homoserine lactone peptide fragment containing the covalently bound heme was generated using CNBr cleavage at the Met65 residue of horse heart cytochrome *c*. This peptide was purified and then refolded at pH 7 in the presence of one equivalent of chemically synthesized residues 66-104 of cytochrome *c* containing a specifically deuterated leucine or methionine residue. Fragment association and aminolysis of the homoserine lactone resulted in high yields of semisynthetic cytochrome *c* with the slight modification of a homoserine at position 65 instead of a methionine. This mutation has been shown to have no effect the redox potential, function and common folding markers of the protein. The final condensation product was purified from the fragments, washed with 10 mM sodium acetate buffer at pH 6.2, divided on a per sample basis, and then lyophilized. Samples were stored at -20 ºC until ready for use. Guanidine hydrochloride solutions were made by dissolving the appropriate amount of solid guanidine hydrochloride in 5 ml of 200 mM sodium acetate buffer, pH 6.2. Doubly distilled water was then added to reach 10 ml and the pH of each solution checked individually and adjusted with 0.1 M NaOH. Guanidine hydrochloride solutions were stored on a per use basis at -80 ºC and only taken out of the freezer 2 hours prior to use. Lyophilized protein samples were dissolved in the appropriate guanidine hydrochloride solution and allowed to stand approximately ten minutes prior to measurement.

FTIR Measurements. Samples were measured using a Bruker Equinox 55 FTIR spectrometer equipped with a liquid nitrogen cooled MCT detector, KBr beamsplitter and the sample compartment continuously flushed with dry nitrogen. Spectra were

collected at 4 cm^{-1} resolution and constructed from 8192 scans. The Blackman-Harris 3-term apodization function was used, a 16 kHz low pass filter and zero filling of 16. The aperture setting was 2000 μm and the liquid transmission cell consisted of CaF$_2$ disks 32 mm in diameter with a Teflon spacer 50 μm thick. For every deuterated sample, a proteo sample was measured directly afterwards and subtracted using a subtraction factor close to unity. After proteo subtraction, the baseline was corrected by fitting a window of at least 100 cm^{-1} to a polynomial and subtracting the polynomial to result in a relatively flat baseline with the peaks of interest.

   UV-Vis Measurement of the 695 nm Band. UV-Vis measurements were carried out on a Cary 300 Bio spectrometer equipped with a temperature bath set at 25 °C. Horse heart cytochrome *c* samples for the UV-Vis measurements were purchased from Sigma and equilibrated in 10 mM sodium acetate buffer, pH 6.2 and lyophilized. The lyophilized protein samples were then dissolved in the appropriate guanidine hydrochloride solution (the same solutions as those used in the FTIR measurements) and allowed to equilibrate for at least ten minutes prior to measurement. The 695 nm band was quantified by taking the spectral window from 550 to 750 nm and flattening the baseline with a linear spline correction using the Genplot program.   Once the baseline was flattened and zeroed a curser was used to obtain the absorbance value at 695 nm. This absorbance value was then divided by the concentration and path length of the cuvette to obtain the extinction coefficient at each guanidine hydrochloride concentration. The concentration was calculated from the absorbance at 528 nm using the extinction coefficient of 10,400 M^{-1} cm^{-1}.

Data Analysis. The data was fit to Gaussians representing chemically meaningful

species as described in Chapter 2. The error reported for the midpoint values are

calculated from fitting three independent data sets. For d_3-Leu98, it was determined

that the folded spectra, acquired at 0 M guanidine hydrochloride, could be fit using

two Gaussians.   Likewise the unfolded spectra, acquired at 6 M guanidine

hydrochloride as well as the free amino acid in buffer, could also be represented by

two Gaussians.   However, as discussed in Chapter 2, the two Gaussians in the

unfolded spectra are believed to represent the two asymmetric stretches and the two

Gaussians in the folded protein most likely reflect overlapping asymmetric stretches of

the two isotopomers. In order to simplify the fitting process, several parameters were

fixed.   For the folded spectra of d_3-Leu98, the relative amplitudes of the two

Gaussians, and the frequencies were fixed. For the unfolded spectra, the relative

amplitudes of the two Gaussians were the only thing fixed. The frequencies of the

unfolded spectra were somewhat restricted to a 4 cm^{-1} window, but could not be

completely restricted due to salt effects causing a slight red shift in the unfolded

frequencies. The linewidths of both the folded and unfolded Gaussians were not held

fixed.   It was ensured that all fits to the spectra at each guanidine hydrochloride

concentration were the best possible fits to the data, yielding the lowest possible least

squared error sum.   Due to the low intensity of the symmetric stretch, only the

asymmetric stretches are reported in this study. Generation of the plots of fraction

folded/unfolded versus denaturant were carried out by first normalizing the intensities

of all spectra to an isosbestic point. An isosbestic point is created when two absorbing

species overlap and offers further proof of a two-state transition (see Figures 5.2 - 5.27). Subsequent normalization was carried out by dividing by the sum of all absorbing species and multiplying by the ratio of extinction coefficients.

The d_3-Met80 spectra measured fit well to one Gaussian in both the folded and unfolded states. The single absorption for d_3-Met80 is in accord with this peak being assigned as the symmetric stretch. The parameters of the folded and unfolded Gaussians were similarly determined from data at 0 M and 6 M guanidine hydrochloride. An attempt to fit the data to a sum of folded and unfolded states yielded a partial unfolding at low guanidine hydrochloride concentrations. It was then revealed that the best fit to the data at low guanidine hydrochloride concentration was in fact not a sum of the folded and unfolded states, but a sum of the folded and another state with slightly different spectral features than the unfolded state. The fits generating the final plot shown in Figure 4.3, arose from fixing the folded Gaussian frequency and restricting the unfolded and intermediate Gaussian frequencies to a 4 cm^{-1} window. Also due to the low signal to noise ratio for the asymmetric stretches, only the symmetric stretch is reported.

4.3 Results

Unfolding of Leu98. As shown in Figure 4.2, each d_3-Leu98 spectra can be fit to a sum of Gaussians representing the folded and unfolded states. The unfolded state grows in, while the folded state decreases as the guanidine hydrochloride concentration is increased. The normalized intensities were calculated by first
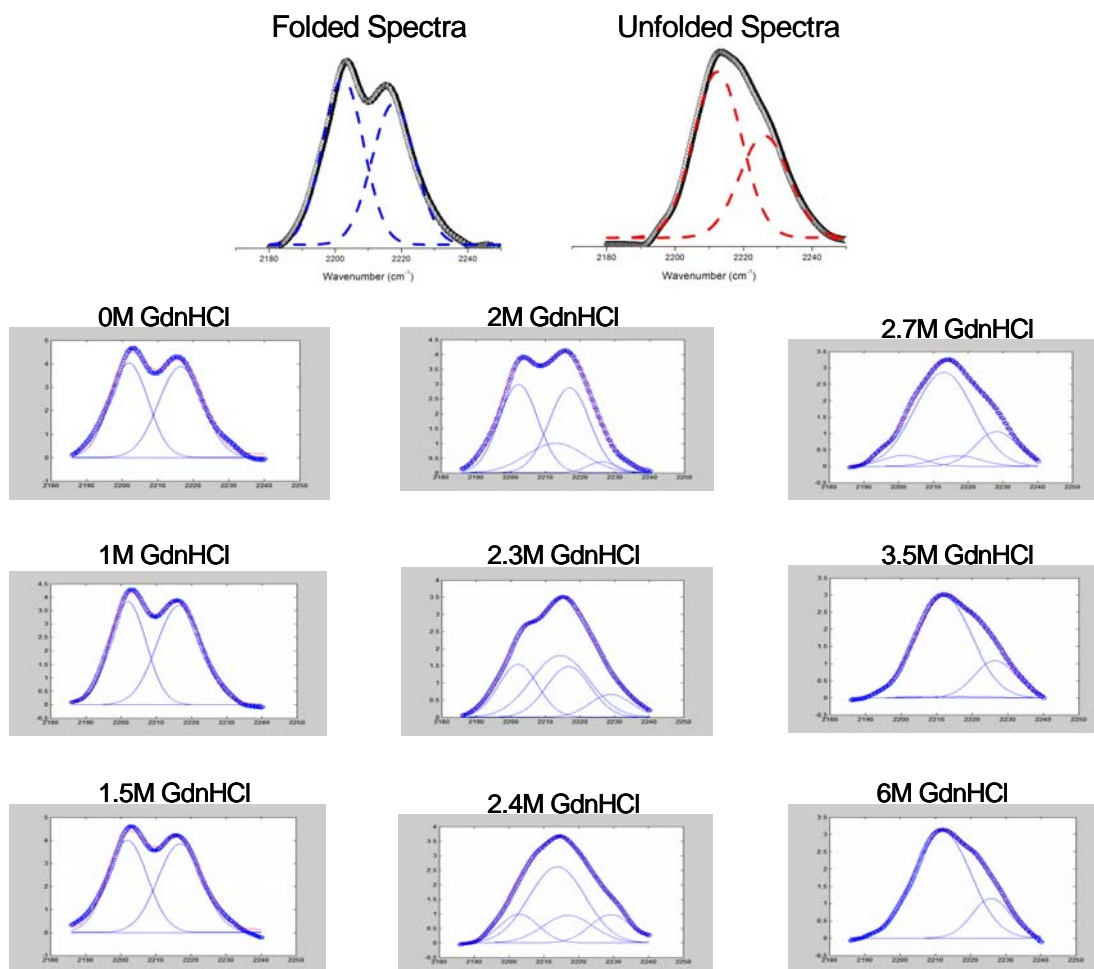
Figure 4.2. Fitting the *d₃*-Leu98 data to a sum of the folded and unfolded spectra. The spectra with approximately half folded and half unfolded is taken at 2.3M guanidine hydrochloride, which is the midpoint value obtained when fitting all the spectra in the titration curve.
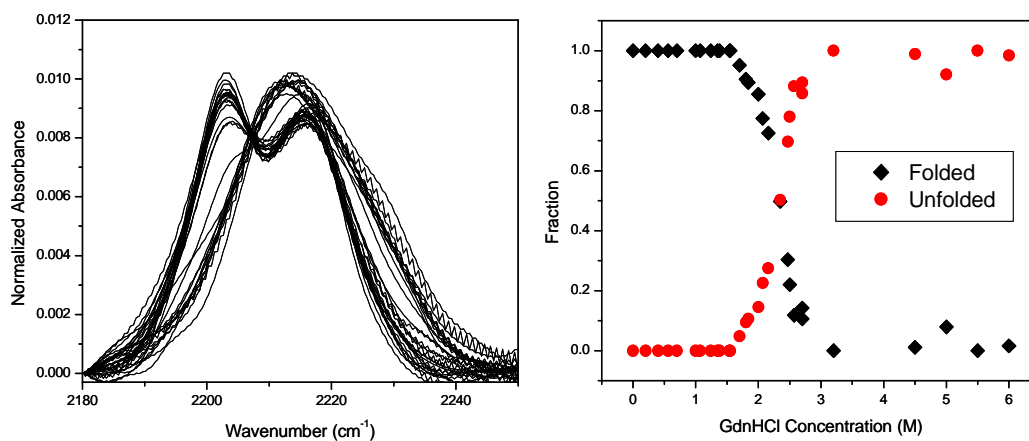
Figure 4.3.   Titration of d_3-Leu98 with increasing guanidine hydrochloride concentration. An isosbestic point at 2208 cm^{-1} is easily identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride shows a two-state transition with a midpoint value of 2.3 ± 0.1 M.

dividing all the individual spectra by the area underneath the peaks of interest. A plot of the resulting spectra reveal an isosbestic point at 2208 cm^{-1}. It was then ensured that all individual spectra had the same intensity at the isosbestic point. After the spectra are normalized for overall intensity using the isosbestic point, the intensities of the individual peaks are divided by the sum of the intensities of all species present. For example, since the folded state consists of two Gaussians, the amplitude of the low frequency Gaussian is monitored as the folded state and divided by the sum of the intensity of that Gaussian and the low frequency unfolded Gaussian at the same guanidine hydrochloride concentration. A final normalization step accounts for differences in extinction coefficient of the low frequency folded and unfolded Gaussians. The resulting plot of normalized intensies vs guanidine hydrochloride concentration yields a sigmoidal transition with a midpoint of 2.3 ± 0.1 M guanidine hydrochloride, as shown in Figure 4.3.

Unfolding of Met80. The d_3-Met80 spectra contained folded and unfolded spectra best approximated with a single Gaussian. An attempt was made to fit the data similar to that of d_3-Leu98, where each individual spectrum was fit by a sum of the folded and unfolded Gaussians. As shown in Figure 4.4, this did not yield a single sigmoidal curve but the appearance of two sigmoidal transitions, one at low guanidine hydrochloride concentration and one close to 2.3 M guanidine hydrochloride. An investigation of the unfolded Gaussian that appeared at low guanidine hydrochloride concentration revealed that the frequency and linewidth were slightly different than that of the unfolded Gaussian at high guanidine hydrochloride concentrations.
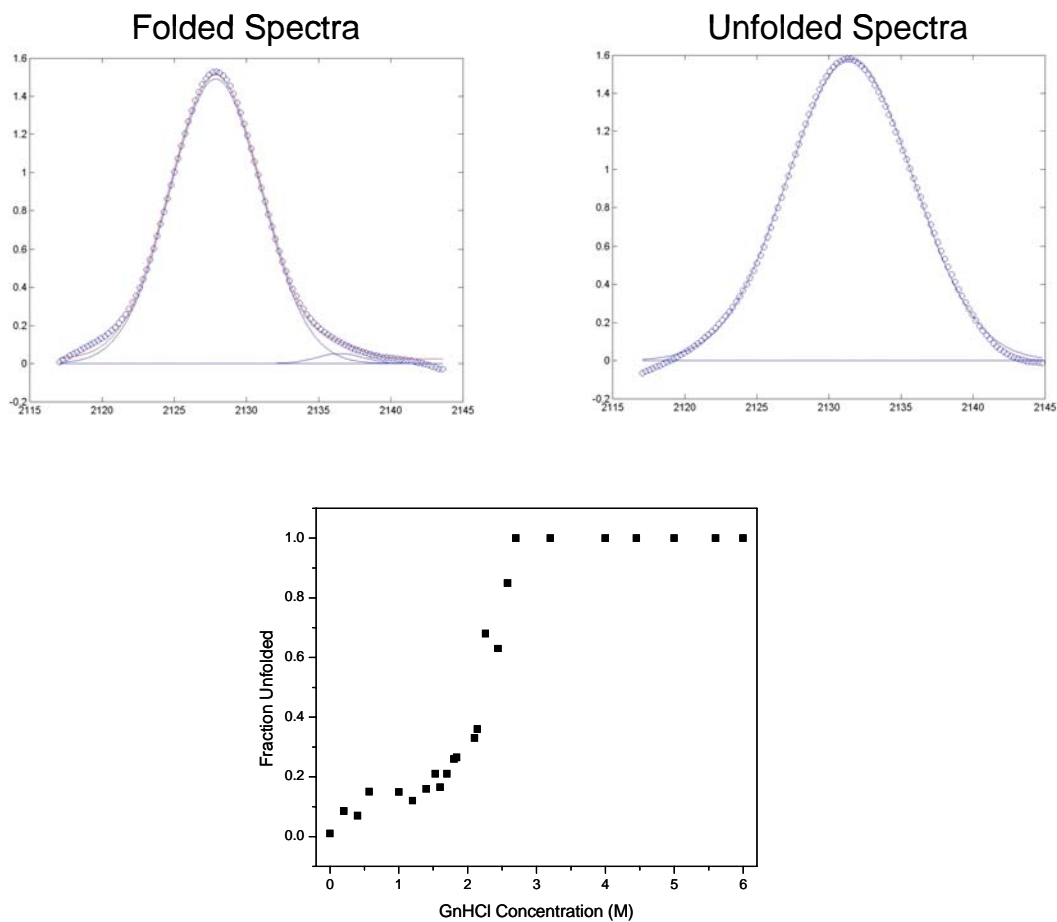
Figure 4.4. An attempt to fit the d_3-Met80 data to a two-state model containing only folded and unfolded spectra. The double sigmoidal curve that resulted implied a third state and prompted a second look at the spectra at low guanidine concentrations.

Therefore, a third Gaussian was added with these slightly different spectral features representing an intermediate species. Each spectrum with increasing guanidine hydrochloride concentration was fit to three Gaussians as shown in Figure 4.5. An isosbestic point was identified at 2130.5 cm^{-1} by plotting the individual spectra together, normalized to peak area in a narrow spectral window only including the folded and unfolded states. The resulting intensities of the folded, unfolded and intermediate states were then divided by the sum of all three intensities and the folded and unfolded states normalized for differences in extinction coefficient. Since the intermediate species is never present to full extent along the titration curve, the extinction coefficient of this species is unknown and no further normalization was carried out with the intensity of that species. Thus the intensities plotted are estimated values for the intermediate species. As shown in Figure 4.6, the folded and unfolded states decrease and increase respectively with increasing guanidine hydrochloride concentration showing a midpoint of 2.3 M guanidine hydrochloride. The intermediate species grows in and out, as one would expect, with a maximal population at 1.2 M guanidine hydrochloride. Similar to the d_3-Leu98 data, the titration was repeated at least three times to ensure reproducibility and the reported midpoints have standard deviation of 0.05 M guanidine hydrochloride.

The Unfolding of the 695 nm Band. Absorbance values were measured at guanidine hydrochloride concentrations similar to those measured for d_3-Met80 and Leu98 and extinction coefficients calculated at each concentration along the titration
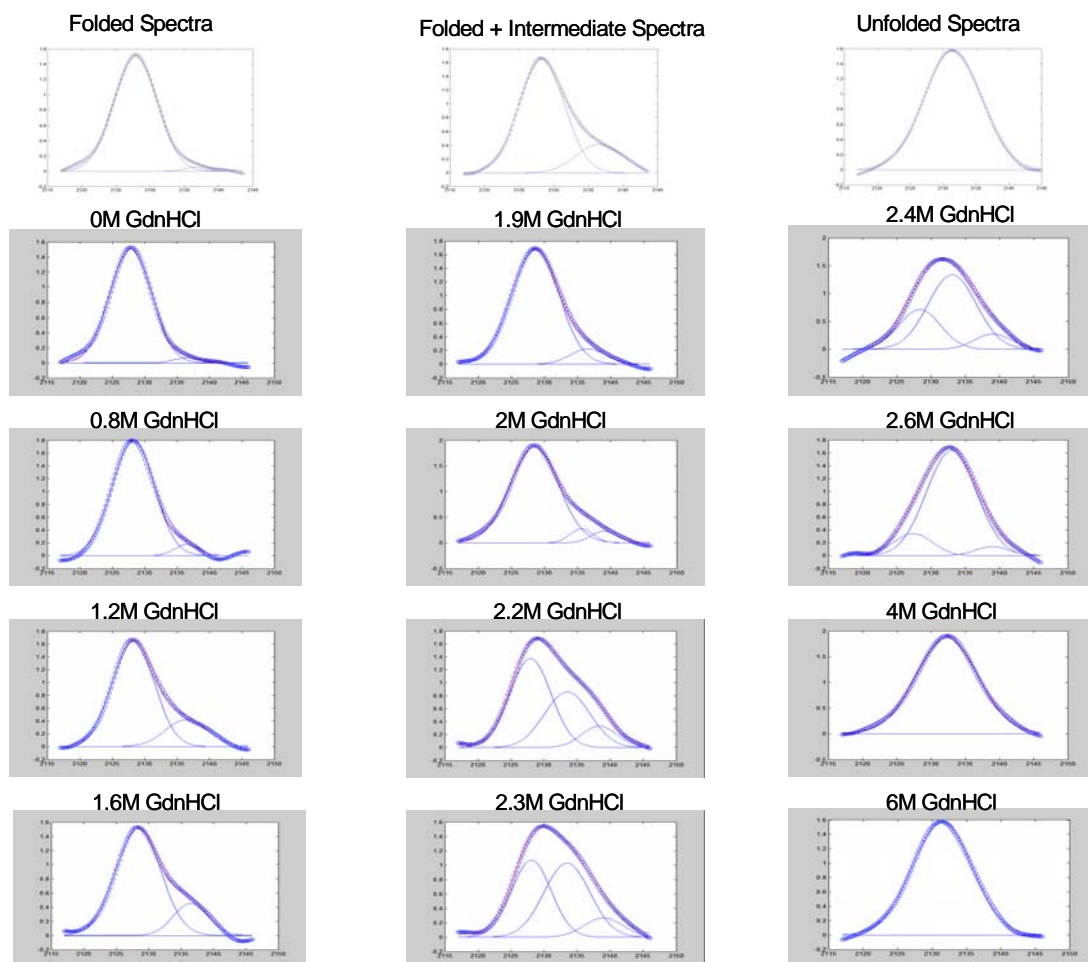
Figure 4.5. Fitting the d_3-Met80 data to three Gaussians representing the folded, intermediate and unfolded states. The intermediate state is shown to reach a maximum at low guanidine hydrochloride concentrations whereas the unfolded state reaches a maximum at high guanidine hydrochloride concentrations.
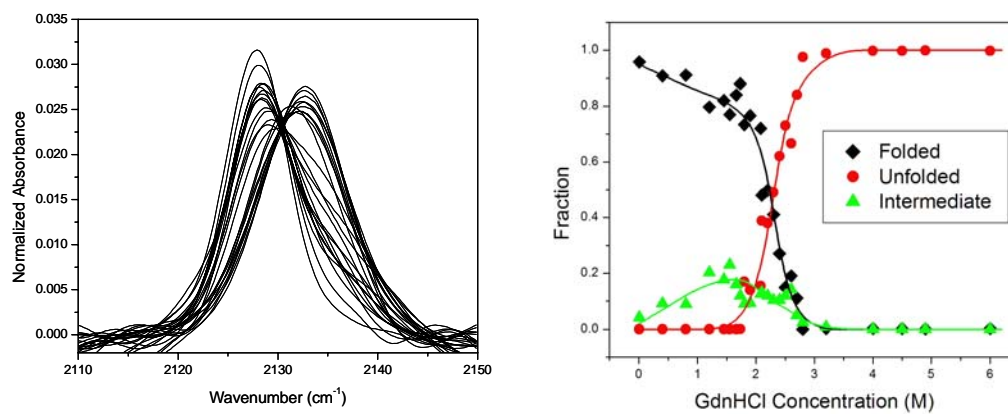
Figure 4.6.   Titration of d_3-Met80 with increasing guanidine hydrochloride concentration. An isosbestic point at 2030.5 cm^{-1} is easily identified when plotting all the spectra normalized to the area underneath the peaks of interest.   This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride shows the presence of a partially unfolded intermediate at low guanidine hydrochloride concentrations and a global unfolding transition with a midpoint value of
2.3 ± 0.1 M.

Figure 4.7. Titration of both the 695nm band and d_3-Leu98 with increasing guanidine hydrochloride concentration. Both show a global transition at 2.3M and the lack of an intermediate at low guanidine hydrochloride concentration.

curve. A plot of extinction coefficient versus guanidine hydrochloride concentration was generated and shows a single transition at 2.3 M guanidine hydrochloride (see Figure 4.7). The reproducibility of the titration curve was checked by repeating the measurement three times, generating a standard deviation of 0.1 M guanidine hydrochloride.

4.4 Discussion and Conclusions

The observation that the two-state transition observed at d_3-Leu98 shares a common unfolding midpoint with d_3-Met80 is a clear indicator of a global unfolding event. Although the majority of equilibrium folding data supports the cooperative unfolding of cytochrome c, recent amide-exchange NMR data indicates a sequential mechanism, with the Met80 loop having a lower ΔG° of unfolding than the C-terminal helix (Xu, 1998). However, chemical exchange happens on relatively slow timescales compared to infrared spectroscopy (milliseconds vs femtoseconds) and would not be able to resolve a rapidly converting substructure of the protein, especially if the structure is solvent exposed. The fact that an intermediate was observed along with cooperative behavior offers a simple explanation for the wealth of contradictory studies which on one side imply cooperative folding behavior, and on the other suggest a sequential folding pathway. In addition, this study illustrates a residue-specific structural characterization of a partially unfolded intermediate, which has proven difficult to detect by other biophysical techniques. Based on the two residues

measured, the partially unfolded intermediate appears to involve only local structural rearrangement.

The d_3-Met80 data reveal a partially unfolded intermediate present in equilibrium with the folded state at low concentrations of guanidine hydrochloride that is not present at the d_3-Leu98 residue. An intermediate reaching maximum population at similar guanidine hydrochloride concentrations was also observed by Russell *et al* using ^{1}H NMR spectroscopy of paramagnetic heme substituents (Russell, 2002). A rigorous investigation of the non-native resonances at low guanidine hydrochloride concentration was carried out using 1D NOE, comparison with alkaline resonances, and chemical modification of all 19 lysine residues in the protein. This analysis showed that the intermediate is similar to that observed for the alkaline species in which misligation of the Met80 ligand occurs. Chemical modification implicates Lys72, Lys73, and Lys79 as being possible candidates for heme ligation instead of Met80 in the intermediate structure. Since these data were taken under similar conditions to those of Russell *et al*, this intermediate most likely involves Met80 misligation. The high frequency of the intermediate spectrum is also indicative of a Met80 residue no longer bound to the heme. Interestingly, the 695 nm band, thought to be indicative of an intact Met80 ligand, shows no sign of an intermediate at low guanidine hydrochloride concentration. Although a similar percentage of molecules present in the intermediate configuration is estimated from both our IR data and the data of Russell *et al*, the 695 nm UV-Vis band data suggests that this population might be smaller. The intensity of the 695 nm band may also be sensitive to the structure of

the heme pocket and geometry of the heme itself (Schejter, 1964). Since the 695 nm band is suggested to be sensitive to many factors and the relative contribution of each factor to the observed intensity is difficult to discern, it is therefore difficult to assign a unique cause for the lack of misligated intermediate in the 695 nm band. Perhaps the misligated intermediate does not involve much structural rearrangement of the heme or heme pocket. A further investigation of the partially folded intermediate in urea will be reported in Chapter 5.

This chapter, in part, has been published by the Journal of American Chemical Society and permission has been granted by the journal and co-authors to include the results of the article in this thesis. The following is a citation of the article:

Sagle, Laura B., Zimmermann, Jörg, Dawson, Philip E., and Romesberg, Floyd E., "A High Resolution Probe of Protein Folding" *Journal of the American Chemical Society*. **2004**, 126(11), 3384-3385.

4.5 References

Bai, Y. W., Sosnick, T. R., Mayne, L., and Englander, S. W., Protein Folding Intermediates – Native State Hydrogen Exchange. *Science*, **1995**, *269*, 192.

Hagihara, Y., Tan, Y. and Goto, Y., Comparison of the Conformational Stability of the Molten Globule and Native State of Horse Cytochrome *c* – Effects of Acetylation, Heat, Urea, and Guanidine Hydrochloride. *J. Mol. Biol*., **1994**, *237*, 336.

Kaminsky, L. S., Miller, V. J., and Davison, A. J., Thermodynamic Studies of the Opening of the Heme Crevice of Ferricytochrome *c*. *Biochemistry*, **1973**, *12*, 2215.

Rush, T. S., and Spiro, T. G. *Spectroscopic Methods in Bioinorganic Chemistry; ACS Symposium Series 692*, American Chemical Society, Washington, DC, 1992, pp 212-219.

Schejter, A., and George, P., The 695 mμ Band of Fericytochrome *c* and Its Relationship to Protein Conformation. *Biochemistry*, **1964**, *3*, 1045.

Segel D. J., Fink, A. L., Hodgson, K. O., and Doniach, S., Protein Denaturation: A Small Angle X-Ray Scattering Study of the Ensemble of Unfolded States of Cytochrome *c*. *Biochemistry*, **1998**, *37*, 12443.

Thomas, Y. G., Goldbeck, R. A., and Kliger, D. S., Characterization of Equilibrium Intermediates in Denaturant-Induced Unfolding of Ferrous and Ferric Cytochromes *c* using Magnetic Circular Dichroism, Circular Dichroism, and Optical Absorption Spectroscopies. *Biopolymers*, **2000**, *57*, 29.

Tsong, T. Y., Acid-Induced Conformational Transition of Denatured Cytochrome *c* in Urea and Guanidine Hydrochloride Solutions. *Biochemistry*, **1975**, *14*, 1542.

Xu, Y., Mayne, L. and Englander, W. S., Evidence for an Unfolding and Refolding Pathway in Cytochrome *c*. *Nature Structural Biology*, **1998**, *5*, 774.

CHAPTER 5. A HIGH RESOLUTION PROBE OF PROTEIN FOLDING: MEASURING PROTEIN FOLDING OF DIFFERENT STRUCTURAL MOTIFS IN CYTOCHROME *C*

5.1 Introduction

Understanding the protein folding process has emerged as one of the most significant and challenging problems in the biological sciences. Through the years, several models have been developed to describe how proteins fold to their native state. In the 1980's, studies on small peptides revealed that in solution a peptide of a given primary sequence folds to the proper secondary structural element, such as an α helix or a β sheet (Shoemaker, 1985). This lead to the framework model of protein folding, where it was thought that as a protein folds, secondary structure forms first which then guides the formation of tertiary structure leading to the final native conformation. However, more recent experimental studies have now shown that secondary and tertiary structure forms simultaneously during the folding of many proteins (Jackson, 1991 and Otzen, 1994). The next model to gain popularity was brought about by experimental thermodynamic as well as theoretical studies that revealed early steps in protein folding most likely involve a hydrophobic collapse, where a substantial amount of energy is gained by the exclusion of water and the close proximity of hydrophobic residues within the protein (Tanford, 1962). Although there is an abundance of evidence that a hydrophobic collapse does indeed take place early in the protein folding process, a non-specific hydrophobic cluster presents the problem of large reorganization energies needed to make the appropriate native contacts (Baldwin, 1989 and Ptitsyn, 1987). The current model of protein folding incorporates concepts from both earlier models. This model, referred to as the nucleation-

condensation model, involves a hydrophobic collapse where some native-like contacts are formed. These native-like contacts then guide the formation of more native contacts along the folding pathway, until the folded structure is obtained (Fersht, 1996 and Uversky, 2002).

Indeed, folding studies of oxidized cytochrome *c* have played a major role in shaping the current model of how proteins fold. Kinetic studies have shown that the protein does undergo a hydrophobic collapse early in the folding process, and some native-like contacts exist in this collapsed state (Shastry, 1998). The collapsed state is thought to resemble a molten globule-like state where both secondary and tertiary structure has been shown to exist in the N and C-terminal helices (Colon, 1996 and Akiyama, 2000). Furthermore, a correlation can be made between energetic stability and the order in which different structures are formed. The N and C-terminal helices of the protein have been shown to be the most energetically stable structures in the protein and are the first to fold (Xu, 1998). As folding continues, less energetically stable, partially folded intermediates have been detected, which for cytochrome *c* have generally involved heme misligation and proline isomerization (Elove, 1994). Although great gains have been made in our understanding of protein folding, due to the experimental challenges of characterizing partially unfolded intermediates, questions still remain about the nature of these intermediates and the role they play in the folding of the protein. Moreover, recent studies have shown that proteins in fact exist in dynamic equilibrium with many substructures even in the native state

(Williamson, 2003). The dynamic nature of protein structure and the influence it has throughout the folding process has not fully been explored.

The folding of oxidized cytochrome *c* has been studied using a variety of techniques and under equilibrium conditions, partially unfolded intermediates have been shown to exist. Circular dichroism, UV-Vis and tryptophan flouresense studies all indicate a lack of partially unfolded intermediates under equilibrium conditions (Myer, 1968). Magnetic circular dichroism results, however, indicate the presence of an intermediate at high denaturant, most likely involving heme misligation, as well as the possibility of an intermediate at low denaturant that is thought to have a molten globule-like structure (Thomas, 2000). Studies employing small angle x-ray scattering have indicated an intermediate accumulates at moderate denaturant concentrations, which is structurally similar to the unfolded state, but contains some native contacts (Cinelli, 2001). Flourescence energy transfer experiments suggest that a compact intermediate persists even at high denaturant (Lyubovitsky, 2001). Recent NMR studies indicate two intermediates are populated during the equilibrium folding of cytochrome *c*, one at low denaturant concentration and one that persists under strongly denaturing conditions (Russell, 2002). Both of these intermediates were attributed to heme misligation. Finally, amide-exchange NMR and 2-D FTIR studies indicate the existence of at least two partially folded intermediates, which were ascribed to the sequential folding of secondary structural elements (Bai, 1995 and Filosa, 2001).

More confusion is created concerning the folding pathway and existence of partially folded intermediates of oxidized cytochrome *c* when one considers the

variety of conditions employed. Under acidic conditions the formation of a compact molten globule-like state is observed, which is not present to a high degree or easily observed at higher pH (Jeng, 1990). The decreased stability of the Met80 loop observed at neutral and high pH is not observed under acidic conditions where the molten globule state is thought to possibly retain Met80 ligation of the heme when the iron is in its low spin form (Goto, 1990). There are also discrepancies concerning the nature of misligated intermediates under different conditions. At neutral pH, guanidine hydrochloride denaturation has implicated histidine residues in misligation events, but at high pH and under urea denaturation, only lysine residues are thought to be involved in misligation events (Russell, 2002).

The confusion involving the structure and nature of partially folded intermediates under equilibrium conditions lies in the wide variety of techniques used and the lack of a technique that can offer both high structural and high temporal resolution. Techniques that have high time resolution and often lack molecular resolution. Techniques possessing high structural resolution, such as NMR-based methods, have proven a very powerful, but studies using conventional NMR techniques are unable to get structural details under denaturing conditions when little native structure is present. Chemical exchange NMR methods are very useful in this respect, but these techniques are unable to distinguish local fluctuations from global unfolding events especially in flexible regions of the protein. Furthermore, the slow timescales of chemical exchange do not allow for the detection of rapidly converting intermediates or the faster events that occur in the folding process.

The incorporation of carbon-deuterium bonds offers residue-specific structural resolution coupled with high time resolution. Not only can one observe a specific amino acid residue within the protein, but the backbone or side chain of that residue can also be compared. In addition, the structural characterization of partially or fully denatured protein can proceed even with large amounts of denaturant present and there are no size limitations on the protein of interest. This infrared based technique also offers the ability to initiate and detect fast events in protein folding. In addition, the technique has the ability to resolve partially unfolded intermediates with high structural resolution. In order to characterize the equilibrium folding of oxidized cytochrome *c* more generally, six residues throughout the protein were deuterated and monitored at different concentrations of denaturant. The six deuterated residues are shown in Figure 5.1 and comprise three structural motifs in the C-terminal half of cytochrome *c*: the Met80 loop, the 60's helix and the C-terminal helix. To gain insight into the role of denaturant when investigating protein folding, the protein was unfolded with both guanidine hydrochloride and urea.

5.2 Materials and Methods

Sample Preparation. As described in Chapter 2, site-selective deuterated cytochrome *c* was prepared via semi-synthesis. Briefly, a cytochrome *c* 1-65 homoserine lactone peptide fragment containing the covalently bound heme was generated using CNBr cleavage at the Met65 residue of horse heart cytochrome *c*. This peptide was purified and then refolded at pH 7 in the presence of one equivalent
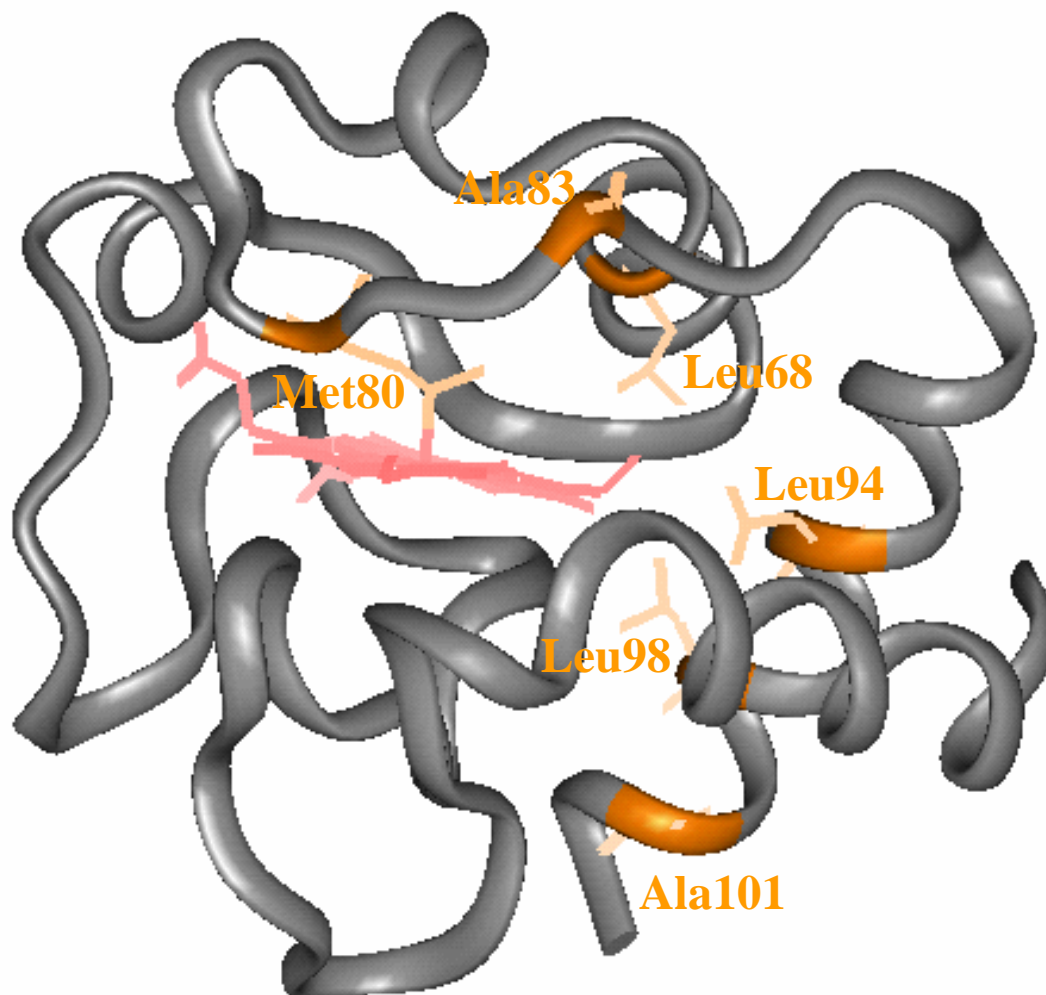
Figure 5.1. Crystal structure of horse heart cytochrome *c* showing the sites of deuteration. Six residues throughout the protein, including three different structural motifs were probed: the Met80 loop, the 60's helix and the C-terminal helix. In every case, the residue examined contained a deuterated terminal methyl group. Sites of deuteration include: two residues in the Met80 loop, d_3-Met80 and d_3-Ala83, one residue in the 60's helix, d_3-Leu68, and three residues in the C-terminal helix, d_3-Leu94, d_3-Leu98 and d_3-Ala101.

of chemically synthesized residues 66-104 of cytochrome *c* containing a specifically

deuterated leucine or methionine residue. Fragment association and aminolysis of the

homoserine lactone resulted in high yields of semi-synthetic cytochrome *c* with the

modification of a homoserine at position 65 instead of a methionine. This mutation

has been shown to have no effect the redox potential, function and common folding

markers of the protein. The final condensation product was purified from the

fragments and washed with 10 mM sodium acetate buffer, pH 6.2 for measurements in

GnHCl and 5 mM potassium phosphate buffer, pH 7 for urea measurements. The

protein solutions were then divided into individual samples and lyophilized. Samples

were stored at -20 ºC until ready for use. Guanidine hydrochloride solutions were

made by dissolving the appropriate amount of solid guanidine hydrochloride in 5 ml of

200 mM sodium acetate buffer, pH 6.2. Doubly distilled water was then added to

reach 10 ml and the pH of each solution checked individually and adjusted with 0.1 M

NaOH. Urea solutions were made by adding solid urea to 100 mM of 5 ml of

potassium phosphate buffer, pH 7 and 0.4 M KCl. After the solid urea was dissolved,

water was added to the 10 ml mark. Since urea increases the pH of the solution,

different amounts of monobasic and dibasic potassium phosphate buffer (along with

0.4 M KCl) were mixed to account for the change in pH caused by urea. The final pH

of all solutions was thoroughly checked and it was insured that the pH did not change

upon addition of the lyophilized protein. The concentration of GnHCl and urea

solutions were all individually checked by measuring their index of refraction

(Creighton, 1989). Guanidine hydrochloride and urea solutions were stored at -80 ºC

and only taken out of the freezer 2 hours prior to use. Lyophilized protein samples were dissolved in the appropriate guanidine hydrochloride or urea solution and allowed to stand approximately ten minutes prior to measurement.

FTIR Measurements. Samples were measured using a Bruker Equinox 55 FTIR spectrometer equipped with a liquid nitrogen cooled MCT detector and a KBr beamsplitter. The sample compartment was continuously flushed with dry nitrogen. Spectra were collected at 4 cm^{-1} resolution and constructed from 8192 scans. The Blackman-Harris 3-term apodization function was used, a 16 kHz low pass filter and zero filling of 16. The aperature was set to 2000 μm and the liquid transmission cell consisted of CaF$_2$ disks 32 mm in diameter with a Teflon spacer 50 μm thick. For every deuterated protein sample, a proteo sample was taken immediately afterwards and subtracted using a subtraction factor close to unity. After proteo subtraction, the baseline was corrected by fitting a window of at least 100 cm^{-1} to a polynomial and subtracting the polynomial to result in a relatively flat baseline with the peaks of interest.

Data Analysis. The data was fit to Gaussians representing chemically meaningful species as described in Chapter 2. The error values reported in Figures 5.2 - 5.27 are the result of the calculated standard deviation values of three independent data sets. For the three leucine residues, d_3-Leu68, d_3-Leu94, d_3-Leu98 it was determined that the folded spectra, acquired at 0 M GnHCl or urea, could be fit using two Gaussians. Likewise the unfolded spectra, acquired at 6 M GnHCl or 9 M urea as well as the free amino acid in the same buffer, was also well fit by two Gaussians. However, as

discussed in Chapter 2, the two Gaussians in the unfolded spectra are believed to represent the two asymmetric stretches and the two Gaussians in the folded protein most likely reflect overlapping asymmetric stretches of the two isotopomers. In order to simplify the fitting process, several parameters were fixed. For the folded spectra of the leucine residues, the frequencies and relative amplitudes of the two Gaussians were fixed. For the unfolded spectra, only the relative amplitudes of the two Gaussians were the only thing fixed. However, the frequencies of the unfolded spectra were restricted to a 4 cm^{-1} window, but could not be completely restricted due to salt effects that caused a slight red shift in the unfolded frequencies. The linewidths of both the folded and unfolded Gaussians were allowed to vary freely. It was ensured that all fits to the spectra at each denaturant concentration were the best possible fits to the data, yielding the lowest possible least squared error sum. Due to the low intensity of the symmetric stretch for leucine and alanine, only the asymmetric stretches are reported. Similarly, due to the low intensity of the asymmetric absorption of methionine, only the symmetric stretch is reported in this study. Generation of the plots of fraction folded/unfolded versus denaturant were carried out by first normalizing the intensities of all spectra to an isosbestic point. An isosbestic point is created when two absorbing species overlap and offers further proof of a two-state transition (see Figures 5.2 - 5.27). Subsequent normalization was carried out by dividing by the sum of all absorbing species and multiplying by the ratio of extinction coefficients.

The d_3-Met80 spectrum was well fit with a single Gaussian in both the folded and unfolded states. This is in accord with this peak being assigned as the symmetric stretch. The parameters of the folded and unfolded Gaussians were similarly determined from data at low and high denaturant concentrations. Fitting the data to a sum of folded and unfolded states yielded a partial unfolding at low denaturant concentrations. It was then revealed that the best fit to the data at low denaturant was in fact not a sum of the folded and unfolded states, but a sum of the folded and another state with slightly different spectral features than the unfolded state. The fits generating the final plots shown in Figures 5.3 and 5.5, arose from fixing the folded Gaussian frequency and restricting the unfolded and intermediate Gaussian frequencies to a 4 cm^{-1} window. The linewidths for all three Gaussians were allowed to vary freely. Also due to the low signal to noise ratio for the asymmetric stretches, only the symmetric stretch is reported.

The folded and unfolded state of d_3-Ala83 could be well fit with a single Gaussian. However, the quality of the fits at low GnHCl and urea concentrations were lower than those at high denaturant. Contour plots also reveal the convolution of additional signals at low denaturant concentrations (see Figure 5.8). Therefore, a second Gaussian was added to represent the intermediate species. Based on the contour plots, the frequency of this second Gaussian was fixed to a 6 cm^{-1} window. For the GnHCl data, the intermediate Gaussian was restricted to a lower frequency than that of the folded and unfolded states. The urea data were fit using an intermediate Gaussian at higher frequency than those of the folded and unfolded states. The linewidths of all

Gaussians were free to vary and due to the low signal to noise ratio for the symmetric stretch, only the asymmetric stretches are reported for this residue.

The d_3-Ala101 data was fit to a two-state model with one Gaussian representing the folded state and one Gaussian representing the unfolded state. Only the frequencies of the folded state were fixed while the linewidths for both Gaussians were free to vary. Similar to d_3-Ala83, the low signal to noise ratio for the symmetric stretch lead to only the asymmetric absorptions being reported for this residue.

5.3 Results

The Met80 Loop. The Met80 loop is particularly interesting from a folding perspective since it offers a unique view of protein-based ligation and misligation. As discussed in Chapter 4, the d_3-Met80 residue when unfolded in guanidine hydrochloride fit best to a three-state model with an intermediate at low GnHCl concentrations. After the accumulation of the intermediate species, which peaks at 1.2 M GnHCl, a second unfolding transition occurs with a midpoint of 2.3 ± 0.1 M GnHCl. Although the d_3-Met80 data is not two-state, a spectral window can be defined that does not include the intermediate at higher frequency. Within this spectral window, from 2120-2133 cm^{-1}, the data are roughly two-state (only containing the folded and unfolded species) and an isosbestic point can be identified, as shown in Figure 5.3. In urea, the overall results are very similar. The data are best fit to the three-state model where the third Gaussian, representing the intermediate species, has very similar spectral features to that observed in GnHCl. As shown in
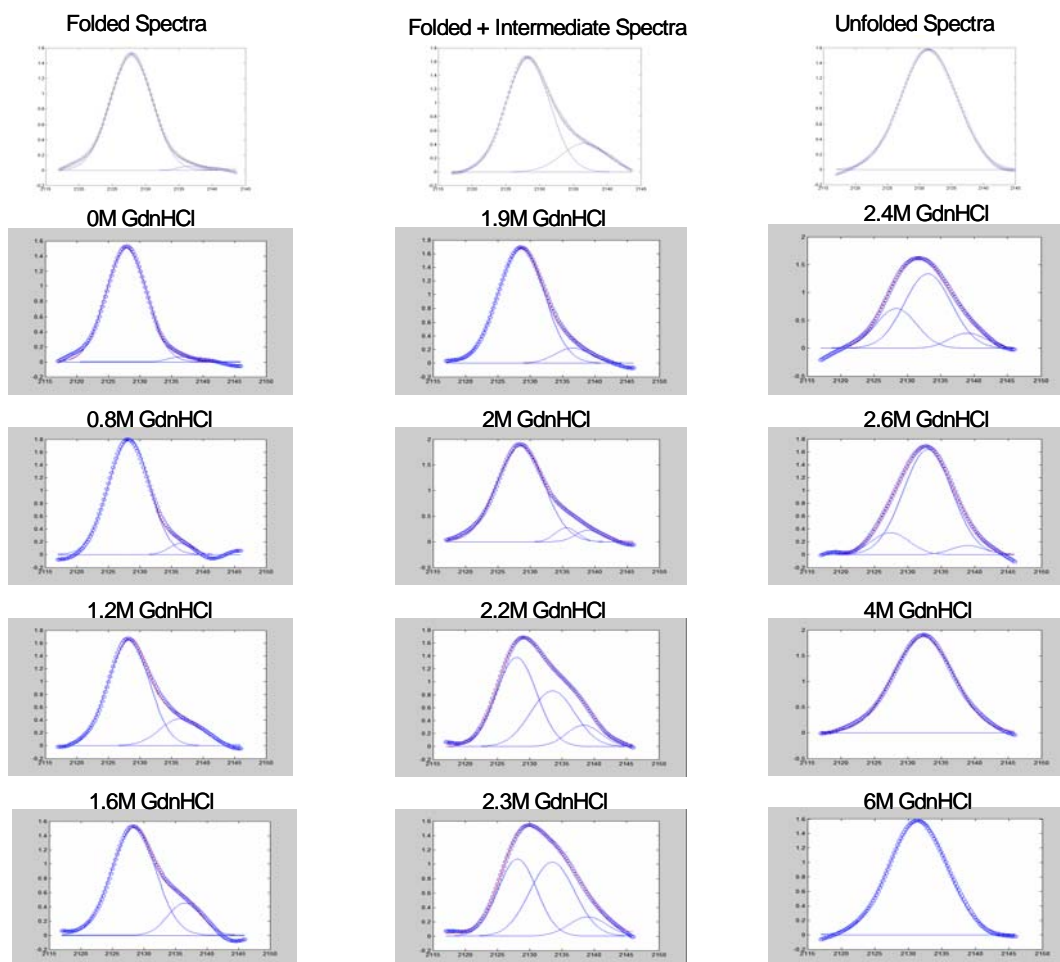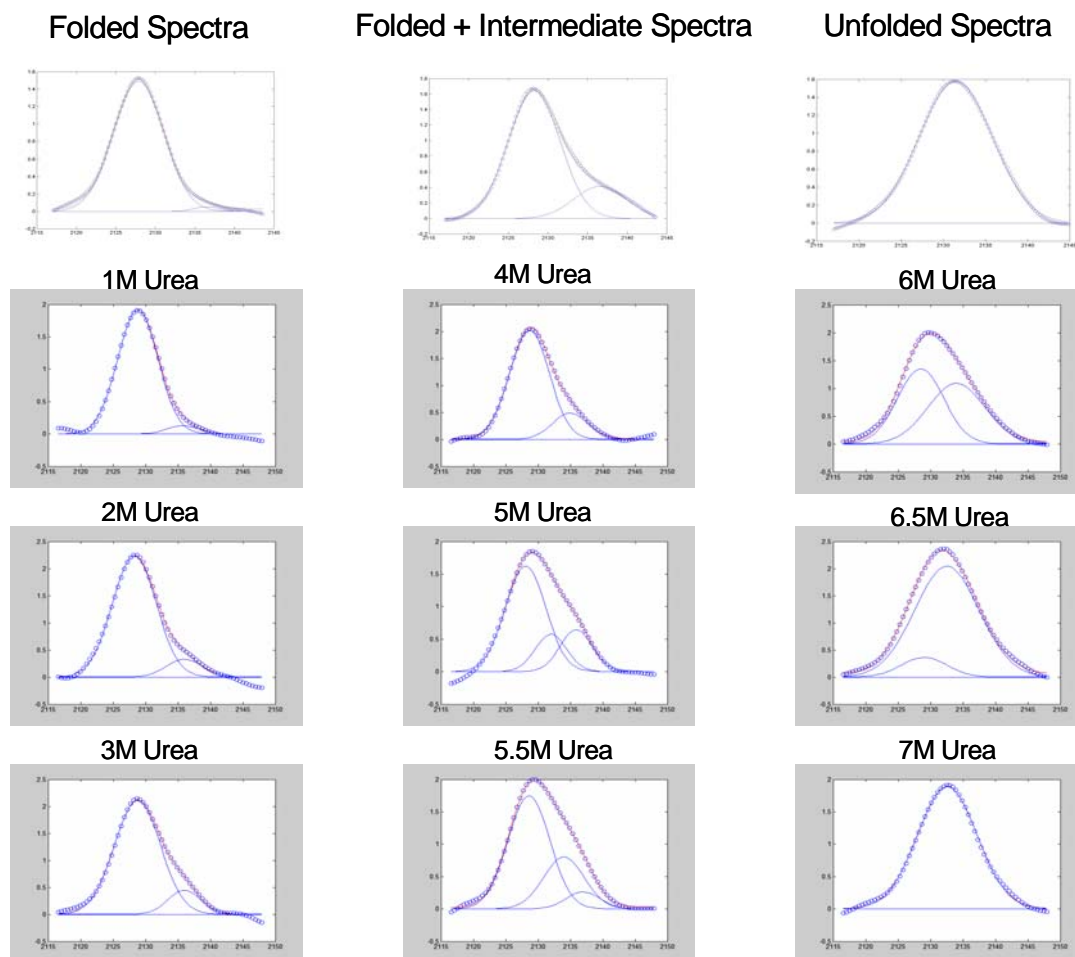
Figure 5.2. The d_3-Met80 data in GnHCl. The d_3-Met80 data in GnHCl was fit to three Gaussians representing the folded, intermediate and unfolded states. The intermediate state is seen to reach a maximum at low guanidine hydrochloride concentrations whereas the unfolded state reaches a maximum at high guanidine hydrochloride concentrations. The frequency of the folded Gaussian was fixed to values between 2127 and 2129 cm^{-1} as well as the frequency of the intermediate Gaussian to values between 2134 and 2138 cm^{-1}. The frequency values of the unfolded Gaussian and all linewidth values were not fixed.
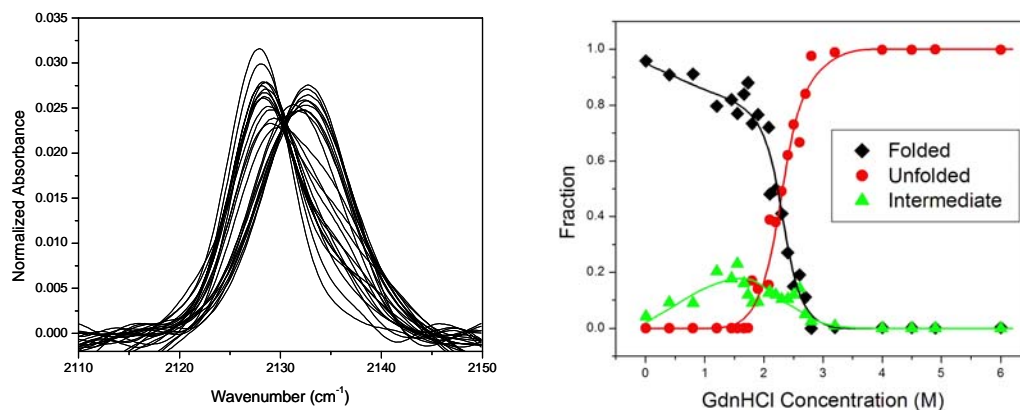
Figure 5.3.   Titration of d_3-Met80 with increasing guanidine hydrochloride concentration.   An isosbestic point at 2030.5 cm^{-1} is easily identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states.   A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride shows the presence of a partially unfolded intermediate at low guanidine hydrochloride concentrations and a global unfolding transition with a midpoint value of
2.3 ± 0.1M.

Figure 5.4. The d_3-Met80 data in Urea. The d_3-Met80 data in urea was fit to three Gaussians representing the folded, intermediate and unfolded states.   The intermediate state is shown to reach a maximum at low urea concentrations whereas the unfolded state reaches a maximum at high urea concentrations, similar to the data in GnHCl. The frequency of the folded Gaussian was fixed to values between 2127 and 2129 cm^{-1} as well as the frequency of the intermediate Gaussian to values between 2134 and 2138 cm^{-1}. The frequency values of the unfolded Gaussian and all linewidth values were not fixed.
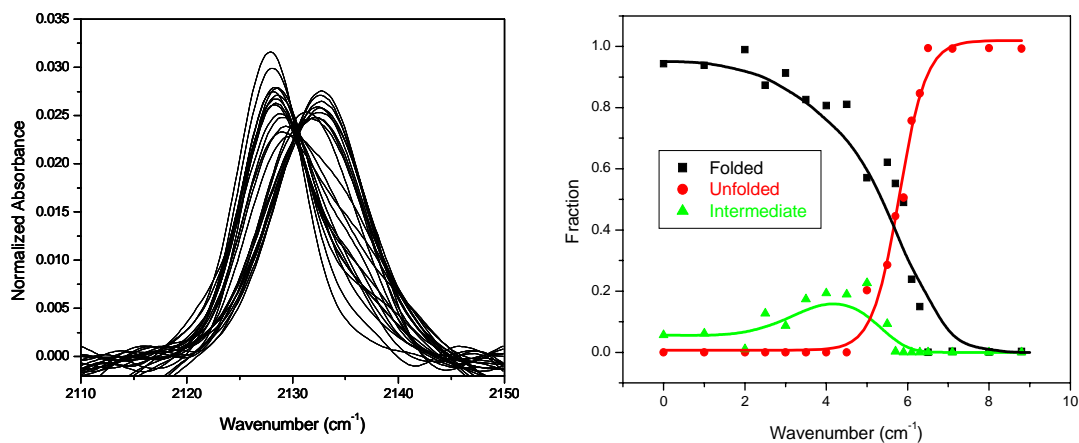
Figure 5.5.   Titration of d_3-Met80 with increasing urea concentration.   An isosbestic point at 2030.5 cm^{-1} is easily identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded, unfolded and intermediate Gaussians shows the presence of a partially unfolded intermediate at low urea concentrations which peaks at 4 M urea. An unfolding transition is also observed with a midpoint value of 5.9 ± 0.1 M.

Figure 5.4, the intermediate species in urea also has a higher frequency than the unfolded species and a similar linewidth to the intermediate observed in GnHCl. The titration curve shown in Figure 5.5 also shows a similar pattern to that observed in GnHCl, with the intermediate species having the highest amplitude at low urea concentrations as well as a larger unfolding transition at higher urea concentrations. The larger unfolding transition has a midpoint value of 5.9 ± 0.1 M urea.

The Met80 loop was also investigated using d_3-Ala83 as a probe. Unfortunately, as discussed in Chapter 3, the crystal structure reveals that the side chain of Ala83 is solvent exposed and thus not sensitive to the protein environment. Therefore, Ala83 shows no significant spectral changes when the protein is unfolded in either GnHCl and urea. Although the spectra look similar at very low and high denaturant concentrations, fitting the spectra to a single Gaussian at intermediate denaturant concentrations yielded inadequate fits. As shown in Figure 5.8, both the GnHCl and urea data indicate another spectral peak is present. This second absorption is shown to be an intermediate species in Figures 5.5 and 5.7 with similar profiles to that of Met80. The intermediate observed with d_3-Ala83 and d_3-Met80 appear to be involved in the same physical process. Thus the d_3-Ala83 data offers a direct observation that at least part of the Met80 loop undergoes a different intermediate conformation.

The 60's Helix. The d_3-Leu68 residue was used to probe of the 60's helix. The data in guanidine hydrochloride was best fit to a two-state model with two Gaussians representing the folded and unfolded states as shown in Figure 5.9. No evidence of additional spectral features could be detected at any denaturant concentration. The
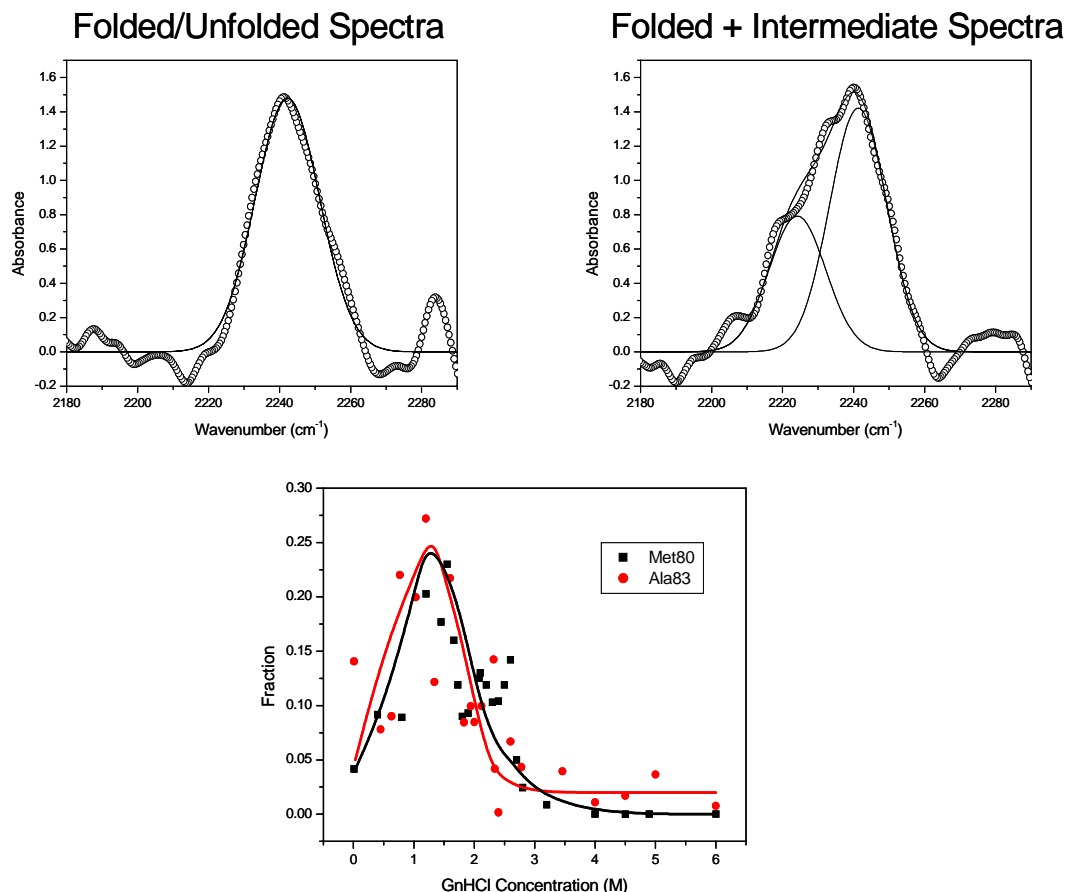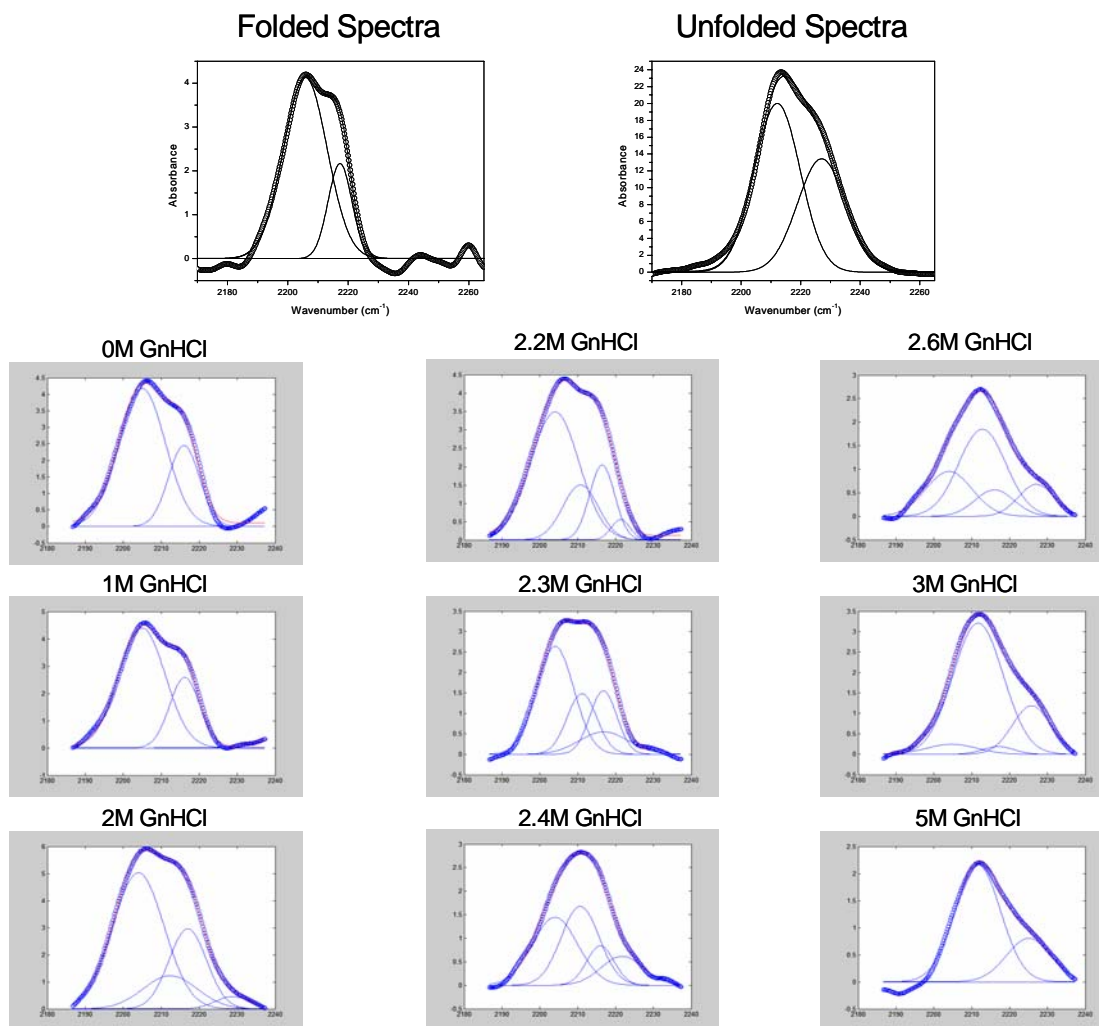
Figure 5.6. The d_3-Ala83 data in GnHCl. The folded and unfolded spectra (low and high concentrations of guanidine hydrochloride) for d_3-Ala83 showed identical spectral features and could be fit using one Gaussian. At intermediate guanidine hydrochloride concentrations a second peak appears, so a second Gaussian at lower frequency was added to represent the intermediate. Since the folded and unfolded spectra are not spectrally distinct, no isosbestic point was observed and the intensities shown are not normalized. However, the intensities of the intermediate for d_3-Ala83 and d_3-Met80 were similar in intensity and normalized to each other for plotting purposes. The frequency of the intermediate Gaussian was fixed to values between 2224 and 2229 cm^{-1}. The frequency values of the folded/unfolded Gaussian and all linewidth values were not fixed. A plot of intensities of the intermediate Gaussian for d_3-Ala83 reveal a similar pattern to that observed at d_3-Met80.
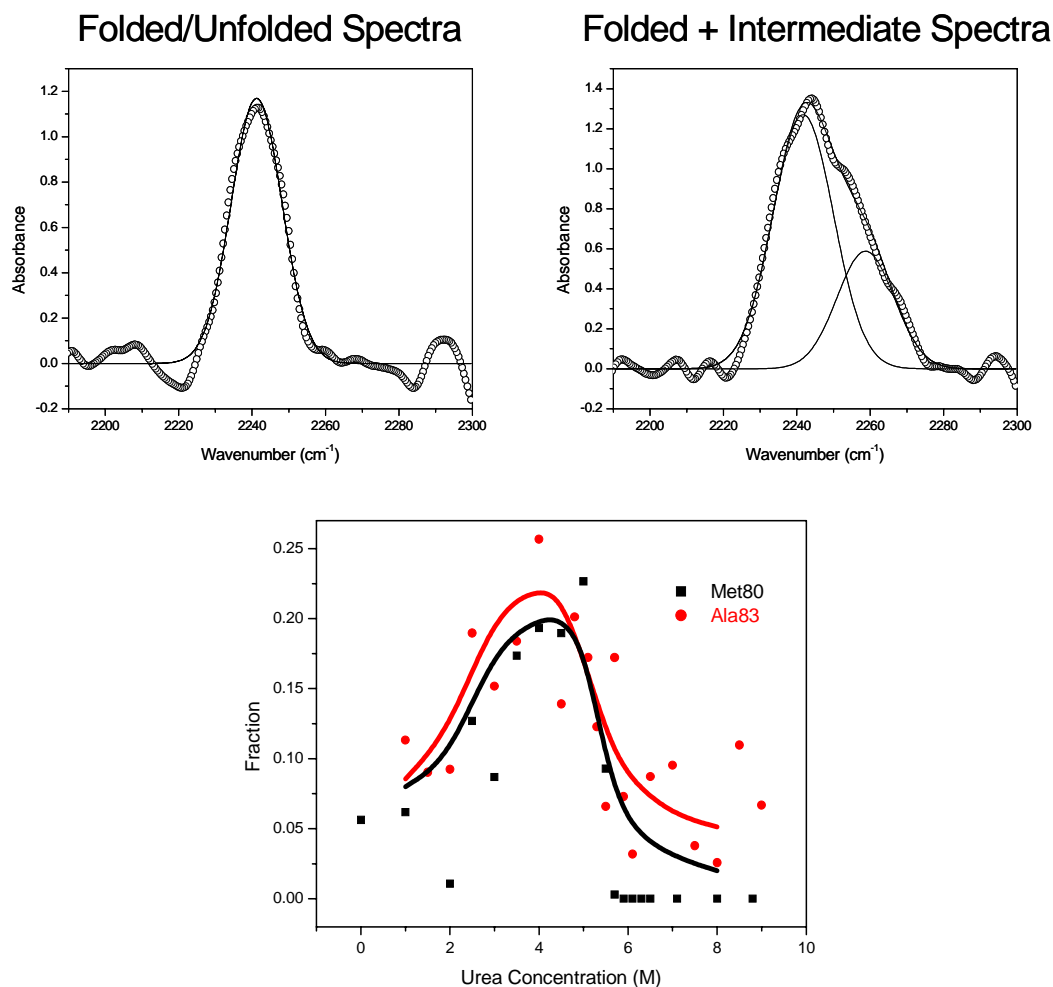
Figure 5.7. The d_3-Ala83 data in urea. The folded and unfolded spectra (low and high concentrations of urea) for d_3-Ala83 showed identical spectral features and could be fit using only one Gaussian. A second peaks appears at intermediate urea concentrations and thus a second Gaussian at higher frequency was added to represent the intermediate. Since the folded and unfolded spectra are not spectrally distinct, no isosbestic point was observed and the intensities shown are not normalized. However, the intensities of the intermediate for d_3-Ala83 and d_3-Met80 were similar in intensity and normalized to each other for plotting purposes. The frequency of the d$_3$-Ala83 intermediate Gaussian was fixed to values between 2250 and 2263 cm^{-1}. The frequency values of the folded/unfolded Gaussian and all linewidth values were not fixed. A plot of intensities of the intermediate Gaussian for d_3-Ala83 reveal a similar pattern to that observed at d_3-Met80.
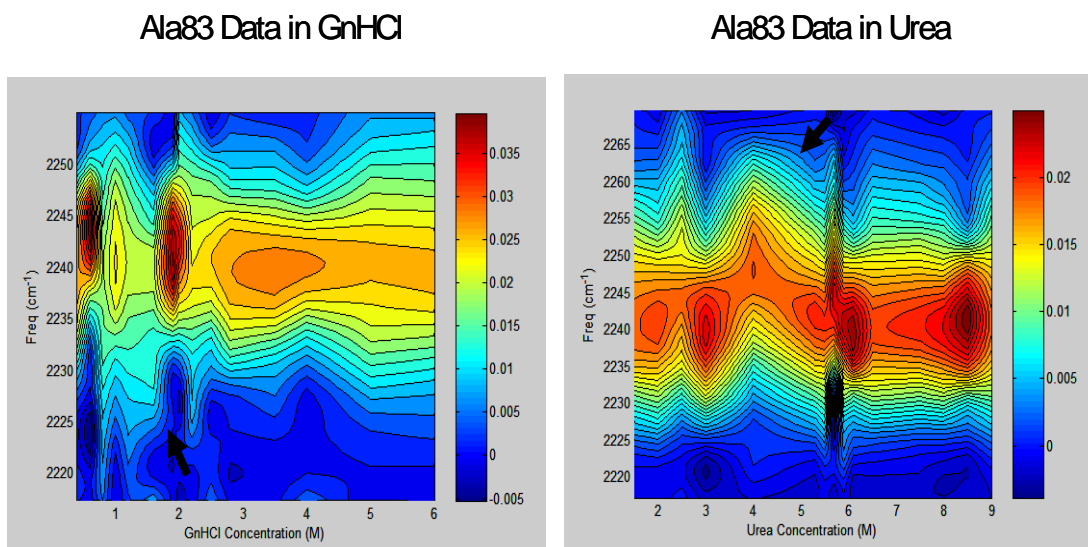
Figure 5.8. Contour plots comparing d_3-Ala83 in GnHCl and urea. Overall the spectra show no significant changes when titrated with denaturant except for the appearance of an intermediate peak at low guanidine hydrochloride and urea concentration. Particularly noteworthy is the difference in intermediate spectral features. The guanidine hydrochloride data consistently shows an intermediate at lower frequency, whereas the urea data indicates an intermediate at higher frequency.
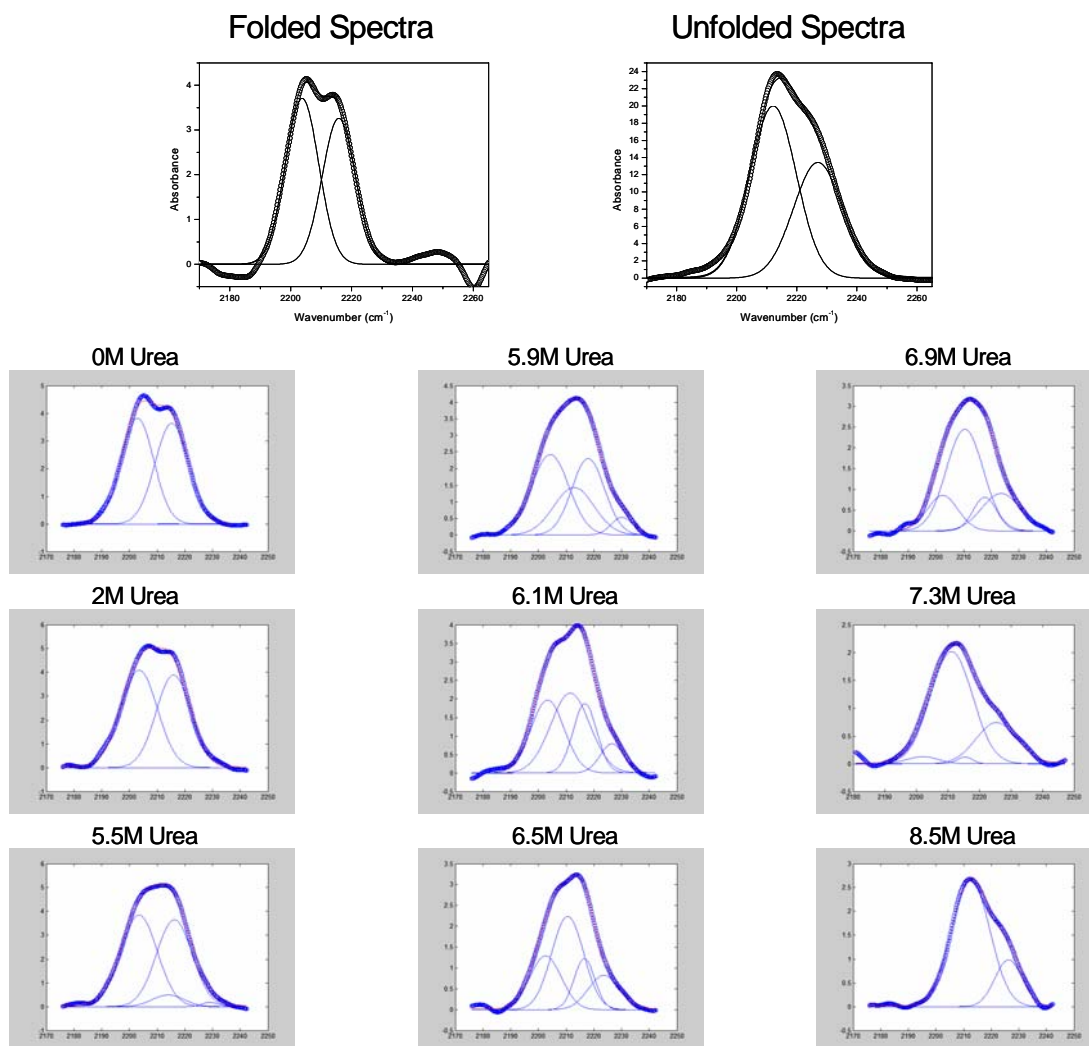
Figure 5.9. The d_3-Leu68 data in GnHCl. The d_3-Leu68 data in GnHCl was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state. The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians. The frequency values of the unfolded Gaussians and all linewidth values were not fixed.
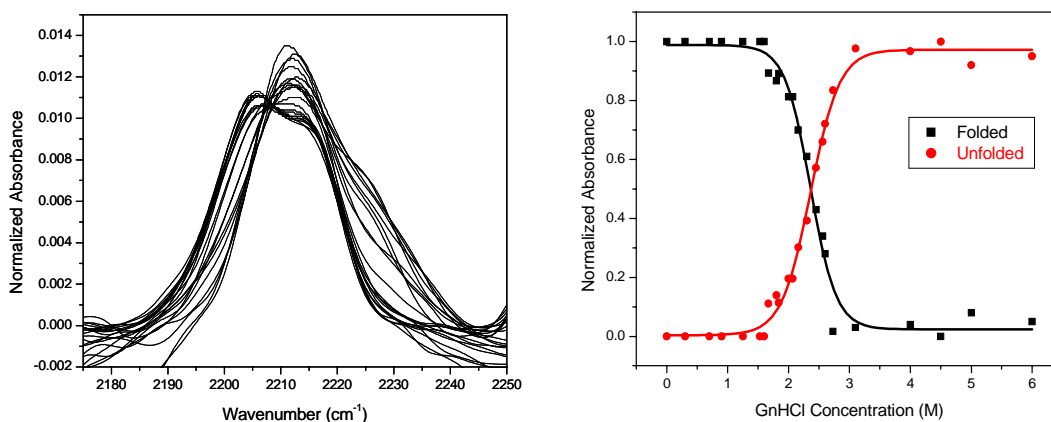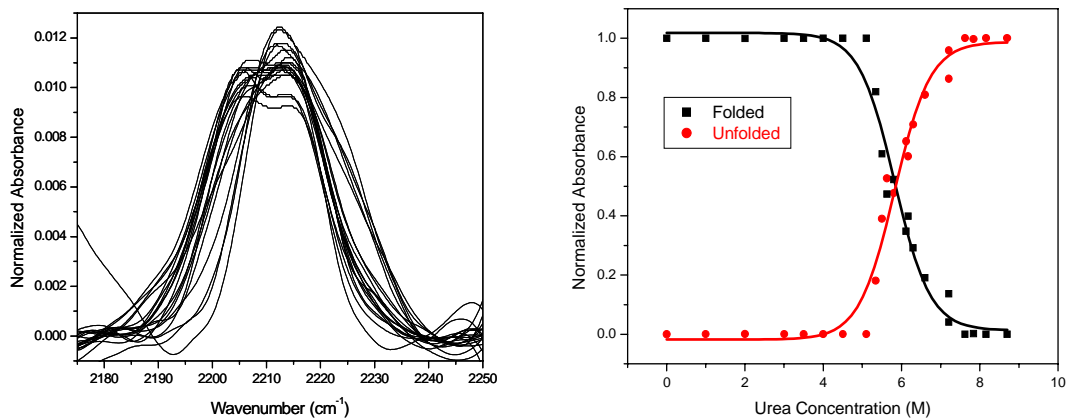
Figure 5.10. Titration of d_3-Leu68 with increasing guanidine hydrochloride concentration. An isosbestic point at 2208 cm^{-1} is readily identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride shows a single two-state transition with a midpoint value of 2.3 ± 0.1 M.

Figure 5.11. The d_3-Leu68 data in Urea. The d_3-Leu68 data in urea was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state. The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians. The frequency values of the unfolded Gaussians and all linewidth values were not fixed.
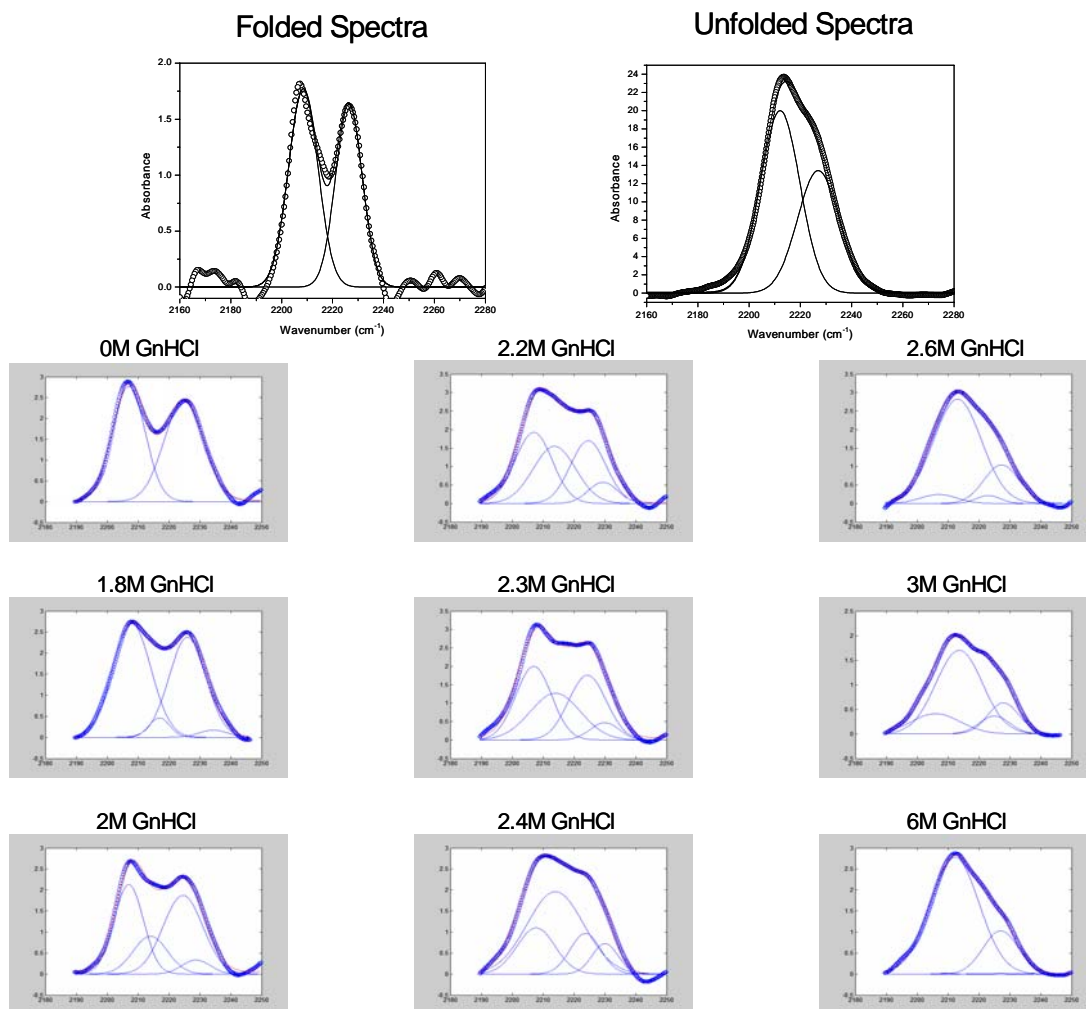
Figure 5.12. Titration of d_3-Leu68 with urea concentration. An isosbestic point at 2208 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus urea concentration shows a single two-state transition with a midpoint value of 5.9 ± 0.1 M.

midpoint for the sigmoidal transition is the same as that observed for the second transition of the Met80 residue, 2.3 ± 0.1 M guanidine hydrochloride. The data in urea is also best fit with a two-state model and the observed midpoint is also similar to that of Met80, with a value of 5.9 ± 0.1 M urea.

The C-terminal Helix. The first residue investigated in the C-terminal helix, d_3-Leu94, lies in the center of the helix and contains the most tertiary interactions with the N-terminal helix. Similar to d_3-Leu68, the folded and unfolded spectra could be fit to two Gaussians and a two-state transition curve generated from the titration. The data in guanidine hydrochloride produced a midpoint value of 2.3 ± 0.1 M GnHCl, similar to the midpoint values obtained for both Leu68 and Met80. In urea, the midpoint value of 6.5 ± 0.1 M is much later than that of Leu68 and Met80, indicating an increased stability of the C-terminal helix.

Investigation of the C-terminal helix continued with d_3-Leu98, which is only three residues from Leu94 towards the C-terminus of the protein. This residue also showed two-state behavior and was fit similar to that of d_3-Leu94 (see Figures 5.16-5.20). The data in guanidine hydrochloride again display a midpoint value of 2.3 ± 0.1 M, similar to that observed for all the other residues examined thus far. The midpoint value of 6.3 ± 0.1 M in urea is slightly lower than that of d3-Leu94.

The last residue examined, d_3-Ala101, is three residues from Leu98 as well as three residues from the C-terminus of the protein. Interestingly, the data in both guanidine hydrochloride and urea indicate C-terminal fraying; where the spectra show a significant amount of unfolded protein even at 0 M denaturant (see Figures 5.21-5.24).
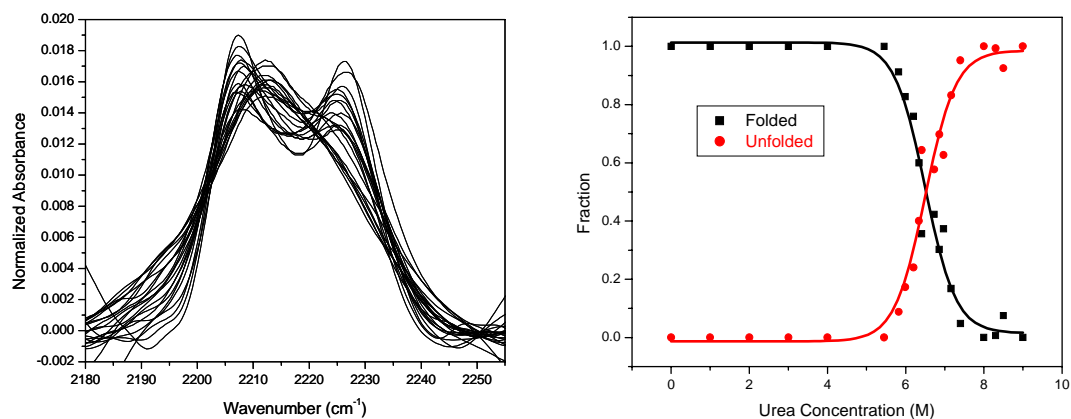
Figure 5.13. The d_3-Leu94 data in GnHCl. The d_3-Leu94 data in guanidine hydrochloride was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state. The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians. The frequency values of the unfolded Gaussians and all linewidth values were not fixed.
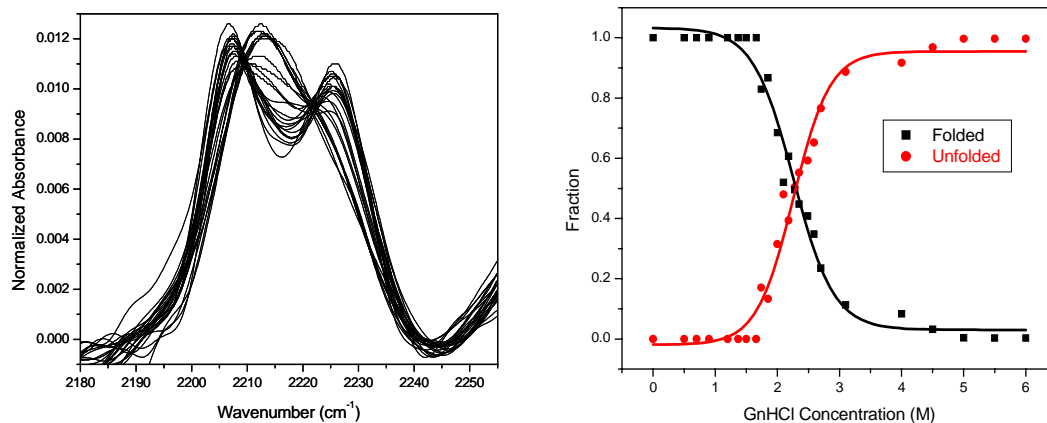
Figure 5.14. Titration of d_3-Leu94 with GnHCl concentration. An isosbestic point at 2221.8 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride concentration shows a single two-state transition with a midpoint value of 2.3 ± 0.1 M.
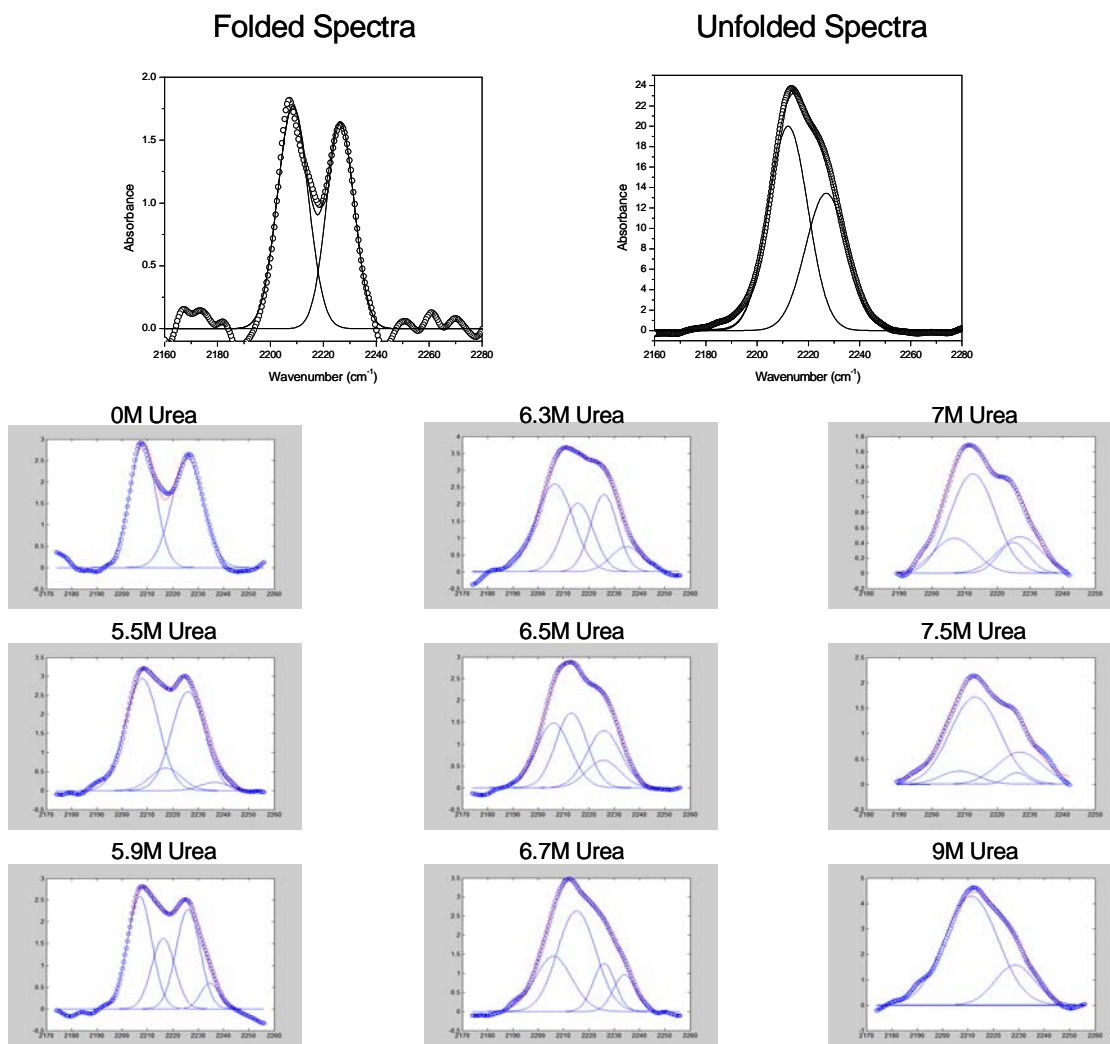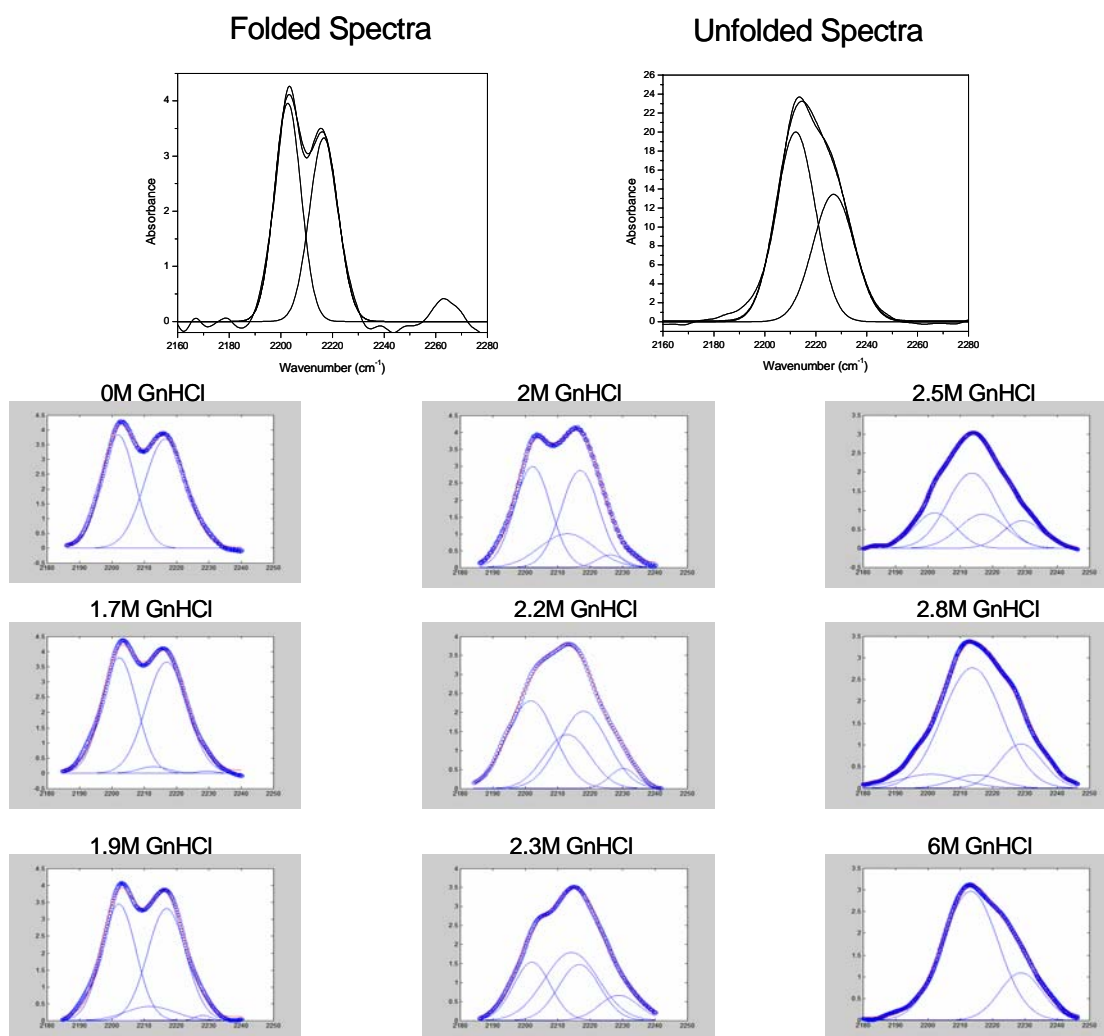
Figure 5.15. The d_3-Leu94 data in Urea. The d_3-Leu94 data in urea was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state. The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians. The frequency values of the unfolded Gaussians and all linewidth values were not fixed.
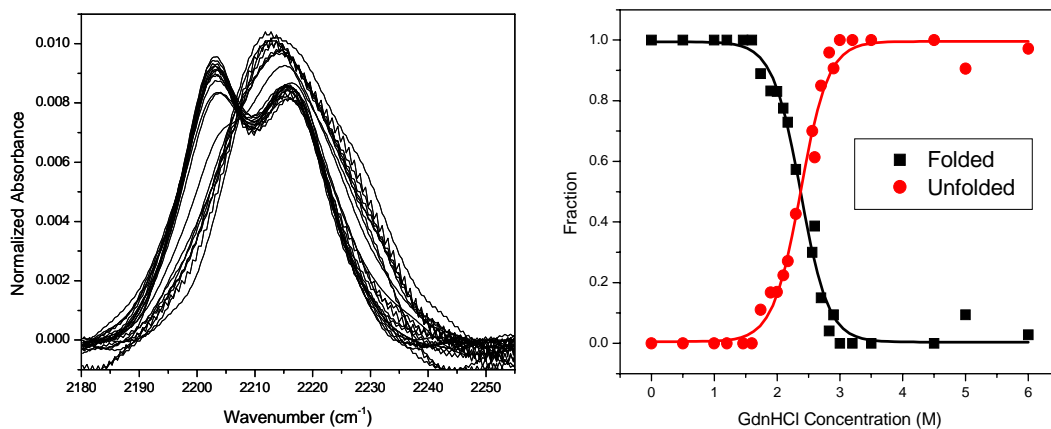
Figure 5.16. Titration of d_3-Leu94 with urea concentration. An isosbestic point at 2222.5 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus urea concentration shows a single two-state transition with a midpoint value of 6.5 ± 0.1 M.
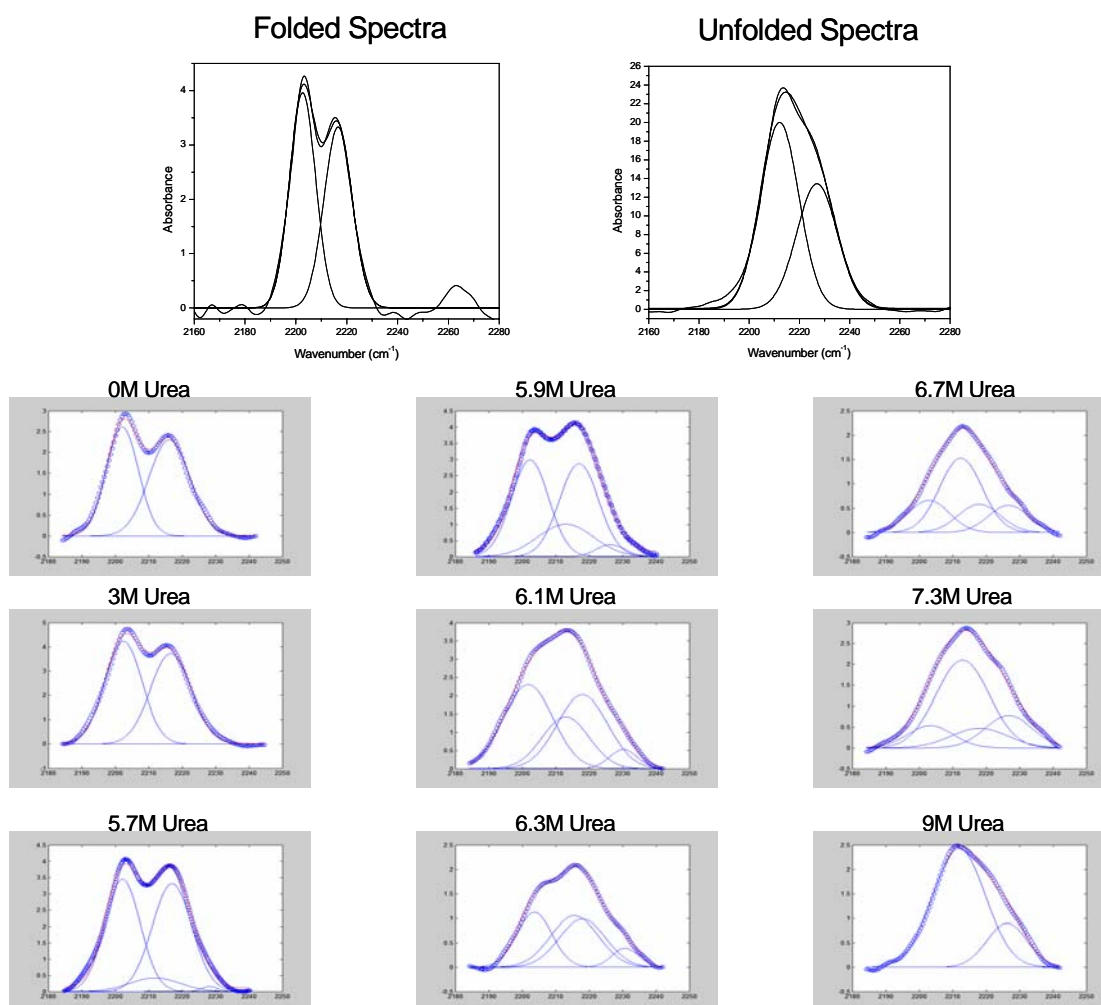
Figure 5.17.   The d_3-Leu98 data in GnHCl.   The d_3-Leu98 data in guanidine hydrochloride was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state.   The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians.   The frequency values of the unfolded Gaussians and all linewidth values were not fixed.

Figure 5.18. Titration of d_3-Leu98 with guanidine hydrochloride concentration. An isosbestic point at 2208 cm^{-1} is readily identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus urea concentration shows a single two-state transition with a midpoint value of 2.3 ± 0.1 M.

**Figure 5.19**. The d_3-Leu98 data in Urea. The d_3-Leu98 data in urea was fit to a two-state model with two Gaussians representing the folded state and two Gaussians representing the unfolded state. The relative amplitudes of both the folded and unfolded states were fixed as well as the frequencies of the folded Gaussians. The frequency values of the unfolded Gaussians and all linewidth values were not fixed.
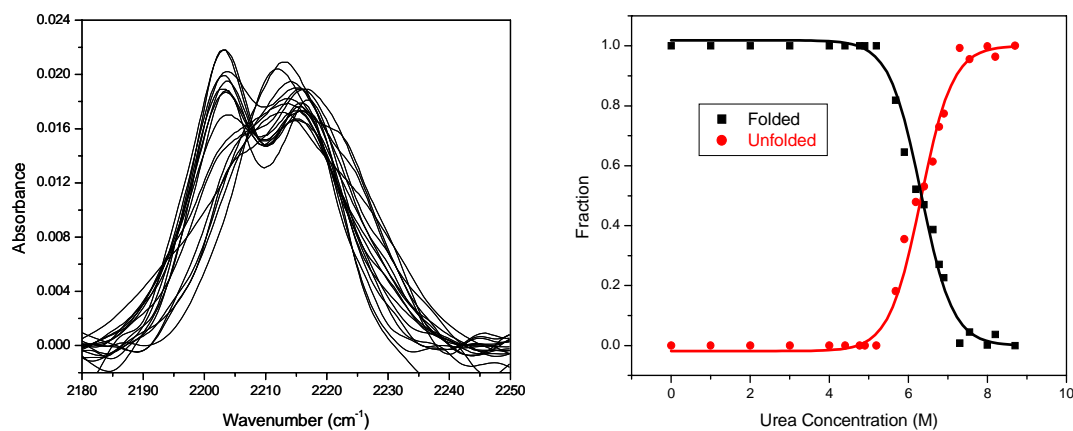
Figure 5.20. Titration of d_3-Leu98 with urea concentration. An isosbestic point at 2208 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. A final plot of normalized intensities for the folded and unfolded states versus urea concentration shows a single two-state transition with a midpoint value of 6.3 ± 0.1 M.
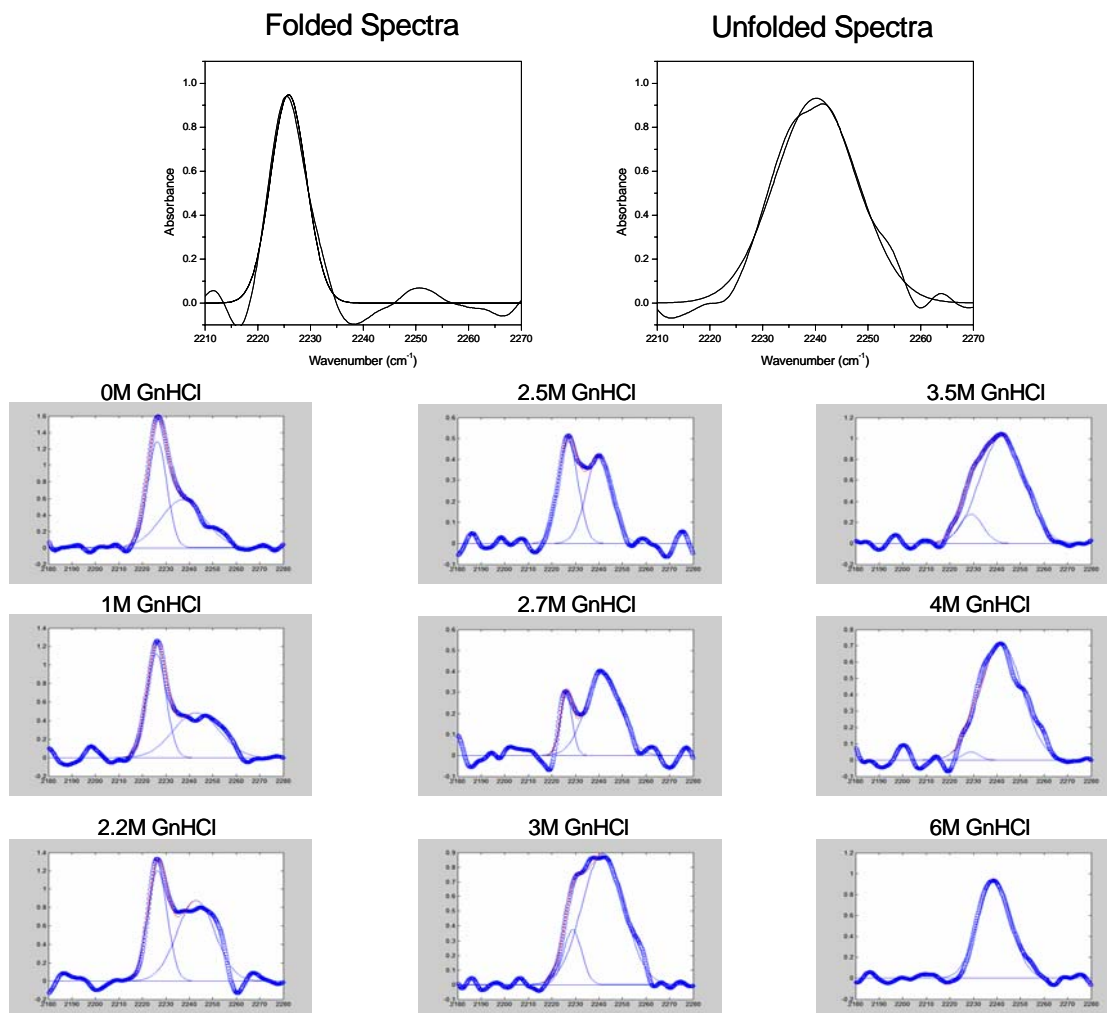
Figure 5.21. The *d₃*-Ala101 data in GnHCl. The *d₃*-Ala101 data in guanidine hydrochlordide was fit to a two-state model with one Gaussian representing the folded state and one Gaussian representing the unfolded state. Due to the lower signal to noise for this residue, the frequencies of the folded state were fixed to values between 2226 and 2229 cm⁻¹. Likewise, the frequency values of the unfolded state were also fixed to values between 2237 and 2243 cm⁻¹. All linewidth values were not fixed.
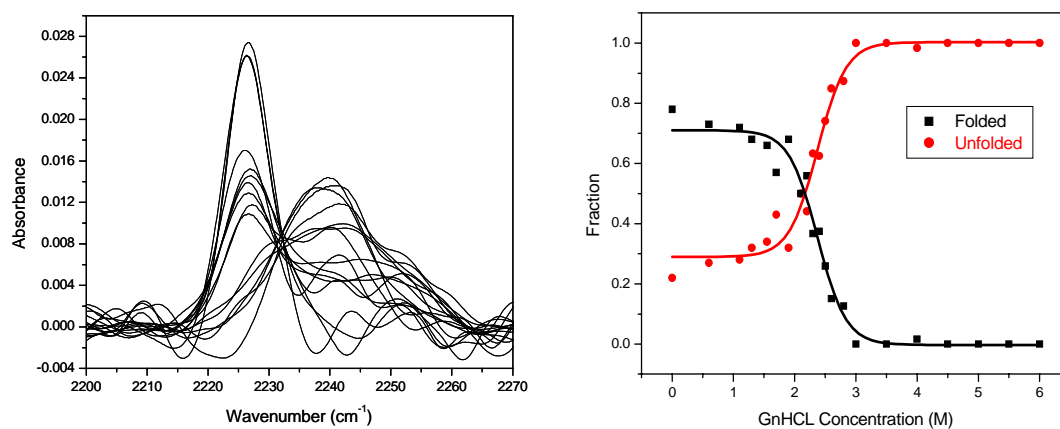
Figure 5.22. Titration of d_3-Ala101 with guanidine hydrochloride concentration. An isosbestic point at 2232 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. Another normalization step is included to account for the difference in extinction coefficient of the folded and unfolded states (the folded state is approximately 2-times more intense). A final plot of normalized intensities for the folded and unfolded states versus guanidine hydrochloride concentration shows a single two-state transition with a midpoint value of 2.3 ± 0.1 M. Note that the asymptotic value at low guanidine hydrochloride concentrations is not zero, but close to 40% unfolded, indicating C-terminal fraying.
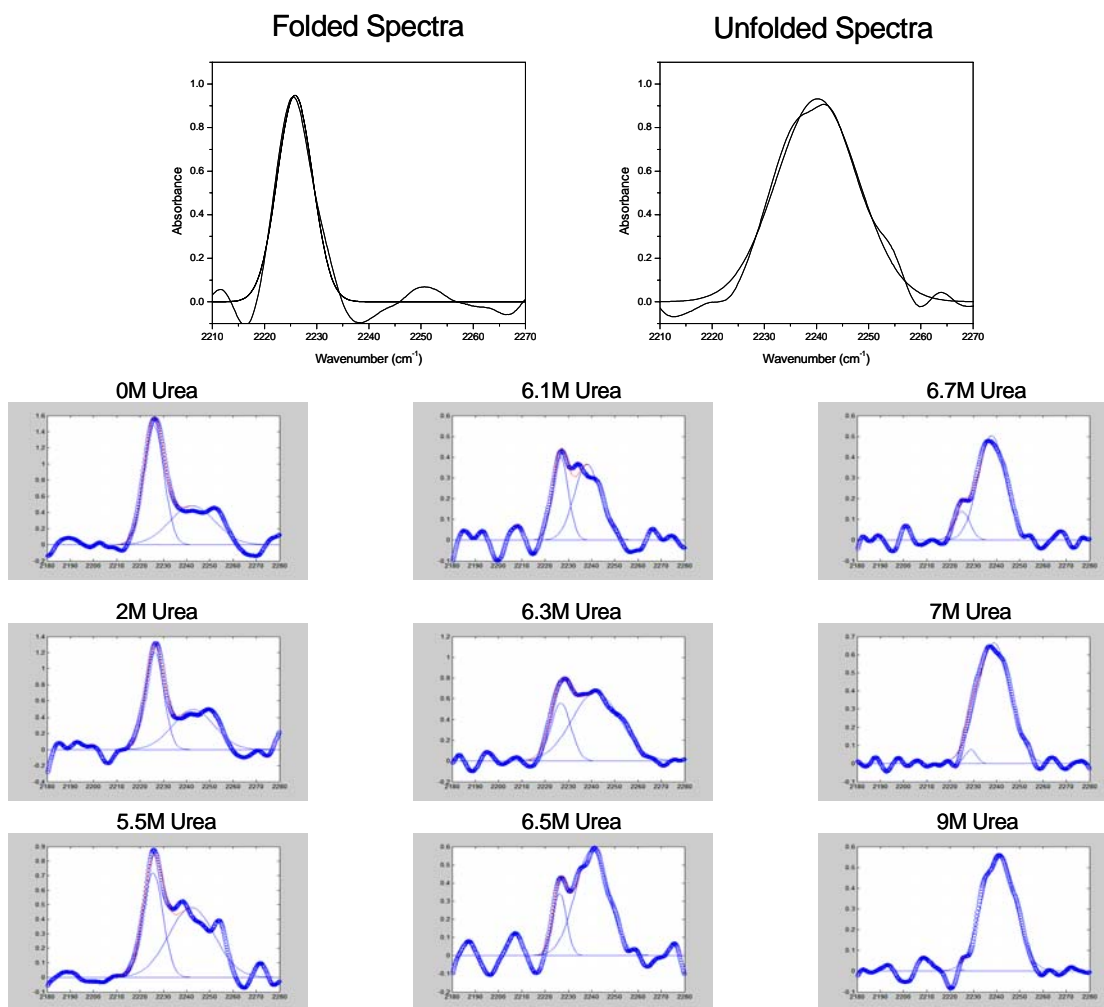
Figure 5.23. The d_3-Ala101 data in Urea. The d_3-Ala101 data in urea was fit to a two-state model with one Gaussian representing the folded state and one Gaussian representing the unfolded state. Due to the lower signal to noise for this residue, the frequencies of the folded state were fixed to values between 2226 and 2229 cm^{-1}. Likewise, the frequency values of the unfolded state were also fixed to values between 2237 and 2243 cm^{-1}. All linewidth values were not fixed.
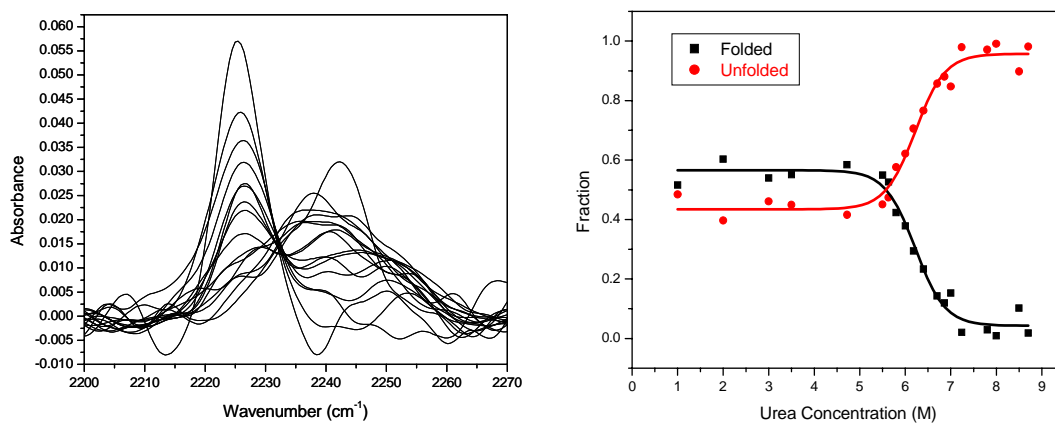
Figure 5.24. Titration of d_3-Ala101 with urea concentration. An isosbestic point at 2232 cm^{-1} is identified when plotting all the spectra normalized to the area underneath the peaks of interest. This isosbestic point is then used to further normalize the intensities of the folded and unfolded states. Another normalization step is included to account for the difference in extinction coefficient of the folded and unfolded states (the folded state is approximately 2-times more intense). A final plot of normalized intensities for the folded and unfolded states versus urea concentration shows a single two-state transition with a midpoint value of 6.05 ± 0.1 M. Note that the asymptotic value at low guanidine hydrochloride concentrations is not zero, but close to 40% unfolded, indicating C-terminal fraying.

Although the unfolding transitions are less defined than those of the other residues in the C-terminal helix due to the substantial amount of unfolded protein at low denaturant concentrations, the guanidine hydrochloride data also exhibit the common midpoint value of 2.3 ± 0.1 M. The urea data, however, yield a midpoint value of 6.1 ± 0.1 M, which is slightly lower than that of both Leu94 and Leu98. The residues in the C-terminal helix, when plotted together as shown in Figure 5.27, indicate that the unfolding in urea occurs from the C-terminal end.

5.4 Discussion and Conclusions

The data overall indicate a different unfolding mechanism in guanidine hydrochloride and urea, as shown in Figure 5.25. With the exception of an intermediate observed at low guanidine hydrochloride concentration involving only the Met80 loop, the unfolding in guanidine hydrochloride is cooperative, with every residue investigated showing the same global transition. In contrast, the urea data show a low level of cooperativity, with residues in the same structural motif, even residues only three apart, yielding small differences in midpoint values. The difference in cooperativity when comparing the two denaturants is also supported by the m values shown in Table 5.1. For the urea data, the m values range from 0.8 - 1.3 kcal/mol-M, but the guanidine hydrochloride data have higher m values, ranging from 1.6 - 2.4 kcal/mol-M. The higher m values obtained for the guanidine hydrochloride data indicate a higher degree of solvent exposure for each transition and thus, a higher degree of cooperativity (Pace, 1975). Denaturant-dependent differences in folding have been
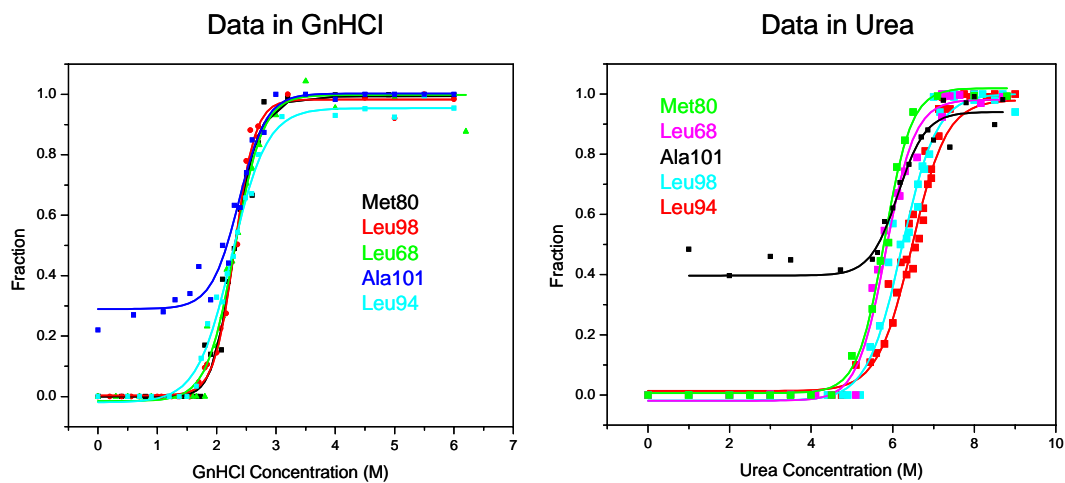
Figure 5.25. Comparison of the unfolding of cytochrome *c* in the two conditions currently investigated, urea and guanidine hydrochloride denaturation. The data in guanidine hydrochloride indicates a cooperative mechanism, with all residues having a global transition at 2.3 M GnHCl. The only exception is the residues in the Met80 loop, *d₃*-Met80 and *d₃*-Ala83, which indicate a common intermediate at low guanidine hydrochloride concentrations (see Figures 5.2-5.5). In contrast, the urea data indicates a sequential mechanism, somewhat similar to the Roder/Englander model.

reported for small peptides and globular proteins where guanidine hydrochloride is shown to be a more effective denaturant, most likely due to its ionic nature (Smith, 1996). Likewise, cytochrome *c* folding studies have also shown that solvent conditions can influence the population of partially folded intermediates and therefore the overall folding pathway. However, the residue-specific study described herein shows a remarkable difference in unfolding mechanism attributed solely to the difference in denaturant used.

The Met80 Loop. The Met80 residue, as discussed in Chapter 4, shows an intermediate at low guanidine hydrochloride concentration and a cooperative unfolding transition at 2.3 M guanidine hydrochloride. Based on previous reports, the intermediate observed at the Met80 residue is most likely a misligated conformer (Russell, 2002). The Met80 loop, being fairly flexible and solvent exposed, has been problematic to obtain structural information under mildly denaturing conditions with conventional techniques. The results presented here for the Ala83 residue is the first attempt to characterize the Met80 loop under conditions where the misligated intermediate is shown to be present. Interestingly, the data clearly shows that at least part of the Met80 loop assumes a different conformation when misligation occurs. Further examination of the spectral features of the misligated conformer reveal a consistent change in frequencies. The Met80 residue shifts to higher frequency upon misligation, whereas the Ala83 residue shifts to lower frequency. Based on the free amino acid solvent data (see Chapter 2) and the x-ray crystal structure, it is presumed that the Met80 residue, when ligated to the heme is in a nonpolar environment and

misligation creates a more polar environment at this residue, most likely moving it into solution (Bushnell, 1990). Similarly, Ala83 is solvent exposed in the folded conformation and appears to be inverted into the lower dielectric protein environment upon misligation. The data indicates that perhaps the Met80 loop undergoes a flipping motion to accommodate misligation.

The unfolding of Met80 in urea also shows an intermediate at low concentrations of denaturant and this intermediate has similar spectral features to that observed in guanidine hydrochloride. From the perspective of the Met80 residue, misligation in urea appears to be similar to that in guanidine hydrochloride; however, the Ala83 residue reveals some interesting differences. Instead of the misligated conformer having a lower frequency than that of the folded state, the frequency is higher (see Figure 5.8). This suggests that although Met80 is still displaced at low urea concentrations, the Met80 loop exhibits a different conformation than that in guanidine hydrochloride. A recent NMR study shows that low concentrations of guanidine hydrochloride results in nearby lysine residues or histidine residues displacing Met80 from the iron center. In urea, only one nearby lysine residue, Lys79, was implicated in misligation (Russell, 2002). Since the Met80 loop would not have to accommodate as many misligated conformers in urea, perhaps only minor structural reorganization is required and the higher frequency of Ala83 is possibly the result of geometric strain (see Chapter 2).

The 60's helix. The unfolding of Leu68 in guanidine hydrochloride provides further evidence for a cooperative folding transition. This is contrary to recent amide-
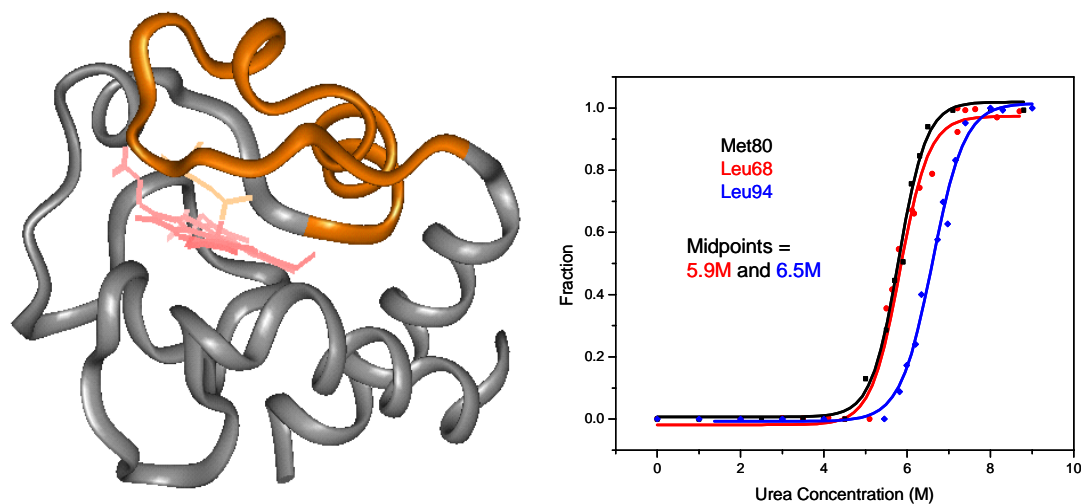
Figure 5.26. Plot of d_3-Met80, d_3-Leu68 and d_3-Leu94 unfolding in urea (see Figures 5.2-5.16). The midpoints of Met80 and Leu68 are similar and distinctly different than Leu94 in the C-terminal helix. This suggests the possibility that Met80 and Leu68 belong to a cooperative folding unit that is less stable than the C-terminal helix.

Table 1: Delta Gº and m values of residues throughout cytochrome *c*.

Amino Acids	Data in GnHCl		Data in Urea	
	Delta Gº (kcal/mol)	m (kcal/mol-M)	Delta Gº (kcal/mol)	m (kcal/mol-M)
d_3-Leu68	5.1 ± 0.7	-2.1 ± 0.3	4.8 ± 1.0	-0.8 ± 0.2
d_3-Leu94	3.8 ± 0.5	-1.6 ± 0.2	9.0 ± 0.6	-1.4 ± 0.1
d_3-Leu98	5.5 ± 1.3	-2.4 ± 0.6	7.9 ± 0.7	-1.3 ± 0.1
d_3-Ala101	4.0 ± 1.4	-1.9 ± 0.4	6.2 ± 3.2	-1.1 ± 0.5

Calculated ΔG^o and m values for all deuterated residues investigated throughout cytochrome *c* exhibiting a two-state transition.   Both ΔG^o and m values were calculated using the linear extrapolation method based on $\Delta G^o = -RTlnK$ and $\Delta G = \Delta G^o + mx$. The equilibrium constants were calculated from the ratio of the normalized amplitudes of the folded and unfolded states. The reported error values result in the evaluation of three independent data sets. Only the residues shown to be two-state are included, since a more detailed model is required for residues that contain a folding intermediate. Note that the m values for the data in GnHCl are higher than those in urea, indicating a more cooperative folding mechanism.   Also note the marked difference in ΔG^o values for the urea data. Both Leu68 and Ala101 unfold at much lower energy compared to Leu94 and Leu98.

exchange NMR measurements carried out under similar conditions where the 60's helix was shown to be approximately 3 kcal/mol less stable than the C-terminal helix (Xu, 1998). Interestingly, the urea data shows a similar trend to that observed in the amide-exchange NMR measurements. A comparison of d_3-Leu68 and d_3-Leu98 shows the same difference of 3 kcal/mol in stability, with the 60's helix being less stable than the C-terminal helix. However, a comparison of the two models of cytochrome c folding yields an important difference; the major unfolding transition of d_3-Met80 appears to be the same as that of d_3-Leu68 (see Figure 5.26). Unfortunately, without a detailed model of how the three states for the Met80 residue are interconnected, it was not possible to accurately calculate ΔG^o and m values for this residue. Nevertheless, the midpoint values suggest that the 60's helix, the Met80 loop and possibly the region in between unfolds as a cooperative unit. Since the amide-exchange NMR studies were not able to distinguish an intermediate at low denaturant concentration and a major unfolding transition, the lower stability reported for the Met80 loop may be due to the averaging of these two species. Although the midpoint values for d_3-Leu68 and d_3-Met80 suggests cooperativity in this region of the protein, there is not a significant difference in m values for d_3-Leu68 when compared to the residues in the C-terminal helix. Based on structural NMR studies, the region close to the Met80 loop is believed to be significantly more flexible and solvent exposed than residues in the C-terminal helix (Qi, 1996). This significant difference in solvent exposure in the folded form for these two regions of the protein may account for the lack of difference in m values upon unfolding.
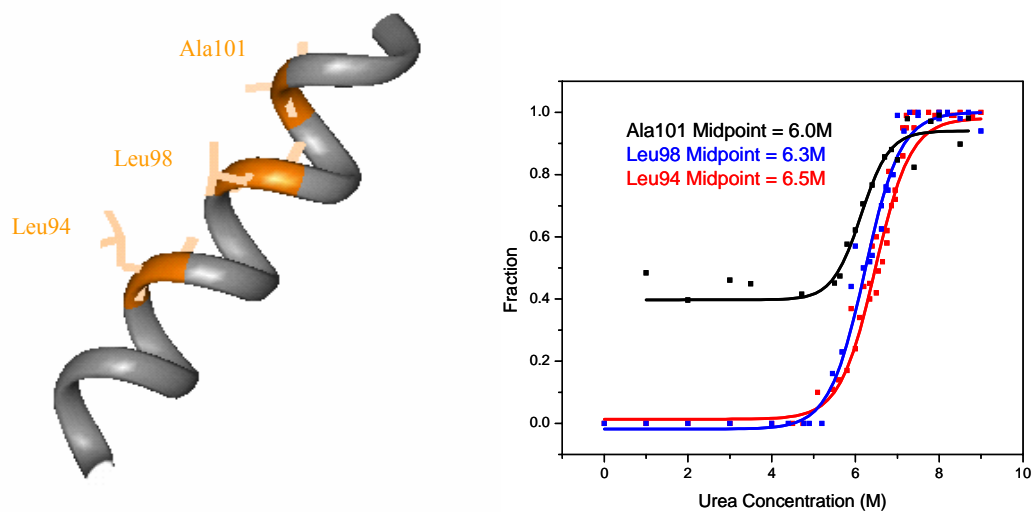
Figure 5.27. The unfolding of the C-terminal helix in urea (see Figures 5.13-5.18). The data presents a sequential unfolding of the C-terminal helix from the C-terminus. Note not only the C-terminal fraying observed at d_3-Ala101 at low concentrations of urea, but also the possibility of residual structure present at this residue at high urea concentrations.

The C-terminal Helix. The three residues in the C-terminal helix also indicate, once again, the striking cooperativity when the protein is unfolded in guanidine hydrochloride. This is in contrast to many recent studies indicating the preferential stability of the C-terminal helix when unfolded under similar conditions (Bai, 1995 and Russell, 2000). Even more surprising is that cooperativity is retained despite clear evidence of C-terminal fraying at d_3-Ala101. The urea data, however, does indicate that the C-terminal helix is more stable than the other motifs measured in this study. Moreover, not only does the urea data imply the sequential unfolding of different structural motifs throughout cytochrome *c*, but it also reveals the sequential unfolding of the C-terminal helix itself. The unwinding of the C-terminal helix from the C-terminus is shown in Figure 5.27. The C-terminal fraying at d_3-Ala101 is further evidence that the C-terminus of the helix possesses decreased stability and this fraying has been observed in other studies (Baxter, 1999). A recent amide-exchange NMR study also showed the C-terminal helix is more stable in the middle than at the ends, in agreement with the data presented here, but this previous work was not able to distinguish between residues Leu94 and Leu98 (Baxter, 1997). The ability of the carbon-deuterium technique to identify small differences in structure between nearly adjacent residues offers a nice example of the increased structural resolution offered by this technique.

Despite the extensive folding studies carried out on cytochrome *c*, some new and interesting observations have resulted from applying this technique. Results of the Met80 loop show that at least part of the loop not only undergoes the same

intermediate as the one observed at Met80, but also populates an intermediate conformation and that it is different in guanidine hydrochloride and urea. Furthermore, data for the 60's helix present the possibility of a new cooperative folding unit consisting of the Met80 loop, the 60's helix and possibly the region in between. Lastly, the high resolution data obtained in the C-terminal helix of the protein is encouraging in light of the need for a folding technique that can offer increased structural resolution of partially folded species.

5.5 References

Akiyama, S., Takahashi, S., Ishimori, K. and Morishima, I., Stepwise Formation of α-Helices During Cytochrome *c* Folding. *Nat. Struct. Biol*., **2000**, *7*, 514.

Bai, Y. W., Sosnick, T. R., Mayne, L., and Englander, S. W., Protein Folding Intermediates – Native State Hydrogen Exchange. *Science*, **1995**, *269*, 192.

Baldwin, R. L., How Does Protein Folding Get Started? *Trends Biochem. Sci*., **1989**, *14*, 291.

Baxter, S. M., Boose, T. L., and Fetrow, J. S., ^{15}N Isotopic Labeling and Amide Hydrogen Exchange Rates of Oxidized Iso-1-Cytochrome *c*. *J. Am. Chem. Soc*., **1997**, *119*, 9899.

Baxter, S. M., and Fetrow, J. S., Hydrogen Exchange Behavior of [U-^{15}N]- Labeled Oxidized and Reduced Iso-1-cytochrome *c*. *Biochemistry*, **1999**, *38*, 4493.

Bushnell, G. W., Louie, G. V., and Brayer, G. D., High Resolution 3-Dimensional Structure of Horse Heart Cytochrome *c*. *J. Mol. Biol*., **1990**, *214*, 585.

Cinelli, S., Spinozzi, F., Itri, R., Finet, S., Carsughi, F., Onori, G., and Mariani, P., Structural Characterization of the pH-Denatured States of Ferricytochrome *c* by Synchrotron Small Angle X-Ray Scattering. *Biophys. J*., **2001**, *81*, 3522.

Colon, W., and Roder, H., Kinetic Intermediates in the Formation of the Cytochrome *c* Molten Globule. *Nat. Struct. Biol*., **1996**, *3*, 1019.

Elove, G. A., Bhuyan, A. K., and Roder, H., Kinetic Mechanism of Cytochrome c Folding: Involvement of the Heme and Its Ligands. *Biochemistry*, **1994**, *33*, 6925.

Fersht, A. R., Optimization of Rates of Protein Folding: the Nucleation-Condensation Mechanism and its Implications. *Proc. Natl. Acad. Sci. U. S. A*., **1995**, *92*, 10869.

Filosa, A., Wang, Y., Ismail, A. A., and English, A. M., Two-Dimensional Infrared Correlation Spectroscopy as a Probe of Sequential Events in the Thermal Unfolding of Cytochromes *c*. *Biochemistry*, **2001**, *40*, 8256.

Goto, Y., Calciano, L. J., and Fink, A., Acid-Induced Folding of Proteins. *Proc. Natl. Acad. Sci. U. S. A*., **1990**, *87*, 573.

Jackson, S. E. and Fersht, A. R., Folding of Chymotrypsin Inhibiter 2. 1. Evidence for a Two-State Transition. *Biochemistry*, **1991**, *30*, 10428.

Jeng, M., Englander, S. W., Elove G. A., Wand, A. J. and Roder, H., Structural Description of Acid-Denatured Cytochrome *c* by Hydrogen Exchange and 2D NMR. *Biochemistry*, **1990**, *29*, 10433.

Lyubovitsky, J. G., Gray, H. B., and Winkler, J. R., Mapping the Cytochrome *c* Folding Landscape. *J. Amer. Chem. Soc.*, **2002**, *124*, 5481.

Myer, Y. P., Conformation of Cytochromes. III. Effect of Urea, Temperature, Extrinsic Ligands, and pH Variation on the Conformation of Horse Heart Ferricytochrome *c*. *Biochemistry*, **1968**, *7*, 765.

Otzen, D. E., Itzhaki, L. S., Elmasry, N. F., Jackson, S. E. and Fersht, A. R., Structure of the Transition State for the Folding/Unfolding of the Barley Chymotrypsin Inhibitor 2 and its Implications for Mechanisms of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.*, **1994**, *91*, 10422.

Pace, C. N., The Stability of Globular Proteins. *Critical Reviews in Biochemistry*, **1975**, *4*, 1.

Ptitsyn, O. B., Protein Folding: Hypothesis and Experiments. *J. Protein Chem.*, **1987**, *6*, 273.

Qi, P. X. R., Beckman, R. A., and Wand, A. J., Solution Structure of Horse Heart Ferricytochrome *c* and Detection of Redox-Related Structural Changes by High Resolution H-1 NMR. *Biochemistry*, **1996**, *35*, 12275.

Russell, B. S., and Bren, K. L., Denaturant Dependence of Equilibrium Unfolding Intermediates and Denatured State Structure of Horse Ferricytochrome *c*. *J. Biol. Inorg. Chem.,* **2002,** *7*, 909.

Shastry, M., and Roder, H., Evidenc for Barrier-Limited Protein Folding Kinetics on the Microsecond Timescale. *Nat. Struct. Biol.*, **1998**, *5*, 385.

Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I. M., Stewart, J. M., and Baldwin, R. L., Nature of the Charged-Group Effect on the Stability of the C-Peptide Helix. *Proc. Natl. Acad. Sci. U. S. A.*, **1985**, *82*, 2349.

Smith, J. S., and Scholtz, M. J., Guanidine Hydrochloride Unfolding of Peptide Helixes: Separation of Denaturant and Salt Effects. *Biochemistry*, **1996**, *35*, 7292.

Tanford, C., Contribution of Hydrophobic Interactions to the Stability of Globular Confirmation of Proteins. *J. Am. Chem. Soc.*, **1962**, *84*, 4240.

Thomas, Y. G., Goldbeck, R. A., and Kliger, D. S., Characterization of Equilibrium Intermediates in Denaturant-Induced Unfolding of Ferrous and Ferric Cytochromes *c* using Magnetic Circular Dichroism, Circular Dichroism, and Optical Absorption Spectroscopies. *Biopolymers*, **2000**, *57*, 29.

Uversky, V. N. and Fink, A. L., The Chicken-Egg Scenario of Protein Folding Revisited. *FEBS Lett.*, **2002**, *515*, 79.

Williamson, M. P., Many Residues in Cytochrome *c* Populate Alternative States Under Equilibrium Conditions. *Proteins: Structure, Function and Genetics*, **2003**, *53*, 731.

Xu, Y., Mayne, L. and Englander, S. W., Evidence for an Unfolding and Refolding Pathway in Cytochrome *c*. *Nat. Struct. Biol.*, **1998**, *5*, 774.