Conformations of unfolded and partially folded peptides and proteins probed by optical spectroscopy

A Thesis

Submitted to the Faculty

of

Drexel University

by

Andrew Michael Hagarman

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

May 2010

Dedications

I would like to dediacte this work to my family who supported me throughout this entire process. To my mother, Patricia, whose undying love and encouragement helped to get through some frustrating times. To my father, James, whose direction and guidance in my early years led me to study chemistry. To my brothers, Patrick and Michael, for their continued support. To my fiancée, Maria, and my son, Matthew, for being the most wonderful things in my life and in the final years dealing with the long hours necessary to complete this process. To my wonderful dog, Jaco, for accepting a decreased amount of walks so I could stay at work later.

I would also like to dedicate this work to my advisor and mentor, Dr. Reinhard Schweitzer-Stenner, for his continued support and guidance through this program, and the unconditional help that was offered along the way. And finally, to all of my colleagues, especially, Thomas Measey, who was there every step of the way and made this journey fun.

Acknowledgements

I would like to acknowledge the people with whom we collaborated and worked with to complete these studies. I would like to thank John Trunk at Brookhaven National Labs for his continued help at beamline U11 to measure SRCD spectra of peptides and cytochrome c. I would like to thank Dr. Harald Schwalbe and his research group for their collaboration, especially Daniel Matthieu and Marie Anders for their assistance in preparing the peptides measured in the NMR experiments and for running the NMR experiments in Frankfurt, Germany. I would like to thank Dr. Tobin Sosnick for enlightening discussions and reference to his coil library website, and Dr. Franc Avbelj for access to his coil library. I would like to thank Dr. Monique Laberge for her collaboration with running MD simulations for the cytochrome c projects. I would like to thank Dr. Carmichael Wallace for his collaboration in preparing the site directed cytochrome c mutants used in our studies.

Table of Contents

List of Figures

List of Tables

Introductory remarks

The work presented herein is in partial requirement for the completion of a PhD thesis from Drexel University Chemistry Department under the guidance of Professor Reinhard Schweitzer-Stenner and comprises two separate projects. The goal of the first project was to determine the intrinsic conformational propensities and distributions of individual amino acids, using optical and NMR spectroscopies. The purpose of the second project was to examine the conformational flexibility and functionally relevant heme distortions of cytochrome c, using optical spectroscopy, in the framework of establishing a baseline comparison for cytochrome c structural changes in biological processes.

Abstract

Conformational plasticity in biomolecules gives rise to unique characteristics. How a protein folds into its native three-dimensional structure has been a long investigated mystery, but it is tied into conformational sampling of polymeric chains of amino acids. One critical piece of information, i.e. intrinsic conformational propensities of individual amino acids in a polypeptide chain, encodes the folding energy landscape of a protein. This funneled landscape facilitates the ability for proteins to fold spontaneously, without randomly sampling the ensemble of accessible conformations. Also, the fact that an essential protein in the electron transport chain, cytochrome c, undergoes conformational changes in many biological processes underscores the importance of conformational heterogeneity in biomolecules.

In order to estimate intrinsic conformational propensities of individual amino acids we use a protocol that allows us to simulate experimental isotropic Raman, anisotropic Raman, FTIR and vibrational circular dichroism spectra and a set of six NMR J-coupling constants by using a superposition of statistically weighted two-dimensional Gaussian distributions representing sterically allowed regions of the Ramachandran space. We use the host-guest motif glycine-x-glycine, where x is confined to a set of amino acids representing aliphatic (A, V, L, M, I), aromatic (F, Y), charged (E, D, R, K) and polar (S, T, C, N) residues. The selection of glycine hosts was imperative to minimize nearest-neighbor effects that would modulate the conformational propensity of the central residue. We have thus confirmed alanine's high propensity to adopt dihedral angles in the PPII distribution and determined that aliphatic and positively charged residues

preferred PPII more than the other investigated residues. Aromatic residues were shown to populate PPII and β-distributions somewhat equally. An anionic residue, glutamic acid behaves like an aliphatic residue at acidic and neutral pH values, which implies a negligible effect of the charged side-chain environment. A similar residue, aspartic acid, with a shorter side-chain, behaves completely different, populating the dihedral angle distributions found in various β-turns and a yet thoroughly unconsidered region centered in the top right quadrant of the Ramachandran plot. The conformational ensembles of polar residues, S, T, C and N were shown to be significantly populated by the dihedral angles found in type I/I' and II/II' β-turns. The distributions obtained here differ somewhat from coil library distributions. Our propensities significantly disagree with what is found from theoretical studies, in particular, MD simulations. We conclude that the high helical propensity found in MD studies is an effect from nearest and non-nearest neighbor interactions.

The conformational changes that cytochrome c undergoes in biological processes have become an area of increasing interest, due to its apparent peroxidase activity, interactions with cardiolipin containing membranes and its involvement in programmed cell death. The oxidized state of this protein adopts a multitude of conformations *in vitro* as well as *in vivo*. The structural stability and thermodynamics of ferricytochrome c have thus been the focus of a substantial amount of research about this protein. We have used far-ultra violet and visible absorption and electronic circular dichroism spectroscopy to structurally and thermodynamically characterize the intermediate states adopted in the process of thermal and alkaline induced (partial) unfolding. The

extent of band splitting caused by electrostatic interactions between the heme group and the protein was determined by a vibronic analysis of the B-band ECD and absorption spectra. We demonstrated that the states III_h and IV are thermodynamically and also conformationally different, contrary to the current belief. With respect to ferricytochrome *c* our results suggest that the overall structure is maintained in the intermediate state populated above 323 K. Conformational changes might involve increasing distances between the heme and aromatic residues such as F82 and a reduced nonplanarity of the heme macrocycle. The band splitting is substantially reduced in the unfolded states, but the heme environment encompassing H18 and the two cysteine residues 14 and 17 is most likely still intact and covalently bound to the heme chromophore. Most importantly, we have shown the need for a comprehensive thermodynamic analysis of all native and non-native states of ferricytochrome *c* under well-defined conditions which would explicitly consider the fact that not only the "ground state" populated at room temperature but also the thermally excited, partially or mostly unfolded states are still pH dependent.

Cytochrome c is in a class of proteins with high redox potentials. Its comparatively high redox potential is stabilized by a hexacoordinated central iron atom in the heme c which is coordinated to a sulfur of a methionine in the surrounding protein matrix at the distal coordination site, as well as by interactions with the internal electric field created by ionizable groups within the heme pocket. Thus, deformations of the heme group are functionally relevant in modulating the redox potential. We have used polarized resonance Raman spectroscopy to exploit the depolarization ratios and normalized

intensities of Raman active bands in the low frequency Soret excited Raman spectrum for an estimation of planar and non-planar deformations of the heme active sites in three different reduced cytochrome c isoforms; horse, chicken and a mutated – to avoid aggregation - *Saccromyces Cerevisae* (yeast). We thus obtained that ruffling was the largest deformation experienced by all investigated hemes with chicken being the most ruffled folloed by horse heart and yeast. Concerning the saddling deformations, the heme group in horse heart was the most followed by yeast, then chicken. We determined that the heme c of chicken experienced the most doming followed by horse heart and yeast. Finally, the heme group of horse heart was determined to be the most propellered. The main saddling and ruffling deformations from crystal and MD structures compare well with our results, whereas MD simulations better account for smaller deformations like doming and propellering, due to the fact that the uncertainty of crystal structures coordinates relates to high error in small deformations.

Section 1 Intrinsic Propensities and Conformational Distributions of Individual Amino Acids Probed by Optical and NMR Spectroscopies

Chapter 1 Background

1.1 Protein Folding

Protein folding, the way a polypeptide chain folds into a stable three-dimensional structure, is a process that is governed by the identity of the amino acids in the chain as well as the influences of the cellular environment. In many cases, the folding and unfolding process of proteins (Figure 1) is critical for proteins to perform biological and cellular function. Unfolded and partially folded states of peptides and proteins are topics of continuous research. Issues such as a) relating the folded to the unfolded state of proteins (*1-4*), b) intrinsic disorder in proteins and its intriguing relation to cellular function and protein misfolding, aggregation and disease (*5-8*) and c) the challenging endeavor of predicting the native folded structure of a protein from the primary amino acid sequence (*9-12*), are studied extensively throughout the protein folding community. The relevance of these points stems from the necessity to understand the dynamics of the folding pathways of proteins and the energetics of folding and unfolding in connection with the potential energy landscape of protein folding pathways. By understanding protein folding pathways, the structures and energetics of transition and intermediate states can be determined. These states, in some cases, are the most functionally relevant and may lead to disease formation, such as in the case of Alzheimer's disease, where the misfolding of amyloid beta leads to aggregation and

depending on the pathway of aggregation, may cause the onset of this neurodegenerative disease (*13*).

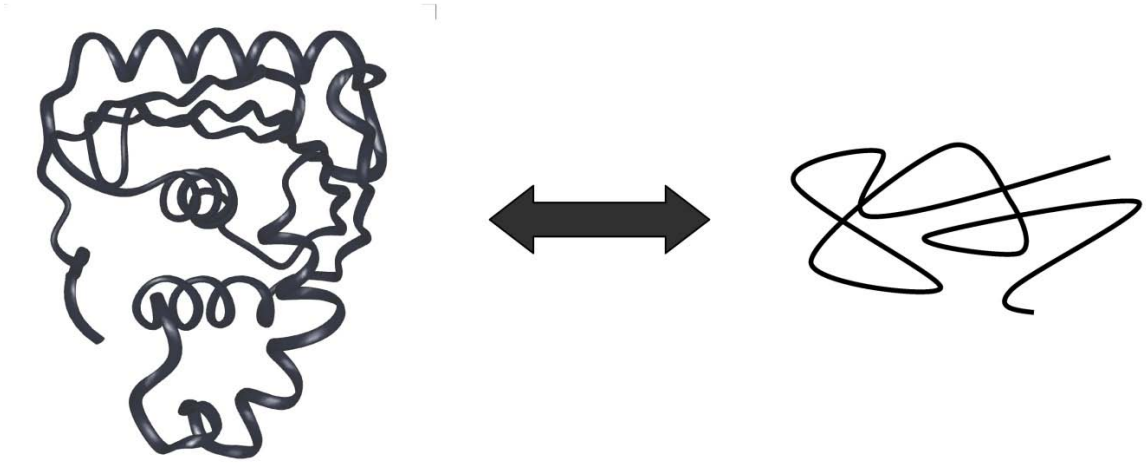


Figure 1: Folded→ Unfolded representation of a protein.

For a long period of time, the research on protein folding was guided by a single theoretical argument and a very decisive experiment. First, in the 1960's, Levinthal stated that a protein sampled 3^N accessible conformations during the folding period (*2*), 3 being the number of conformers adopted by each residue and N is the number of amino acids in the protein, so that the time it takes a protein to fold after a random walk through the energy landscape would be longer than the time of the universe. Based on the fact that proteins fold spontaneously and on near millisecond timescales, Levinthal concluded that the conformational sampling of protein folding was, in fact, not random. Second, Anfinsen et al. discovered that the unfolding of ribonuclease A is reversible in that it refolds to its native structure if exposed to folding promoting conditions (*1*). This observation led to the prediction that the folded state of proteins

can be determined from their primary amino acid sequence. How proteins reach their folded state remains the subject of continual research.

Wolynes and colleagues proposed a funneled free energy of folding landscape under folding conditions (*3, 4*). Proteins start in their unfolded state at a high energy (top of the funnel) and follow a gradient down the slope of the funneled landscape, like a 'rolling ball' until the protein reaches its final folded state at the energy minimum of the funnel. As the protein follows the gradient down the funnel it accumulates favorable interactions that lower its energy, which promote a more compact structure and reduce the conformational entropy. In this funneled landscape protein folding gathers speed as the transition from U to N progresses, which allows proteins to fold reliably and cooperatively. In this model, the protein avoids frustration of non-native contacts, which would trap the protein in small wells on the wall of the free energy funnel, because these contacts are not energetically favorable and hence do not interact at all, whereas native contacts are favorable, thus making the walls of the funnel smooth. The folding and unfolding process is directly related to the free energy landscape of the protein. In cases where the landscape has the correct shape, only a minimal amount of conformations are accessible. Since this free energy surface is preset by the primary amino acid sequence, nature has designed a way for proteins to fold rapidly and into stable structures. Many small, single domain proteins have been characterized to undergo a two-state, reversible (U$\leftrightarrow$N) and often times cooperative, folding mechanism, where at any point in time the protein is either fully folded or fully

unfolded. A thermodynamic intermediate is not significantly populated or cannot be detected within experimental uncertainty in such reactions.

Energetic factors that govern the process of protein folding are still a matter of debate. Hydrophobic collapse and hydrogen bonding are two energetic factors that predominate folding although the extent of which each contributes to the folding process is unknown. Much of the argument in favor of hydrogen bonding contributing to folding comes from results that state that intrapeptide hydrogen bonding is favored over peptide-water hydrogen bonds (*14-16*). Although Kauzmann, in particular, has argued that hydrogen bonding contributes little and may actually disfavor protein folding and thus the hydrophobic effect provides the major free energy contribution to protein stability (*17*). Some results have supported this notion. Klotz and Franzen measured approximately 0 kcal/mol for the enthalpy of the interpeptide hydrogen bond in N-methyl acetamide (*18*), which was corroborated by Susi and Ard for a different system (*19*). These finding led the authors to conclude that interpeptide hydrogen bonds in aqueous solution have a small intrinsic stability. Currently the protein folding community is at odds over these issues, but recently hydrogen bonding has been considered more as the energetic factor promoting protein folding. It has been shown for some proteins which were natively unfolded and were forced to fold upon the addition of organic osmolytes, that the peptide backbone is the dominant contributor to folding (*20-22*). Since the backbone lacks hydrophobic groups, it was concluded that hydrogen bonding drives protein folding in these cases (*23*).

1.2 Conceptual Characterization of Folded and Unfolded Proteins and Peptides

The Ramachandran plot was developed as a tool for identifying regular secondary structures in proteins and peptides. It allows the identification of sterically allowed pairs of dihedral angles, ϕ and ψ (Figure 2), which amino acid residues can adopt in a polypeptide chain.

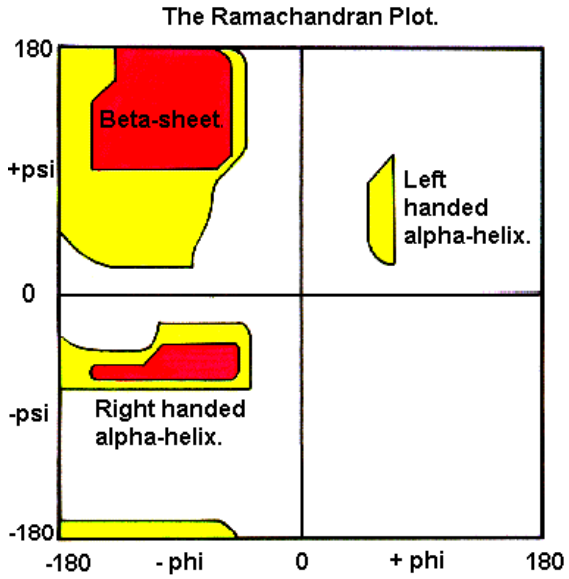


Figure 2: Representative Ramachandran plot showing the sterically allowed regions found mainly in folded native proteins.

The simplest peptide unit with a side chain other than hydrogen where the bonds encompassing the dihedral angles are not free rotating entities is the so-called alanine dipeptide (Figure 3). The term residue applies in this context to describe an amino acid in a polypeptide with a definite set of dihedral angles. The first so-called dihedral angle of residue *i* is phi (ϕ) (Figure 3), which is defined by a rotation around the N(H)-C_α bond by two planes designated by the atoms ($C'_{(i-1)}$, $N(H)_{(i)}$, $C_{\alpha(i)}$) and ($N(H)_{(i)}$, $C_{\alpha(i)}$, $C'_{(i)}$). The second dihedral angle of residue *i*, psi (ψ), is defined by a rotation around the C_α-C' bond by two planes designated by the atoms ($N(H)_{(i)}$, $C_{\alpha(i)}$, $C'_{(i)}$) and ($C_{\alpha(i)}$ $C'_{(i)}$, $N(H)_{(i+1)}$).

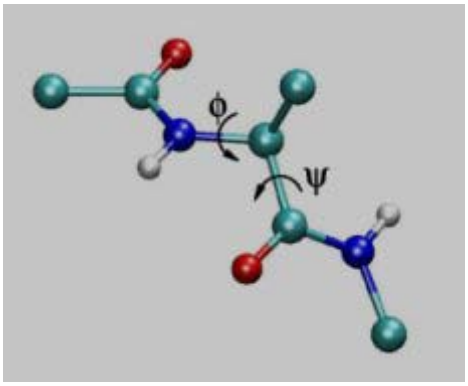


Figure 3: Acetyl-Alanine-Methyl Amide (alanine dipeptide) showing the backbone dihedral angles ϕ and ψ.

The alanine dipeptide was therefore used as a model system to identify the regions of the Ramachandran space for which the repulsive interactions between the atoms of the peptides are minimized, so that they can be sampled by the alanine residue with detectable probability. The thus obtained sterically allowed regions of the Ramachandran plot (space) are dominated by extended (upper left quadrant) and right-handed helical (lower left quadrant) conformations (Figure 2) (*24*). Ramachandran was able to calculate the sampling of pairs of ϕ and ψ angles within remarkable accuracy using a simple steric model, which has been substantiated upon the development of high resolution techniques such as X-ray crystallography. Conversely, the calculated conformations allowed for glycine, whose side chain induces negligible steric constraints, were much more numerous and covered a majority of the (ϕ,ψ) space. The Ramachandran plot for alanine was thought to represent the conformational sampling for other naturally occurring amino acids excluding glycine and proline, due to the similarities in plots calculated for the remaining naturally occurring amino acids.

In qualitative agreement with Ramachandran's rather simple modeling, folded native proteins are indeed mostly comprised of α helices and β sheets (strands) with pairs of dihedral angles in the lower and upper left quadrants of the Ramachandran plot, respectively (Figure 2). Other know structures such as a variety of turns and coil regions that occur in folded proteins populate other various regions of the conformational subspace (Figure 2), but residues do not populate these structures as frequently as α helices and β sheets (strands). The conformational space is much more restricted for these structures than the 'free alanine space' due to hydrogen bonding. The situation changes drastically when disordered and unfolded proteins are considered. The Ramachandran plots that were calculated by Flory show a comparatively large distribution of the allowed conformational space for alanine (*25*). A large portion of the total conformational ensemble occurs in the upper left part of the plot in Flory's calculations. The same plots from coil libraries (Figure 4) where the sampling of alanine in regular structures (α-helices and β-sheets) is excluded from a search of the Protein Data Bank so that the distributions represent the sampling in unfolded segments, indicate a somewhat more restricted conformational ensemble for alanine than what Flory and Ramachandran calculated. The extended region (upper left quadrant) is indeed sampled though not entirely at the canonical β sheet (strand) coordinates. A subsection of the upper left quadrant of the Ramachandran space becomes predominantly populated with ϕ angles centered at ~-70°. This preferential sampling implies a dominant structure of unfolded peptides and proteins.

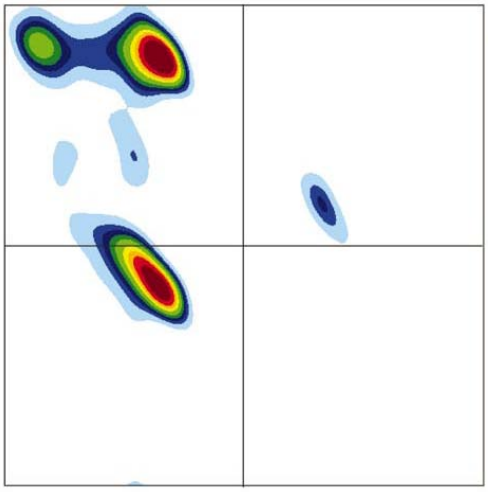


Figure 4: Taken from ref (*26*) and modified. Coil library distribution for alanine.

1.3 Unfolded State

A detailed characterization of the unfolded and folded states is necessary for a complete understanding of the folding mechanism. Short lived intermediates, or transient states were previously thought to be more important, but much attention has been devoted to the unfolded state recently. The structure of folded peptides and proteins is generally characterized by high resolution techniques such as X-ray crystallography and NMR spectroscopy. Due to the dynamic nature of disordered systems they cannot be crystallized which makes x-ray crystallography impractical for structural characterization of unfolded peptides and proteins. Nevertheless, much attention has been devoted recently to studying and understanding unfolded proteins. Five different types of systems have been used to study the unfolded state of peptides and proteins, as stated previously by Kallenbach and coworkers (*27*): 1) Proteins in the presence of denaturing agents (*28*), 2) Coil libraries, where residues in regular secondary structures have been excluded from the PDB (*26, 29-33*), 3) Short peptides, so that no regular secondary structure can be maintained (*34*), 4) Charged oligopeptides that cannot support a folded

structure due to electrostatic interactions of the side chains (*35*), and 5) Natively unfolded proteins, which have been referred to frequently as intrinsically disordered proteins (IDP) (*36*).

Classical textbook knowledge of the unfolded state of proteins stems from polymer theory, which treats unfolded proteins like a flexible polymer in a so-called good solvent (*25, 28*). First, Tanford related the hydrodynamic radius of an unfolded protein to the expression, n^{γ}, where n is the number of monomer units (residues) (*28*). The exponent, γ, was consistent with values found for polymers (*25*). Flory then hypothesized that there were no single preferential conformations in unfolded proteins, based on calculations of Ramachandran maps of alanine (*25*). These notions implied that the sterically accessible part of the conformational space is randomly and equally sampled by individual residues, so that the unfolded state of proteins can be described by the term 'random coil' (*37*). However, the term 'statistical coil' is preferred by Scheraga to not get the false impression that the conformational space is isoenergetic as it was suggested by Ramachandran and Flory (*38*). This led to the conclusion that nucleation parameters of secondary structure formation do not depend significantly on intrinsic properties of amino acids, since local minima in the Ramachandran space associated with the different secondary structures are sampled with comparable probability by all residues. The nucleation process in helix↔coil transitions is accounted for by the parameters, σ and v, in Zimm-Bragg (*39*) and Lifson-Roig (*40*) theories, respectively. For example, Zimm-Bragg theory states that the product, σs, where s is the statistical weight of a helical conformation, reflects the probability of forming a helix in a segment of

three amino acids (*39*). The Zimm-Bragg model accounts for cooperativity so that the probability of a given residue to adopt either a helix or coil is affected by the preceding unit. This implies that propagation of a helix is more favorable opposed to nucleation of a helix from a coil unit, i.e. $\sigma << 1 < s$. Lifson-Roig theory refined the previous model so that the condition of long-range interactions stabilizing helices was taken into account. Where the Zimm-Bragg theory was able to only consider two consecutive units, this extension allows for the consideration of three consecutive residues. In folded peptides and proteins, propensity values for individual residues reflect the fact that nucleation has already occurred. Without the presence of nearest-neighbor and long-range interactions the nucleation parameter should reflect intrinsic properties of individual amino acids. The Ramachandran plots calculated from the simple steric model do not indicate any difference in nucleation parameters from different residues while coil library distributions indicate the opposite.

1.4 Random Coil Theory

The 'Random Coil Theory' originates from classifications of polymer chains. The random coil conformation is one that samples all accessible conformations randomly in the absence of an environment that promotes stabilization. It was accepted for many years since the work of Flory and Tanford that unfolded proteins are interpreted as random coils due to the fact that the hydrodynamic radius of the proteins investigated by Tanford in strong denaturant related to that of polymeric random coils (*28*). The notion of unfolded proteins being 'random coils' has been contested many times based on the observation of residual structure in unfolded peptides and proteins. Experimental

evidence challenging the 'random coil' theory was reported by Tiffany and Krimm more than 40 years ago (*41*). These authors measured ultra-violet circular dichroism (UVCD) spectra of poly-L-glutamic acid and poly-L-lysine in their charged states, and based on the similarities of their spectra with that of poly-L-proline (Figure 5), concluded that charged amino acids predominantly sample the polyproline II (PPII) region of the Ramachandran space.

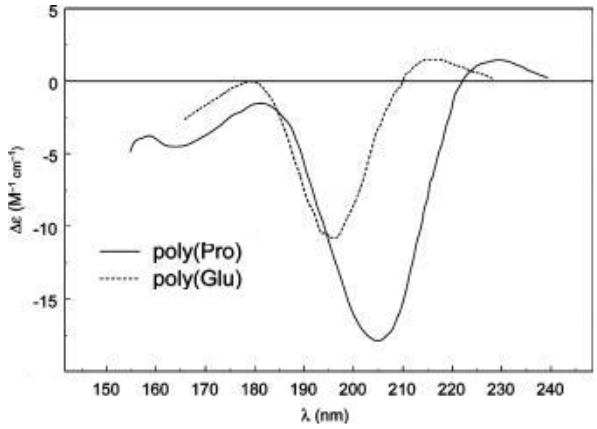


Figure 5: Taken from ref (*42*) and modified. Far-UVCD spectra for poly-L-proline (solid line) and polyglutamic acid (dotted line).

Krimm and coworkers also observed the resemblance between CD spectra of proline peptides and proteins unfolded using denaturing agents, leading them to hypothesize that the conformational manifold of unfolded peptides and proteins is dominated by PPII conformations (*43*). PPII structures exhibit dihedral angles similar to those adopted by poly-L-proline with all the residues in the *trans* conformation. PPII conformations cluster around canonical dihedral angles, ϕ=-70° and ψ=150° considered to be found in the distribution of extended structures, seen in Figure 2 (*44*). The structure of a peptide

with these dihedral angles can be visualized in Figure 6 compared to the structures of an α-helix and an antiparallel β-strand.

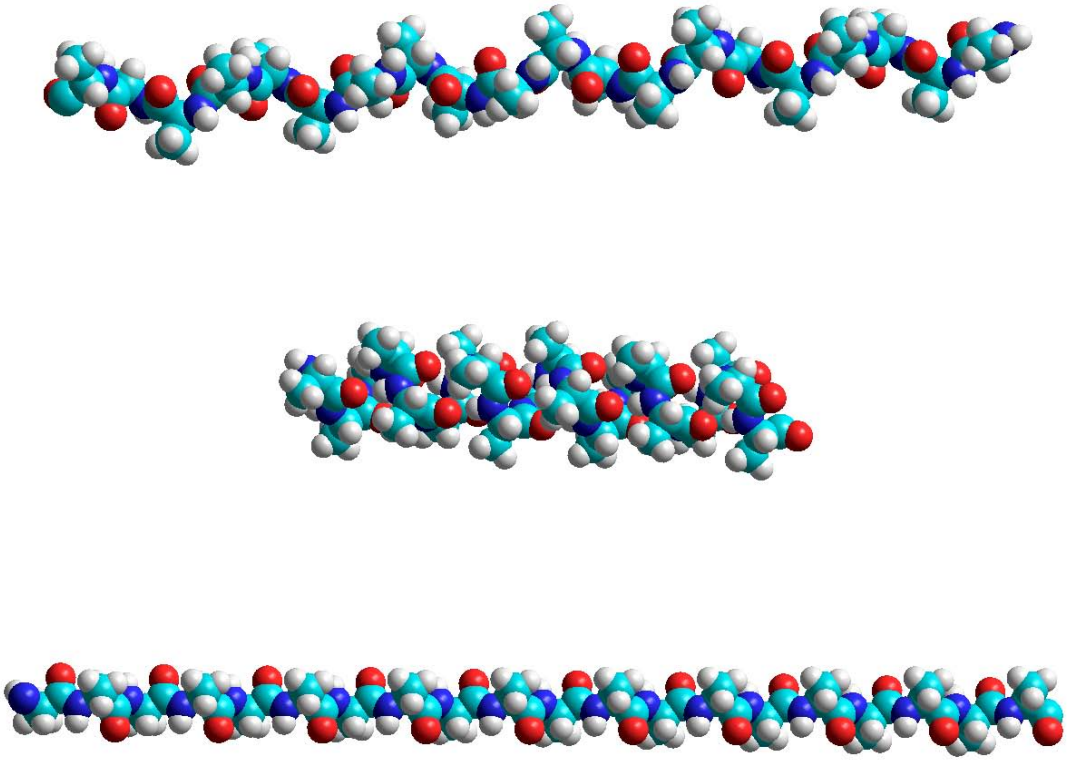


Figure 6: Alanine oligomers (20-mer) in representative conformations: polyproline II (top), α-helix (middle) and antiparallel β-strand (bottom).

No more evidence disproving the random coil theory accumulated after the work of Tiffany and Krimm, and thus this notion died down. It was only after Dukor and Keiderling presented similar findings as Tiffany and Krimm, using VCD spectroscopy on the same peptides, that the random coil theory became a central point of contention again (*45*). The notion of residual structure in the unfolded state of peptides and proteins then experienced a revival. The preferential sampling of PPII-like

conformations by individual amino acids has been identified many times since the exploratory work of Tiffany and Krimm (*35, 46-52*).

The determination by Tanford and Flory of unfolded peptides and proteins being 'random coils' and the contrasting determination of residual structure in the unfolded state led to the establishment of the so-called "reconciliation problem". This led to many studies that attempted to resolve the reconciliation problem. Measurements of radius of gyration with small angle x-ray scattering (*53, 54*), NMR (*55*) and viscometric techniques (*28*) consistently produce an exponent of ~0.6 which is consistent with a random coil polymer conformation. Rose and coworkers, however, used Monte Carlo simulations to show that proteins with a high degree of internal structure, in fact, have end-to-end and mean radii of gyration distances that are consistent with random coil behavior (*56*). Thus, they conclude that the scaling properties are not really sensitive to residual structure. Sosnick and coworkers used NMR residual dipolar coupling constants (RDCs) which are more sensitive to measuring local structure in an aim to reconcile this 'problem' (*57*). They formulated a statistical coil model based on conformational sampling from coil libraries and compare the results to experimental RDCs in chemically denatured states. The agreement between these two values is inherent upon the extended region (β and PPII) being predominantly sampled. Their model accounts for the presence of local structure while retaining a random coil like scaling behavior for the radius of gyration. Even if local deviation from random sampling occurs, the Flory Isolated Pair Hypothesis (FIPH) still applies for longer polypeptide chains. The FIPH states that individual pairs of (ϕ,ψ) angles are sterically insensitive to their neighbors

(*25*). Since its establishment, this hypothesis has since been a central idea of contention within the protein folding community.

1.5 Conformational Propensities of Amino Acids

Various strategies to study the unfolded state of proteins and how they relate to the respective folded structures have been developed over the past 40 years. One method of these investigations involved the exploration of conformational propensities of amino acid residues (*9*). By virtue of an understanding of the preference of amino acids and peptide fragments for sampling specific regions of the sterically allowed conformational space associated with regular secondary structures, insight could be gained into how proteins, in their earliest stages, begin to fold. The conformational propensity of a given amino acid is reflected by its preference to sample parts of the sterically allowed regions of the Ramachandran space (Figure 2) (*24*). A quantitative measure of the structural propensity of individual residues can be estimated by determining the occurrence of that residue in stable secondary structures in folded proteins (*9*). This propensity value certainly reflects, although to a limited extent, that of an intrinsic property of the amino acid, but it also reflects the influences of neighboring residues as well as that of long range interactions, so that these propensities do not reflect the nucleation process that occurs in earliest stages of protein folding where long-range interactions have no effect (*58-60*).

The use of coil libraries is generally accepted to estimate conformational propensities of amino acids (*30, 32, 57, 61*). By analyzing coil libraries, which reflect the distributions of dihedral angles for amino acids throughout proteins listed in the Protein Data Bank, the

distributions of dihedral angles for that residue can be visualized (*26, 62*). In some cases residues contained in regular secondary structures have been excluded to focus on segments that are considered unfolded (*30, 32*). Jha et al. extended the restrictions to exclude regular secondary structures and residues neighboring regular structures and found the distributions to be dominated by sampling of the extended region (*61*). These distributions were considered to be a reliable estimation of conformational propensities, due to the assumption that non-local interactions were 'averaged out' when a sufficiently large basis was being sampled (*63*). However, as demonstrated (*31, 32*), the distributions of individual residues significantly differ depending on whether regular structures are considered (*61, 64*). Even in the simplest cases where all regular structure and amino acids neighboring regular structures were excluded, the distributions of individual residues are affected by neighboring amino acids, so that the distributions do not reflect intrinsic propensities (*27, 30, 65-68*). Figure 7 shows an example of this variable sampling extracted from Sosnick's sampling library (*69*). The distributions for a central amino acid in a tripeptide are extracted from the PDB on this website. The program allows one to consider all types of nearest neighbors encompassing all types of structures. The plots in Figure 7 exhibit the conformational distribution of alanine flanked by either alanines or glycines. The N denotes that only segments of coil:unfolded were chosen to be sampled in the program. Apparently, the two distributions are significantly different. This observation and ample experimental and theoretical evidence of the occurrence of nearest- and even second-nearest-neighbor interactions between residues clearly indicate the necessity to experimentally

determine intrinsic properties of amino acids, which apparently cannot be determined from coil libraries. Furthermore, these plots indicate a breakdown of the isolated pair hypothesis.

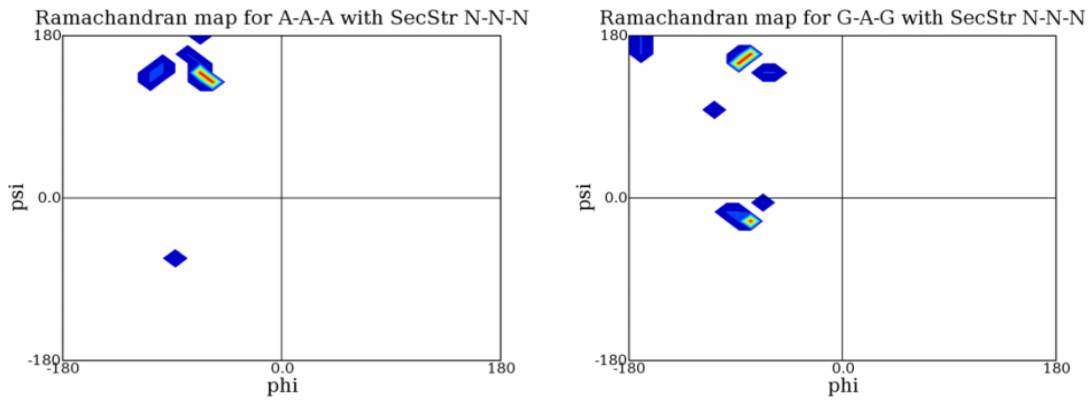


Figure 7: Conformational distributions for alanine with alanine and glycine nearest neighbors extracted from Sosnick's website (*69*) sampling only coiled regions.

Intrinsic propensities of individual residues can be better inferred from unfolded proteins, though nearest neighbor interactions still perturb these properties. A common method developed over the last 15 years to estimate intrinsic propensities of individual amino acids is to utilize short peptides with no certain secondary structure as model systems for the unfolded state, as first proposed by Dyson and Wright (*34*). Short peptides lack the ability to adopt folded structures due to the absence of long range interactions needed to form regular folded structures. Hence, these peptides can be studied under native folding conditions while remaining unfolded. Much attention has been devoted recently to the study of residual structure in short peptide segments. Ho and Dill showed with the use of replica exchange molecular dynamics that, out of a set of 133 peptides (octamers) from six different proteins, nearly 1/3 of the set converged to a preferred structure. Of the converged set, 85% of the investigated fragments

resembled native like structures (*70*). Short peptides are flexible and dynamic entities, though preferential sampling of the (ϕ, ψ) space for individual residues indeed exists, as confirmed by experimental and theoretical results (*64, 71-73*). Many experimental studies have thus far focused on spectroscopic data of predominantly alanine containing peptides. Alanine has been frequently argued to be a model for the backbone in unfolded proteins and peptides due to its abundance and its high propensity to form helical conformations in proteins. This residue has been used for investigating the breakdown of the isolated pair hypothesis as well (*65*). While the studies about alanine-based peptides have been nearly exhausted, the remaining naturally occurring amino acids have yet to be thoroughly studied. The same can be said for theoretical studies. In what follows a brief history about the conformational propensity of alanine in short peptides will be detailed. Then, studies about the propensities of residues other than alanine will be highlighted.

1.6 Experimental Studies of Alanine Propensities

Many conflicting results have emerged regarding the conformational sampling of alanine residues in unfolded peptides, though recently a somewhat consistent picture has surfaced based on experimental results. Nearly ten years ago Wouterson et al. estimated, using nonlinear time-resolved vibrational spectroscopy, that trialanine populates PPII with a mole fraction ~0.8 (*74*). Kallenbach and associates then employed NMR and CD spectroscopy to show that the alanine residues in a peptide commonly referred to in the literature as the XAO ($X_2A_7O_2$) peptide, where O is ornithine and X is diaminoisobutyric acid, predominantly adopted a PPII structure (*47*). These results

about the preferential sampling of alanine in short and unfolded peptides have subsequently been corroborated many times. The Kallenbach group concluded as well from and NMR and CD study that alanine in the pentapeptide, GGAGG adopts a PPII structure (*75*). Much work has emerged from the Schweitzer-Stenner group which reported a dominant PPII sampling for alanine in short peptides using a plethora of techniques. The results are as follows. Eker et al. simulated the Raman, FTIR and VCD spectra of tri- and tetraalanine and determined that trialanine exhibits a 50:50 mixture of PPII and β, while the alanines in tetraalanine prefer PPII (*76, 77*). They validated their results with temperature dependent far-UV CD spectra (*78*). Schweitzer-Stenner et al. later constrained their simulations of vibrational spectra with NMR $^{3}J(H^{N}H^{\alpha})$ coupling constants for a series of short alanine containing peptides; AAA, AAAA and AAKA. They reported that alanines in these peptides exhibit PPII fractions of >0.60.

An even shorter unit, namely the alanine dipeptide, was shown to predominantly sample the PPII region as well (*79*). Barron and coworkers used Raman optical activity measurements to show that alanine oligomers from 3-7 residues were predominantly adopting PPII-like structures (*80*). It was also shown using ROA measurements that an alanine based peptide similar to XAO, namely OAO, adopted a PPII-like structure (*80*). Asher and coworkers used UV resonance Raman to show that a longer helical peptide, referred to as AP (AAAAA-$(AAARA)_3$-A) unfolds to a predominantly PPII-like structure (*81*).

The high PPII propensity of alanine reported in the above studies has been challenged based on results from experimental and theoretical studies. Particularly, conflicting results emerged concerning the sampling of PPII in the XAO peptide. Scheraga and coworkers used Monte Carlo simulations to conclude that PPII was not a dominant conformation for the alanines in XAO (*68, 82*). The authors proposed a significant population of various types of turn structures as well as PPII, making the sampling of these residues more of an ensemble (*68*). The Pande group conducted small angle X-ray scattering (SAXS) measurements on this peptide and derived a radius of gyration of 7.4 Å. By assuming a "random walk scaling" between radius of gyration and end-to-end distance, their reported value corresponds to an average end-to-end distance of 18.1 Å which is inconsistent with a substantial sampling of PPII, for which one would get an end-to-end distance of around 32 Å (for pure PPII, radius of gyration = 13.1 Å, end-to-end distance = 32.04 Å) (*83*). Schweitzer-Stenner and Measey then reported a compromise for these inconsistencies by simulating experimental vibrational spectra using J-coupling constants and the parameters of Scheraga and coworkers as constraints (*84*). The authors concluded that the alanines in XAO indeed sample approximately 50% PPII-like structures though the peptide adopts a heterogeneous distribution, with a population of 26% for a variety turn-structures and 23% for β-strand. Using these results they were also able to reproduce the end-to-end distance reported by Pande and coworkers. Also, their results indicate that alanine might adopt turn-like structures in the presence of charged residues.

Recent studies about the conformational sampling of alanine in short peptides have grown increasingly complex, in that they combine various techniques like NMR and vibrational spectroscopy as well as theoretical methods so that, for instance, distributions of conformations can be considered instead of representative structures. Graf et al. used a set of 7 NMR scalar coupling constants that relate to ϕ and ψ dihedral angles of a given residue independently through different Karplus relations combined with distributions derived from MD simulations to show that the central residue in trialanine populates mostly a PPII-like distribution with a mole fraction of 0.90 (*85*). They extended the study to longer alanine oligomers peptides and confirmed that alanine preferentially adopts dihedral angles in the PPII region of the Ramachandran space in agreement with the results of Barron and coworkers (*80*). Schweitzer-Stenner then used the coupling constants of Graf et al. as constraints to simulate the experimental Raman, IR and VCD amide I' profiles based on an excitonic coupling model considering distributions of conformations and obtained a PPII fraction of 0.84, which is close to the value reported by Graf et al. (*86*). The studies about the conformational propensity of alanine in short peptides are numerous, and even though the results do not agree in all cases, most cases indicate a high PPII propensity for alanine, which is becoming the accepted result in the experimental community.

1.7 Theoretical Studies of Alanine Propensities

Results from theoretical studies about the conformational propensity of alanine in short peptides vary much more than those from experiments. Results from MD simulations vary the most, depending on the choice of force field. Han et al. reported a dominant

sampling of a PPII structure with dihedral angles of (ϕ,ψ) = (-93°, 128°) for the alanine dipeptide explicitly hydrated with four water molecules, by comparing experimental and *ab initio* calculated Raman, VCD and ROA spectra (*49*). Using the GROMOS96 force field, Mu et al. found that the extended region (PPII and β-like structures) was predominantly sampled by solvated trialanine (*87*). The authors then used an OPLS force field with a TIP5 water model and reported that the individual populations for PPII and β-like structures were 0.65 and 0.12, respectively (*88*). Osman and coworkers and Rose and coworkers reported that PPII is the default conformation for polyalanines peptides in water (*89, 90*). Beck et al. reported 'intrinsic' propensities of amino acids using *in lucem* MD simulations in the host-guest motif GGXGG (*91*). The authors report a significantly lower PPII propensity (0.16) and an α-helical propensity above 50% for alanine. Results from numerous experimental techniques cannot rationalize a high helical propensity for alanine in short peptides. Best and Hummer recently used modified force fields to predict the sampling of polyalanine peptides (*92*). The authors used the Amber03 and Amber99SB and modified Amber03 and Amber99SB force fields to calculate the secondary structure fractions of Ala_5. They reported that Ala_5 populates the PPII-region, at most, at a mole fraction of 0.5 using the original and corrected force fields (*92*). The authors previously stated that the force fields only needed refinement, and that the differences between their results and those obtained from experimental studies, specifically the Graf *et al.* study (*85*), were due to differences in parameters in the respective Karplus equations (*93*). Contrary to other MD studies, Garcia and coworkers found that alanine oligomers shorter than 8 residues adopt, to a significant extent, a

PPII structure, using a reparameterized AMBER force field (*94*). Interestingly, Sosnick and coworkers used many different force fields to calculate the MD populations for the central residue in trialanine and found a significantly reduced PPII propensity for most force fields, except for the OPLS-AA-97 force field, which reported >80% PPII propensity (*66*). Pappu and coworkers used Monte Carlo methods to estimate the conformational sampling of individual residues in short peptides (*73*). They estimate that alanine has a nearly equal preference for PPII and β-like structures and is rather context independent.

Although Molecular Dynamics and other theoretical techniques provide an essential method for visualizing the structure and dynamics of peptides and proteins, there are certainly drawbacks to these methods. Much of the discrepancies in results from MD calculations reflect the variability of empirical force fields. The most commonly used force fields, though based on the same empirical potential energy function, contain different associated parameters. These force fields are often calibrated by transferring the results of quantum mechanical calculations on peptides. This method is impractical due to the limitations in such calculations, where a high degree of accuracy can only be achieved for a small number of atoms. Even in the case of short peptides, *ab initio* and DFT calculations require a large amount of time if solvent (water) is considered explicitly. Thus, many results are reported on the gas phase structures, which in turn compromises the calibration of force field parameters (*95*). Consequently, these force fields underestimate electrostatic interactions when applied to larger molecules like proteins and longer peptides.

1.8 Rationale for the Stabilization of PPII

A physical rationale for amino acid residues to prefer PPII-like conformations in the unfolded state of peptides have not yet been fully revealed. Solvent hydration (*46, 50, 96, 97*), steric effects (*65, 98*), side chain-backbone interactions (*46, 97*) or some combination of these (*99*) have all been proposed as influencing the bias of unfolded peptides to sample the extended region of the Ramachandran space. Coil libraries even suggest that residues sampling the PPII conformation are in highly solvated regions of proteins (*100*). Details are briefly described below. Though many of these assumptions are speculative, there are some experimental data to confirm the notion that solvent hydration plays a significant role in conformational sampling of individual residues. Eker et al. reported a shift in PPII/ β equilibrium upon dissolving AcAA in DMSO as opposed to D_2O, with the former populating β-like structures more than the latter (*96*). The same group reported an effect upon solvating with D_2O instead of water, proposing that the hydrogen bonding network plays an important role in stabilizing PPII structures (*78*). Creamer and coworkers reported that D_2O stabilizes PPII more so than H_2O as well (*101*). Kallenbach and coworkers reported a conformational change upon solvating $AcGGAGGNH_2$ in neat TFE, which confirms the notion of the hydration shell playing a direct part in stabilizing PPII structures (*102*). The authors reported that the PPII fraction of this peptide decreases linearly according to solvent polarity with the following hierarchy; water > methanol > ethanol > 2-propanol. Avbelj and coworkers reported a substantial increase in the $^3J(H^N, H^\alpha)$ coupling constants for alanine and valine in non-aqueous solvents, CCl_4 and 1,4-Dioxane. This large increase in these values reflects a destabilization of PPII.

With respect to the stabilization of the PPII conformations by water, the results of DFT calculations for an alanine dipeptide complexed with water molecules indicate a pivotal role of a bridge comprised of two hydrogen bonded water molecules connecting the CO and NH groups of neighboring peptide units (*49*). An alternative explanation for the stability of PPII has been provided by Drozdov et al. (*103*) These authors performed a very thorough Monte Carlo simulation with an all-atom OPLS force field to explore the conformations sampled by an alanine dipeptide in explicit solvent (i.e. water). Their results suggest that water molecules connecting CO and NH groups of adjacent peptides via hydrogen bonding do not contribute significantly to the solvation free energy of the peptide. Moreover, these simulations revealed that solvation itself does not yield to a preference of PPII over helical conformations. The contribution of solvation to the stabilization of PPII was instead found to be a rather indirect one, in that it neutralizes electrostatic interactions between non-bonded pairs of atoms. As a consequence, steric interactions favoring PPII become predominant. It is unclear whether the results of Drozdov et al. contradict those of Garcia, who performed MD simulations for various polyalanine peptides in explicit water (*94*). This work led the author to conclude that PPII is stabilized because it allows an optimal hydration of the considered peptides. Whatever the correct model is, it can be expected that the water molecules in the hydration shell can be perturbed by the amino acid side chains. It is noteworthy in this context that the results of Drozdov et al. indicate that the equilibrium between PPII and β is, in fact, directly affected by solvation. In agreement with this notion, Law and

Daggett recently concluded, based on MD results, that water bridges do not stabilize PPII structures (*104*).

Rose and coworkers introduced a method in which hydrophobic accessible solvent areas, termed CHASA, could be determined, to rationalize the stabilization of PPII by water molecules hydrogen bonded to N and O backbone atoms in polypeptide chains (*105*). They determined that alanine in a polyproline II conformation allows the proper geometry for the NH and CO of the peptide backbone to be solvated by water, in turn, decreasing the hydrophobic accessible solvent area. With longer β-branched side-chains, the solvation of the backbone is disrupted and thus PPII is destabilized. Polar side-chains counterbalance the PPII stabilization due to their solvation free energies.

1.9 Conformational Propensities of Residues other than Alanine

Conformational propensity studies on residues other than alanine are limited in number. However, some experimental and theoretical results have emerged in the past. Eker et al. used the amide I band profiles to obtain average conformations for a series of AXA peptides (X= different natural amino acids) (*71*). The authors reported that residues such as P, K and E exhibit a preference for PPII, residues V, W, Y, S, H and M exhibit a β-strand preference and residues A and L exhibit a mixture of PPII and β. Hagarman et al. used a more quantitative approach based on far-UVCD spectra and NMR $^{3}J(H^{N},H^{\alpha})$ coupling constants to explore the conformations of AX and XA dipeptides in terms of a two-state model by using representative conformations for PPII and β-strand (*106*). Their results suggest that only alanine exhibits a clear PPII propensity (~0.63), while residues with bulky aliphatic side chains, such as valine and isoleucine,

clearly prefer a β-strand conformation. However, these results should be treated with caution because the propensities might reflect the special environment of the C-terminal. In a different context, namely P_3XP_3, Kelly et al. found the qualitative PPII propensity hierarchy of X residues to be P>Q>A>(G,L,M)>N>(I,V) (*97*). Shi et al. (*72*) used an approach similar to that of Hagarman et al. (*106*) to analyze the NMR $^3J(H^N,H^\alpha)$ constants of different X-residues in the host-guest peptide motif $AcGGXGGNH_2$. Their two-state model was based on representative conformations inferred from the maxima of PPII and β-strand type distributions in the coil library of Avbelj and Baldwin (*64*). The obtained results suggest that nearly all amino acid residues, with the exception of histidine and methionine, exhibit PPII propensities above 0.5. If these results were true, PPII would be the default conformation in unfolded proteins and peptides. Pappu and coworkers confirmed this notion for most amino acid residues, but predicted a substantially reduced PPII propensity for valine and isoleucine (*73*). Their comprehensive study about propensities of individual amino acids included calculations of propensities with a model solely based on peptide-solvent and repulsive peptide-peptide interactions. The author's results include the propensities all of the natural amino acids considering also different neighboring residues (G, A and P).

1.10 Protocol

The protocol used for the current study combines NMR and vibrational spectroscopic techniques to determine the conformational distributions of various amino acid guest residues in GxG (x labels the guest residue) peptides (*48, 86*). The selection of GxG over GGxGG host-guest motifs was preferable because it is more amenable to our techniques

in the following ways. The addition of extra residues on the termini complicates the amide I' profiles. One might argue that end effects may modulate the conformational propensity of the central residue, but, as shown before by Eker et al. where a set of tripeptides were determined to have the same structures in three different protonation states, these effects are negligible (*77*). End effects actually resolve the splitting of the amide I' profiles further. Below, we address the sampling of the central residue in GEG upon changes in protonation state of the peptide. We selected glycine as neighbor to minimize nearest neighbor interactions, so that the obtained propensities can really be considered as intrinsic, which is not strictly the case for the propensity values obtained for homopeptides such as trialanine and trivaline. The determination of intrinsic propensities of amino acid residues in water is necessary for construction of a reference system based on which their context-dependence in more complex systems can be explored. Moreover, the knowledge of intrinsic propensities will allow computational biochemists to calibrate molecular mechanics force fields so that they can be used for the simulations of unfolded, intrinsically disordered peptides and the formation of secondary structures (*92, 107, 108*). For the guest residue, x, we selected a set of amino acids with aliphatic (A, V, L, I), aromatic (F and Y), charged (E, R and D) and polar (S, T, C and N) residues. We measured six different NMR J-coupling constants of the central amino acids, which exhibit different ϕ and ψ-dependencies (*85*). The Karplus relations of these coupling constants can be seen in Figure 8.

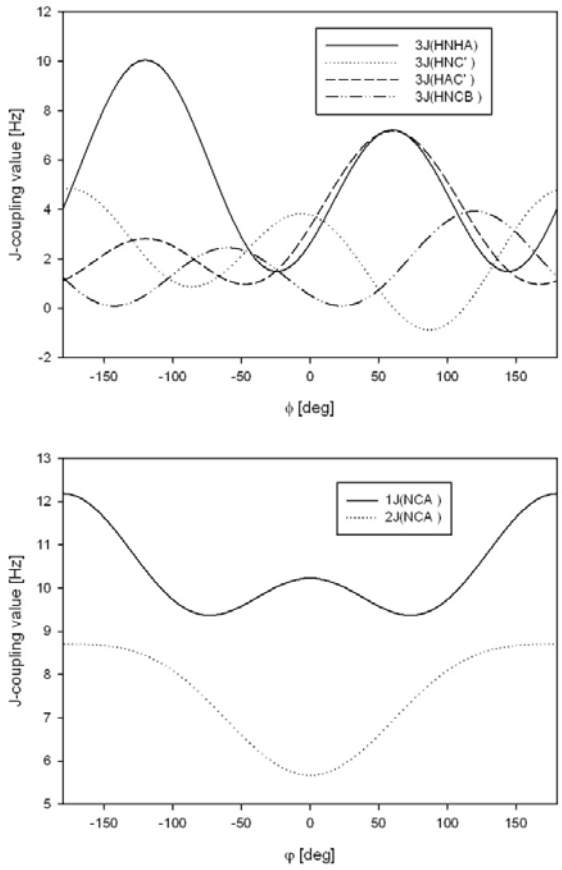


Figure 8: Graphical representation of the ϕ (top) and ψ (bottom) angle dependent coupling constants as indicated in the legends, based on different Karplus relations published in refs (*109-113*).

Additionally, we measured the amide I' profiles by means of polarized Raman, FTIR and vibrational circular dichroism (VCD) spectroscopies for the set of peptides. A global analysis of these data were performed by fitting them to distribution models which can be described as a superposition of two-dimensional Gaussian functions reflecting the peaks of different secondary structure conformations in the Ramachandran plot, i.e. PPII, different types of β-strands, right- and left-handed helical and a variety of turn structures. This model is very sensitive to variations in centers and fractions of distributions due to the simultaneous fit of four ϕ value and two ψ value J-coupling constants with different Karplus parameters. The results obtained from the analysis of

the residues in the first set were complemented by analyzing the amide I' band profiles and the canonical $^{3}J(H^{N},H^{\alpha})$ coupling constant of K and M. In addition to obtaining the propensities of the investigated residues for adopting the considered conformations we also obtained the distribution functions for the latter, which together reflect the respective Gibbs energy landscape of the residue in an aqueous environment.

Chapter 2 Theoretical Background

2.1 Amide I'

The so-called amide I band is one of the most intense peptide bands in FTIR and Raman spectra. Other peptide normal modes include amide II, III, IV, V and S. All of these normal modes are structurally sensitive, though amide I is the most frequently used to probe secondary structure of peptides and proteins due to its intensity in vibrational spectra (*114, 115*). The amide I is naively thought to be purely a CO stretching (COs) vibration though it is known from normal mode analyses that its eigenvector contains contributions from NH in-plane bending (NHipb) and CN stretching (CNs) (*116, 117*) (Figure 9).

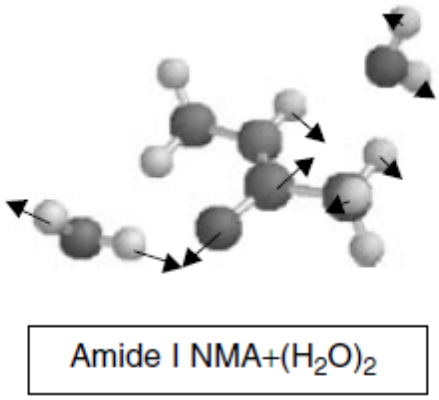


Figure 9: Representation of the amide I normal mode in N-methylacetamide (NMA) hydrogen bonded to two explicit waters.

When peptides are investigated in aqueous solution this normal mode mixes with the bending mode of water within the hydration shell (*49, 116, 118, 119*). This effect and the contribution of NHipb are eliminated if deuterated water (D_2O) is used as the solvent. Concomitantly, a slight downshift of the amide I occurs. This band is now commonly referred to as amide I'. The frequency (wavenumber) of this normal mode is

highly structure sensitive, due to hydrogen bonding and concomitant changes in the force constant of the carbonyl bond (*120*). Hydrogen bonding is relevant in this context because of its affect in stabilizing PPII structures. Moreover, the coupling between adjacent amide I oscillators in polypeptides makes this band an optimal candidate in structural studies (*121-123*). The latter effect is the key determinant of the amide I profiles of short, unfolded peptides. Measey et al. enhanced the ability to reproduce experimental amide I profiles by reporting the IR oscillator strength, the Raman cross-section and the wavenumber of local amide I modes for different side-chains and neighboring residues (*124*). In what follows we describe how we exploit the vibrational (excitonic) coupling between amide I modes for a structural analysis of short peptides.

2.2 Excitonic Coupling

The coupling of adjacent amide I modes is a well-known phenomenon (*121*). The delocalization of local amide I modes occurs via through-bond and through-space coupling (*122, 123*). There are two concepts which describe the coupling of adjacent amide I modes. The first one is classical and based on normal mode calculations. The eigenvectors calculated for the higher and lower frequency amide I modes in a two oscillator system are attributed to in-phase and out-of-phase combinations of the two interacting modes (Figure 10).

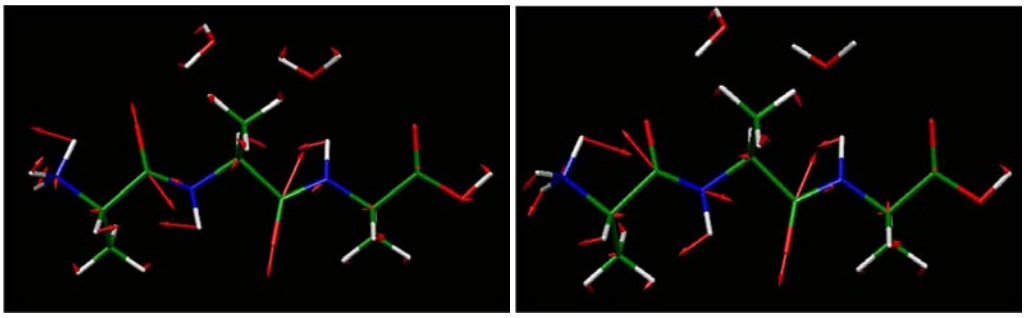


Figure 10: In-phase (left) and out-of-phase (right) combinations of adjacent amide I oscillators in trialanine.

The second concept of interaction utilized to describe the coupling between vibrational oscillators is quantum mechanical, where the excited states of the involved normal modes are mixed, owing to the aforementioned through bond and through space coupling. Such an excitonic coupling has been frequently described in the literature (*52, 125-130*). Hence, this section will be confined to a basic description using only the most important concepts of the excitonic coupling model and how it can be applied to tripeptides. The Schrödinger equation for amide I of a polypeptide containing *n* residues can be written as:

$$\left(\hat{H}_0 + \hat{H}_{ex}\right)|\chi'\rangle = E|\chi'\rangle \tag{1}$$

where the excitonic Hamiltonian, $\hat{H}_{ex}$, accounts for coupling of local oscillators. $\hat{H}_0$ is the unperturbed Hamiltonian, E is the eigenenergy and $|\chi'\rangle$ is the wavefunction. For a tripeptide the total Hamiltonian (two amide I oscillators) is written as:

$$\hat{H}_T = \begin{pmatrix} \tilde{\nu}_1 & \Delta_{12} \\ \Delta_{21} & \tilde{\nu}_2 \end{pmatrix}$$

(2)

The elements $\tilde{\nu}_1$ and $\tilde{\nu}_2$ are the wavenumber positions of the local uncoupled amide I oscillators and the off-diagonal element Δ denotes the excitonic coupling operator. Torii and Tasumi calculated the (ϕ,ψ) dependence of the expectation values of Δ, by using the results of *ab initio* based normal mode calculations on a blocked glycine tripeptide (*122*). The diagonalization of the total Hamiltonian yields vibrational eigenfunctions expressed as linear combinations of the vibrational basis functions χ_j of the uncoupled individual oscillators:

$$\chi_i' = \sum_{i=1}^{n-1} \alpha_{ij} \chi_j$$

(3)

The coefficients α_{ij} represent the amplitude of the *i*-th excitonic wave function at the *j*-th residue and *n* corresponds to the number of individual amide I oscillators. The corresponding Raman tensor and transition dipole moment of the *i*-th excitonic state are written as:

$$\hat{\alpha}_i' = \sum_{i=1}^{n-1} \alpha_{ij} \hat{\alpha}_j$$

(4a)

$$\vec{\mu}_i' = \sum_{i=1}^{n-1} \alpha_{ij}\vec{\mu}_j$$

(4b)

The Raman tensor $\hat{\alpha}_i$ and the dipole moment $\hat{\mu}_i'$ for the transition into the *i-th* excitonic state can be used to calculate the isotropic Raman, anisotropic Raman and IR intensity, respectively. In order to perform this calculation however, the Raman tensors $\hat{\alpha}_j$ and dipole moments of individual amide I normal modes have to be transformed into a common coordinate system. Schweitzer-Stenner has previously detailed the rotation of local amide I oscillators and their Raman tensors into a basis reference system positioned at the C-terminal amide nitrogen (*125*). Subsequently, the isotropic Raman, anisotropic Raman and IR amide I band profiles can be calculated as a superposition of Gaussian profiles representing the respective bands of the excitonic modes:

$$I_{iso}(\Omega) = \sum_{i=1}^{n-1} \left[45\beta_{s,i}^2 f_i\right]$$

(5a)

$$I_{aniso}(\Omega) = \sum_{i=1}^{n-1} \left[7\gamma_{s,i}^2 f_i\right]$$

(5b)

$$I_{IR}(\Omega) = \sum_{i=1}^{n-1} \left[I_{s,i}^{IR} f_i \right]$$

(5c)

where

$$f_i = \frac{1}{\sigma_i \sqrt{2\pi}} e^{\frac{-(\Omega-\Omega_i)^2}{2\sigma_i^2}}$$

(6)

Ω_i is the eigenenegry of the *i-th* excitonic state and σ_i is the corresponding halfwidth of the band profile. The expressions $\beta_{s,i}^2$ and $\gamma_{s,i}^2$ are derived from the Raman tensor where:

$$\beta_{s,i}^2 = \frac{1}{9} (Tr\ \hat{\alpha}_i')^2$$

(7a)

$$\gamma_i^2 = \frac{1}{2}[(\hat{\alpha}'_{xx,i} - \hat{\alpha}'_{yy,i})^2 + \left(\hat{\alpha}'_{yy,i} - \hat{\alpha}'_{zz,i}\right)^2 + \left(\hat{\alpha}'_{zz,i} - \hat{\alpha}'_{xx,i}\right)^2] +$$

$$\frac{3}{4}[(\hat{\alpha}'_{xy,i} - \hat{\alpha}'_{yx,i})^2 + (\hat{\alpha}'_{yz,i} - \hat{\alpha}'_{zy,i})^2 + (\hat{\alpha}'_{zx,i} - \hat{\alpha}'_{xz,i})^2]$$

(7b)

and

$$I_i^{IR} = 9.2 * 10^{-39} \frac{\vec{\mu}_i \cdot \vec{\mu}_i}{\tilde{\nu}_0}$$

(8)

2.3 Calculation of VCD Profiles

The calculation of the VCD profile is less straightforward than that of the IR and Raman profiles. The rotational strength for a peptide with *n-1* residues can be expressed as (*125*):

$$R_i = Im\left[\sum_{j=1}^{n-1} \alpha_{ij}\vec{\mu}_j \left(\sum_{k=1}^{n-1} \alpha_{ij}\vec{m}_j - \frac{i\pi}{2}\left(\sum_{l=1}^{n-2}\sum_{m=1}^{n-1} \tilde{\nu}_{ij}\vec{T}_{lm} \times (\alpha_{il}\vec{\mu}_l - \alpha_{im}\vec{\mu}_m)\right)\right)\right]$$

(9)

Where $\vec{m}_j$ is the intrinsic magnetic moment associated with the amide I of the *j-th* residue and $\vec{T}_{lm}$ is the distance vector of the *l*- and *m-th* local oscillators. The term $\tilde{\nu}_{ij}\vec{T}_{lm} \times (\alpha_{il}\vec{\mu}_l - \alpha_{im}\vec{\mu}_m)$ on the right hand side of equation 9 describes the chirality brought about by the magnetic moment at residue *l* induced by the transition dipole moment at residue *m*. Higher order terms on the right hand side of the equation would describe interactions that would occur when n>2. It is important to note that the rotational strength depends on the amplitude of α_{ij} of which the *j-th* residue contributes to the *i-th* excitonic state which implies that it is directly proportional to the strength of excitonic coupling. Hence, the VCD profile can be calculated as a superposition of Gaussian bands:

$$\Delta\varepsilon(\Omega) = \frac{\tilde{\nu}_0}{2.3 * 10^{-38}} \sum_{i=1}^{n-1} [R_i f_i]$$

(10)

Where $\tilde{\nu}_0$ is the first moment of the amide I' band profile. This formalism has been used several times to successfully calculate Raman, IR and VCD amide I' band profiles (*76, 86, 125, 131*) and will subsequently be used in the work presented herein.

Chapter 3 Materials and Methods

3.1 Peptides

L-glycyl-L-alanyl-L-glycine (GAG), L-glycyl-L-glutamic acid-L-glycine (GEG), L-glycyl-L-lysyl-L-glycine (GKG), L-glycyl-L-phenylalanyl-L-glycine (GFG), L-glycyl-L-methionyl-L-glycine (GMG), glycyl-L-tyrosyl-L-glycine (GYG), and L-glycyl-L-leucyl-L-glycine (GLG) were purchased from Bachem Biosciences Inc. (King of Prussia, PA) and used without further purification. L-glycyl-L-valyl-L-glycine (GVG), L-glycyl-L-aspartic acid-L-glycine (GDG), L-glycyl-L-seryl-L-glycine (GSG), L-glycyl-L-Isoleucyl-L-glycine (GIG), L-glycyl-L-threonyl-L-glycine (GTG), glycyl-L-asparagyl-L-glycine (GNG), and L-glycyl-L-cystyl-L-glycine (GCG) were custom synthesized by Genscript Corp. (Piscataway, NJ) at >98% purity and subsequently dialyzed in an acidic medium in 100 MWCO dialysis bags purchased from Spectrum Laboratories Inc. (Rancho Dominguez, CA) and lyophilized in a home-made lyophilization apparatus to remove traces of trifluoroacetic acid (TFA) used in the synthesis of the peptides. TFA has a peak in the FTIR and Raman spectra at ~1670 cm^{-1} which contaminates the amide I' band and complicates the deconvolution of the spectra, so elimination of TFA from the peptide samples is necessary. The peptides were dissolved in D_2O at a 0.2 M peptide concentration and were acidified (<pD 2.2) with small aliquots of DCl. The pD values were calculated using the method of Glasoe and Long, where pD= pH + 0.4 (*132*).

3.2 Isotopically Labeled Peptides for NMR Experiments

For NMR-experiments, GAG, GVG, GFG, GEG, GLG, GSG, GCG, GDG, GIG, GNG, GRG, GTG and GYG peptides were ^{13}C carbonyl labelled at residue 1, uniformly ^{13}C and ^{15}N labelled at residue 2 and ^{15}N labelled at residue 3. The C-terminal residue was manually attached

to a chlorotrityl resin. The synthesis was carried out on an Applied Biosystems 433A peptide synthesizer using standard Fmoc chemistry. Peptides were purified by reversed-phase HPLC. Products were characterized using electrospray ionization mass spectrometry and analytical HPLC. Resins and activating agents were purchased from Novabiochem (Darmstadt, Germany). All Isotopically labelled Fmoc-protected amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA). All solvents were of analytical grade and dried over molecular sieves if necessary. All NMR samples were prepared by dissolving the peptides in 90% H_2O / 10% D_2O and the pH was adjusted to 2 with HCl.

3.3 Vibrational Spectroscopy

The set ups for polarized Raman, FTIR and VCD experiments have been explained in detail in previous publications (*84, 133*). For polarized Raman experiments a Ramascope, a confocal Raman microscope with a back thinned CCD camera, from Renishaw Inc. (Hoffman Estates, IL) was used. Raman spectra were collected polarized parallel (X) and perpendicular (Y) to the polarization of the excitation laser beam. The Raman samples were excited with a 514.5 nm laser beam obtained from a mixed gas tunable Ar/Kr laser obtained from Spectra Physics (Santa Clara, CA). VCD and FTIR spectra were measured in a ChiralIR instrument obtained from Biotools Inc. (Jupiter, Fla.). Aqueous samples were placed in a 20 µm calcium fluoride (CaF_2) demountable cell for measurement. Both spectra (VCD and IR) were collected in one measurement, using an 8 cm^{-1} resolution and a total of 720 minutes of acquisition time. FTIR spectra were background corrected using a D_2O spectrum.

3.4 NMR Spectroscopy

NMR measurements of isotopically labelled GAG, GVG, GFG, GSG, GEG, GLG, GCG, GDG, GIG, GNG, GRG, GTG and GYG have been carried out on a Bruker 400 MHz Avance II spectrometer equipped with a 5 mm HCN triple resonance probe with z-Gradients. All measurements were performed at 298 K. Spectra were acquired and processed using the program TopSpin Version 2.1. Acquisition and processing parameters are given Table 1.

Proton assignment for all GxG peptides could be obtained from 1D spectra and $^{1}H,^{1}H$-TOCSY spectra using a DIPSI-2 (*134*) mixing sequence (Figure 11).

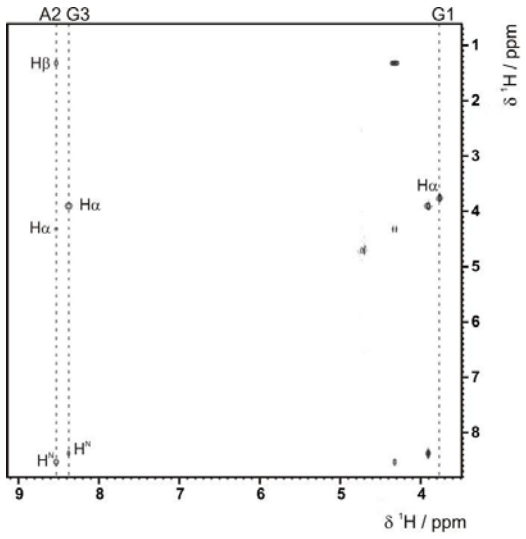


Figure 11: Taken from ref (*48*) and modified. Example for a $^{1}H,^{1}H$-TOCSY NMR experiment measured to obtain resonance assignments (shown for H_3N^+-GAG-COO^-).

The acquisition and processing parameters for all multi-dimensional NMR are available Table 1.

experiment	F[a]	sw_1[b]	t_1[c]	sw_2[d]	t_2[c]	n_t[d]	d_1[e]	mix[f]	^{15}N[g]	^{13}C[g]	td_1[h]	td_2[h]
^{13}C decoupled 1D	400	4400	8192	-	-	64	2.5	-	-	74.6		128k
$^{1}H,^{1}H$-TOCSY	400	4000	4096	4000	256	8	1.5	60	-	-	4096	1024
J-modulated $^{1}H,^{15}N$-HSQC	400	4400	2048	1280	64	4	2.0	5 - 200	117	176 / 51	2048	256
CO-coupled (H)NCαHα	400	2200	4096	1459	128	8	2.0	-	119.5	54	4096	1024
HNCO[CA]-E.COSY	400	4000	4096	450	64	32	2.0	-	119.5	176	4096	1024
HNHB[CB]-E.COSY	400	4000	4096	4000	128	32	2.0	-	119.5	120	4096	1024
soft HNCA-COSY	400	4400	4096	3217	64	32	1.5	-	119.5	54	4096	1024

Table 1: Acquisition and processing parameters for multidimensional NMR experiments. [a] Proton Lamor Frequency in MHz, [b] spectral width in Hertz, [c] number of complex points, [d] number of transients, [e] relaxation delay in seconds, [f] mixing time in milliseconds: DISPSI-2 mixing time in the case of the TOCSY, delay for evolution of NC^{α} coupling in the case of the J-modulated $^{1}H,^{15}N$-HSQC, [g] carrier frequency in ppm, [h] number of processing points for the Fourier Transform

The $^{3}J(H^{N},H^{\alpha})$ coupling constant was obtained from a ^{13}C-decoupled 1D spectrum using presaturation for solvent suppression. The exact coupling constant was obtained by Lorentzian deconvolution (Figure 12). Depending on the labelling scheme, either only the $^{1}J(N,C^{\alpha})$ or both the $^{1}J(N,C^{\alpha})$ and $^{2}J(N,C^{\alpha})$ coupling were obtained by measuring a J-modulated $^{1}H,^{15}N$-HSQC (*109*). The intensities for different mixing times were fitted to extract the coupling constant (Figure 12). Analysis of the x residue NH resonance yields the $^{1}J(N,C^{\alpha})$coupling. The same analysis could be performed for the C-terminal glycine NH resonance yielding the $^{2}J(N,C^{\alpha})$ coupling constant. The fitting was performed using the program SigmaPlot and the equation stated in Figure 12B. $^{3}J(H^{\alpha},C')$, $^{3}J(H^{N},C^{\alpha})$, $^{3}J(H^{N},C^{\beta})$ and $^{3}J(H^{N},C')$ coupling constants were obtained from E.COSY type spectra (see Figure 12C-F) (*110-113*). Exact peak positions were determined by deconvolution of the

respective 1D trace. A full set of the obtained coupling constants is given in Table 2 and Table 3, in the Results sections.

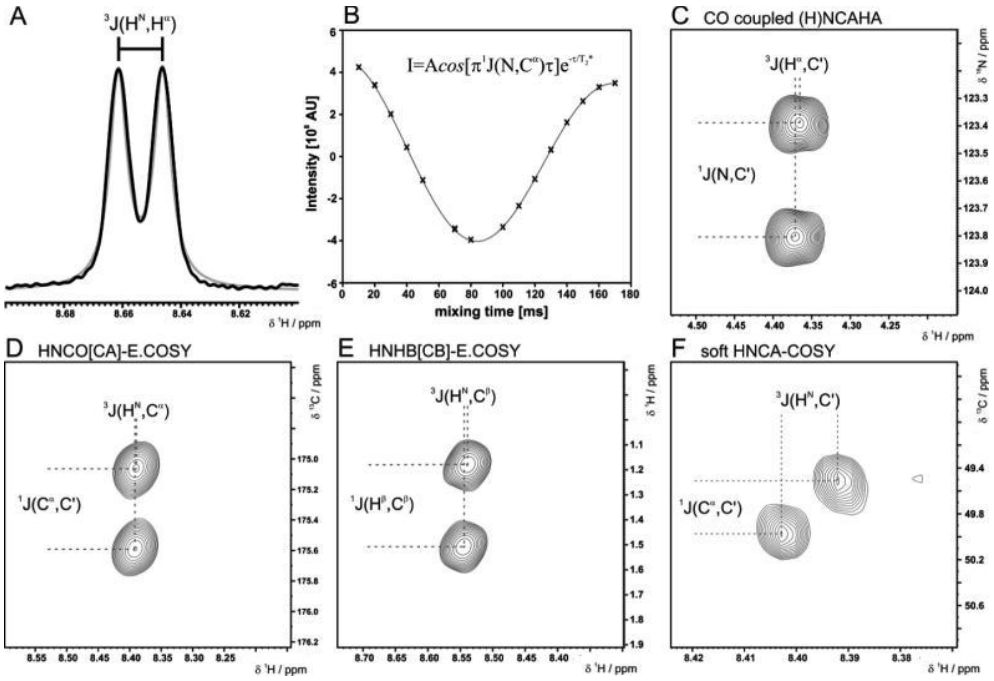


Figure 12: Taken from ref (*48*) and modified. Determination of coupling constants. (A) ^{13}C decoupled 1D for the measurement of $^3J(H^N,H^\alpha)$. The deconvolution result is shown in gray. (B) Intensity fit of J-modulated ^{1}H,^{15}N-HSQCs yielding either only $^1J(N_i,C^\alpha_i)$ or both $^1J(N_i,C^\alpha_i)$ and $^2J(N_iC^\alpha_{i-1})$, depending on the labelling scheme. For the example given, only $^1J(N_i,C^\alpha_i)$ was determined. (C) CO-coupled (H)NCAHA for the measurement of $^3J(H^\alpha,C')$. (D) HNCO[CA]-E.COSY for the measurement of $^3J(H^N,C^\alpha)$. (E) HNHB[HB]-E.COSY for the measurement of $^3J(H^N,C^\alpha)$. (F) Soft HNCA-COSY for the measurement of $^3J(H^N,C')$.

The ^{1}H NMR spectra of GKG and GMG were measured with a 500 MHz Varian FT-NMR instrument equipped with a 5 mm HCN triple resonance probe. All spectra were acquired and processed using Varian's VNMR software (v 6.1), with presaturation applied to suppress solvent signals. Peptides were dissolved in 90% H_2O / 10% D_2O at concentrations of 0.1 M, and acidified with DCl to a pD of 1.5. The D_2O (Sigma-Aldrich) contained 0.05 wt % of 3-(trimethylsilyl) propionic-2,2,3,3-d_4 acid (TSP), which was used as an internal standard. 64 transients were averaged for each sample at 298 K. The

$^{3}J(H^{N},H^{\alpha})$ coupling constants were determined by deconvolution of the amide proton doublets, using the program MULTIFIT (*135*).

3.5 Ultra Violet and Synchrotron Radiation Circular Dichroism

We used a Jacso J810 Spectrapolarimeter purged with gaseous nitrogen to collect Electronic Circular Dichroism spectra. Aqueous peptide samples were measured with a concentration of 10 mM in a 50 um quartz cell purchased from International Crystal Laboratories (Garfield, NJ). The spectra were measured in the range from 180-240 nm with a scan speed of 500 nm/min, a data pitch of 0.1 nm, a 1 sec response and a 5 nm bandwidth on medium sensitivity. Ten accumulations were obtained at 25°C. The temperature was controlled with a Peltier solid-state heating and cooling mechanism. Each spectrum was solvent corrected in the Jasco spectral analysis program.

Synchrotron radiation circular dichroism (SRCD) experiments were performed at Brookhaven National Laboratories in the National Synchrotron Light Source building at beamline U-11 using peptide concentrations ranging from 0.01-0.1 M, a 0.012 cm path length, 1 nm resolution, and an average of three scans. All spectra were manually correct for background variations.

Chapter 4 Data Analysis

The analysis of the amide I' band profiles exploits excitonic coupling between the two local amide I' modes in the tripeptides, which increases the splitting between them and redistributes IR and Raman intensities. The underlying theory explained previously as well as the formalism and the empirical parameters used for the simulation of amide I' band profiles have been described in detail in numerous papers (*52, 127-130, 136*). While earlier studies analyzed amide I' profiles in terms of average or representative conformations (*131*), we recently linked this analysis to a statistical model which describes the conformational manifold of the central residue of tripeptides in terms of an ensemble of superimposed two-dimensional Gaussian distribution functions (*86*):

$$f_j = \left(\frac{\chi_j}{2\pi\sqrt{\left|\hat{V}_j\right|}} \right) e^{-0.5\cdot\left(\vec{\rho}-\vec{\rho}_j^{\,0}\right)^T \hat{V}_j^{-1}\left(\vec{\rho}-\vec{\rho}_j^{\,0}\right)} \tag{11a}$$

where

$$\vec{\rho} = \begin{pmatrix} \phi \\ \psi \end{pmatrix} \tag{11b}$$

and

$$\hat{V}_j = \begin{pmatrix} \sigma_{\phi,j} & \sigma_{\phi\psi,j} \\ \sigma_{\phi\psi,j} & \sigma_{\psi,j} \end{pmatrix} \tag{11c}$$

The vector $\vec{\rho}_j^0$ points to the position of the maximum of the *j-th* distribution in the (ϕ,ψ) space and χ_j is its statistical weight, which we use as quantitative measure of a residue's intrinsic propensity for the conformation *j*. The diagonal elements of the matrix $\hat{V}_j$ are the halfwidths at half-maximum of the j^{th} distribution along the coordinates ϕ and ψ and the off-diagonal elements reflect correlations between variations along the two coordinates. If $\hat{V}_j$ is diagonal, the ϕ,ψ projection of the distribution is an ellipse with its main axes parallel to the ϕ and ψ axes.

The expectation value of any observable x (spectral intensities, rotational strengths, J-coupling constants) can be written as:

$$\langle x \rangle = \frac{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} x \cdot f(\phi,\psi)\, d\phi\, d\psi}{Z} \qquad (12)$$

where Z denotes the canonical partition sum.

Schweitzer-Stenner successfully applied this approach to estimate the intrinsic propensities of the central residues of trialanine and trivaline (*86*), and it has now been used in the present study for simulating the amide I' profiles as well as the obtained J-coupling constants. Considering that no stable secondary structures (helices and sheets) can be adopted by tripeptides, we use the classifications of turns instead. The centers of the distributions were defined to be located in the following sub-sections of the Ramachandran plot: (1) PPII (-60° > $\phi_{max,1}$ ≥ -90°; 180° ≥ $\psi_{max,1}$ > 100°), (2) parallel β-

strand (pβ) (-100° > $\phi_{max,2} \geq$ -130°; 140° > $\psi_{max,2} \geq$ 100°), (3) antiparallel β-strand (aβ) (-130^{o} > $\phi_{max,3} \geq$ -180^{o}, 180° > $\psi_{max,3} \geq$ 100°), (4) the transition region between antiparallel β-strand and PPII (aβt) (-90^{o} > $\phi_{max,3} \geq$ -130^{o}, 180° > $\psi_{max,3} \geq$ 140°), (5) dihedral angles of the *i+1* residue in type I β-turns (these are the dihedral angles adopted by residues in right-handed α-helices) (-50° > $\phi_{max,5}$ > -80°; -20° > $\psi_{max,5}$ > -40°), (6) dihedral angles of the *i+1* residue in type I' β-turns (these are the dihedral angles adopted by residues in left-handed α-helices) (80° > $\phi_{max,6}$ > 60°; 40° > $\psi_{max,6}$ > 20°), (7) dihedral angles of the *i+2* residue in type I and II' β-turns (these are the dihedral angles adopted by residues considered in the extended right-handed α-helix distribution) (-70° > $\phi_{max,7}$ > -110°; 20° > $\psi_{max,7}$ > -20°), (8) dihedral angles of the *i+2* residue in type I' and II β-turns (these are the dihedral angles adopted by residues considered in the extended left-handed α-helix distribution) (110° > $\phi_{max,8}$ > 70°; 20° > $\psi_{max,8}$ > -20°) and (9,10) so called γ-turns (classic and inverse, respectively) ($\pm$80° > $\phi_{max,9,10}$ > $\pm$60°; $\mp$50° > $\psi_{max,\ 9,10}$ > $\mp$60°). Other sub-ensembles have been considered based on the results of Makowska *et al.* with dihedral angles of (80° > ϕ_{max} >50°; 170° > ψ_{max} > 40°) (*68*). The 9th and 10th conformations have been considered to account for recent experimental and theoretical evidence that amino acids can populate this part of the (ϕ,ψ) space (*137, 138*). The designations "anti-parallel" and "parallel" β-strand solely indicate that the corresponding region contain the canonical dihedral coordinates found in the corresponding sheet structures.

For each conformation, the corresponding amide I' band profiles were simulated as the sum of two Gaussian profiles assignable to vibrational transitions into two delocalized excitonic states. To account for inaccuracies in peptide concentrations, the experimental

IR and VCD band profiles were scaled according to the dipole strengths obtained by Measey et al. (*124*) The total intensities of the considered ensemble were then calculated by equation 12. The overlap of amide I' with bands assignable to the carbonyl stretching mode of the C-terminal and side chain modes (e.g. for F and Y) was accounted for by fitting the former and the latter with empirical Gaussian and Lorentzian band profiles, respectively.

In a first step, we employed a two-state model encompassing Gibbs energy minima in the PPII and β-strand region of the Ramachandran space. Following Shi et al. (*72*), we obtained the (ϕ,ψ) coordinates of the distribution maxima of the corresponding conformational sub-ensembles from the coil library of Avbelj and Baldwin (*62, 64*). The centers, widths and statistical weights of the two superimposed Gaussian distributions were then adjusted to optimize the agreement between the simulated and experimental amide I' band profiles. In agreement with earlier results (*84, 131, 133*), the shape and the magnitude of the VCD couplet were found to be the most sensitive of all the spectral parameters to variations among the investigated amide I' profiles. In each case, we considered only a single β-strand sub-ensemble associated with one of the three β-strand regions introduced above. Subsequently, the distribution parameters were fine-tuned to minimize the root mean square deviation (RMSD) between the experimentally observed and calculated J-coupling constants, which is written as:

$$RMSD = \sqrt{\sum_{j} \frac{\left(J_j(\exp) - J_j(sim)\right)^2}{J_j(\exp)^2}} \qquad (13)$$

where $J_j(exp)$ and $J_j(sim)$ are the *j-th* experimental and simulated coupling constants, respectively. Generally, this procedure yielded somewhat satisfactory reproductions of the experimental data. In a second step, however, motivated by earlier results for trialanine and trivaline (*86*) and inspection of coil libraries, we attempted a further improvement of our simulations by admixing fractions of turns and additional conformations, which yielded improvements for the fits to the experimental data. Though small in some cases, the admixture of dihedral angle distributions found in various β-turns and γ-turns improved the RMSD of the J-coupling constants, yet did not jeopardize the agreement between the experimental and simulated vibrational spectra.

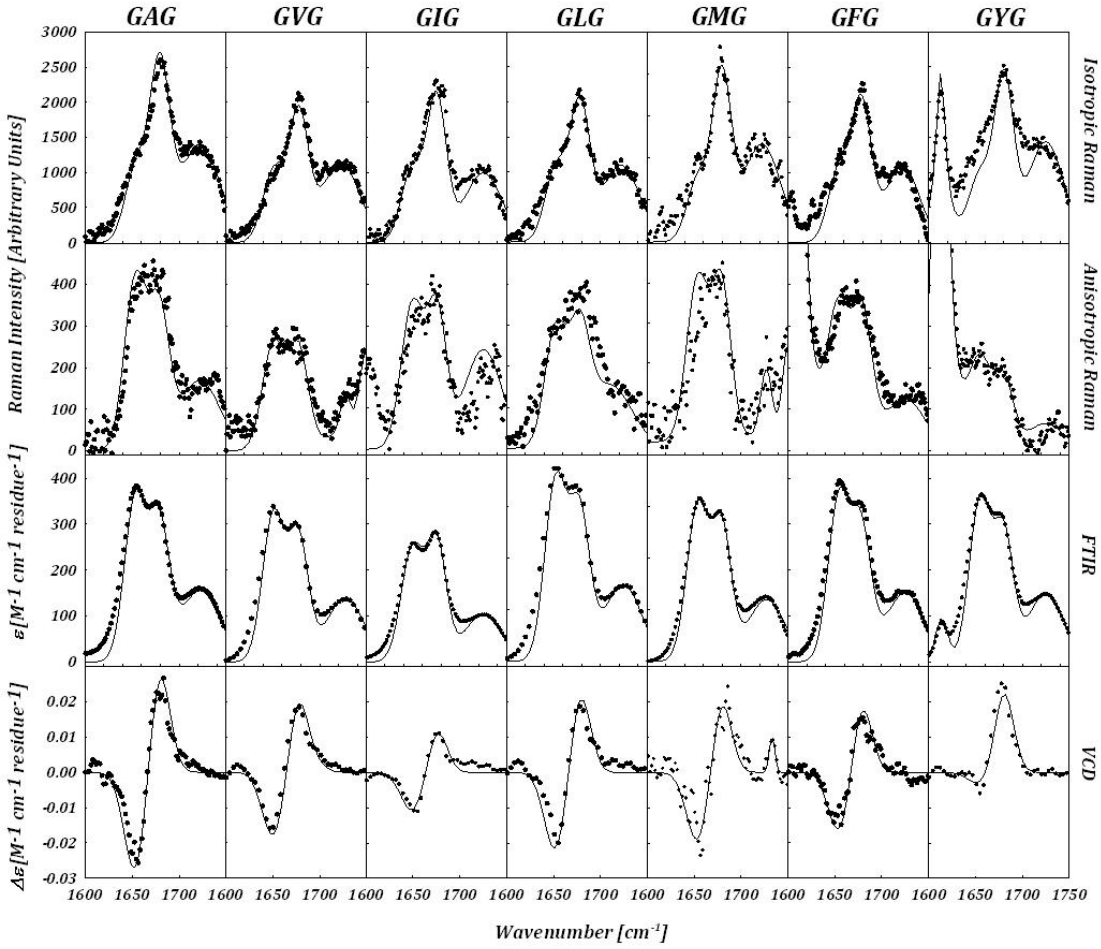
Chapter 5 Results

5.1 Spectral Characteristics

We measured the amide I' band profiles of IR, isotropic Raman, anisotropic Raman, and VCD spectra for a set of GxG peptides in D_2O at acidic pD. D_2O was used to avoid the overlap with the rather strong IR band of water at 1640 cm^{-1} and the vibrational mixing between amide I and H_2O bending modes (*119*). Acidic conditions were selected since GFG is not soluble at concentrations required for the current study and the fact that the $^3J(H^N,H^\alpha)$ constants of some other peptides could not be determined at near neutral pD due to fast H↔D exchange. We chose unblocked, rather than blocked peptides, because the terminal charges increase the difference between the two amide I' bands, and thus the spectral resolution. As shown in our earlier papers, the terminal charges do not significantly affect the conformation of the central residue if the latter is aliphatic or aromatic (*71, 77*). The influence of the terminal charges on charged side chains (i.e. E) will be analyzed below. We also measured the aforementioned set of six J-coupling constants using similar solution conditions.

Figure 12 represents an example of the NMR experimental spectra and Figure 13 and Figure 14 display the full set of vibrational spectra for aliphatic, aromatic, charged and polar residues. Table 2 and Table 3 list the complete set of measured J-coupling constants for the residues corresponding to those presented in Figure 13 and Figure 14, respectively. The regions corresponding to right and left handed α-helices and the extended regions around them have been appropriately termed for that set of dihedral angles' occurrence for the *i+1* and *i+2* residues in various types of β-turns (*139*). It should also be mentioned in this context that the coordinates for PPII and antiparallel β-

strand are similar to those found in various type VI β-turns, but the work of Hutchinson and Thornton (*139*) indicates that the frequency of these is very low so we retain the classification of PPII and aβ in Table 2 and Table 3. For the description of the data analysis below, in some cases we use the term *state* for each Gibbs energy minimum of residue x in the (ϕ,ψ) space. This minimum corresponds to a maximum of a two-dimensional Gaussian distribution function, which describes a sub-ensemble of conformations.



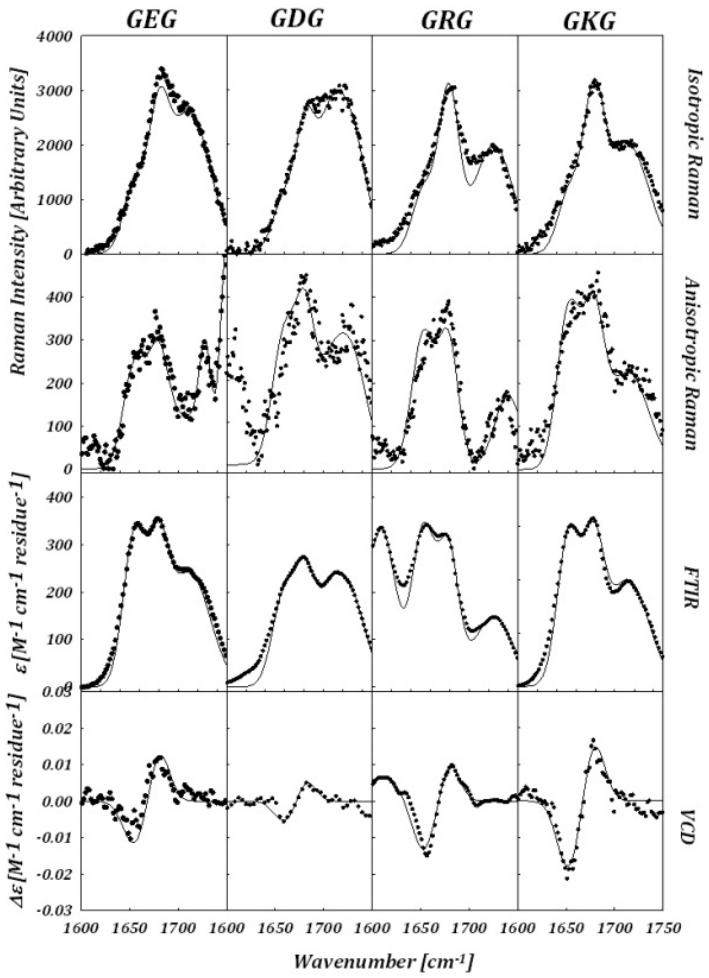


Figure 13: Experimental (dotted line) and simulated (solid line) isotropic (top row) and anisotropic (second row) Raman, IR (third row) and VCD (bottom row) spectra of GAG, GVG, GIG, GLG, GMG, GFG and GYG (top figure) and GEG, GDG, GRG and GKG (bottom figure) (as indicated at the top of each column) from 1600-1750cm^{-1}, encompassing the amide I' region. The fitting parameters and methods of the simulations are described in the text.

	A	V	I	L	M	F	Y
PPII	0.79	0.40	0.42	0.56	0.64	0.42	0.43
-70							
150	-74	-81	-85	-76	-74	-80	-82
	152	170	170	150	160	160	170
Aβt	-	0.38	0.32	0.24	-	0.40	0.32
-110							
150	-	-99	-98	-98	-	-100	-94
		170	120	160		160	170
aβ	0.06	-	-	-	0.36	-	0.04
-140							
140	-115	-	-	-	-120	-	-145
	120				160		75
(*i+1*) residue	0.05	0.04	0.08	0.04	-	0.10	-
Type I β-turn							
-60	-60	-60	-50	-50	-	-70	-
-30	-30	-30	-45	-30		-30	
(*i+1*)residue	-	0.07	0.08	0.10	-	-	0.06
Type I' β-turn							
60	-	60	45	55	-	-	55
30		30	45	30			75
(*i+2*) residue	-	-	0.10	-	-	-	0.10
Type I/II' β-turn							
-90/-80	-	-	-85	-	-	-	-78
0			0				0
(*i+2*) residue	-	-	-	-	-	-	-
Type I'/II β-turn							
90/80	-	-	-	-	-	-	-
0							
$\gamma_{inverse}$	0.05	0.11	-	0.03	-	0.04	-
-80							
70	-80	-80	-	-80	-	-85	-
	60	60		70		50	
$\gamma_{classic}$	0.05	-	-	0.03	-	0.04	0.05
75							
-65	80	-	-	80	-	85	55(15)
	-60			-70		-50	-75(5)
$^3J(H^NH^\alpha)$	6.10	7.48	7.51	6.81	7.10	7.52	7.43
	6.11±0.02	7.46±0.08	7.47±0.01	6.78±0.01	7.08±0.02	7.45±0.02	7.37±0.03
$^3J(H^N,C')$	1.19	1.04	1.06	0.98	-	1.00	1.00
	1.18±0.07	0.91±0.12	0.86±0.18	0.84±0.09	-	0.88±0.15	0.74±0.24
$^3J(H^\alpha,C')$	1.90	2.39	2.45	2.53	-	2.22	2.62
	2.02±0.10	2.33±0.15	2.34±0.08	2.45±0.03		2.20±0.27	2.47±0.09
$^3J(H^N,C^\beta)$	2.09	1.76	1.71	1.91	-	1.84	1.75
	2.32±0.06	1.59±0.06	1.39±0.20	1.75±0.15	-	1.79±0.17	1.37±0.11
$^1J(N,C^\alpha)$	11.28	11.16	10.95	10.97	-	11.18	11.40
	11.28±0.07	11.24±0.02	10.88±0.07	10.96±0.10	-	11.48±0.04	11.26±0.13
$^2J(N,C^\alpha)$	8.35	8.17	7.93	8.14	-	8.23	8.14
	8.51±0.03	8.01±0.02	7.91±0.05	8.24±0.09	-	8.28±0.04	8.08±0.12

	E	D	K	R
PPII **-70** **150**	0.54	0.04	0.50	0.62
	-80 160	-70 150	-66 160	-78 157
Aβt **-110** **150**	0.26	0.28	-	0.10
	-97 150	-100 165	-	-100 135
aβ **-140** **140**	0.04	0.14	0.41	0.06
	-140 165	-135 170	-115 160	-140 145
(*i+1*) residue **Type I β-turn** **-60** **-30**	0.08	-	0.09	0.08
	-50 -10	-	-65 -30	-65) -50
(*i+1*) residue **Type I' β-turn** **60** **30**	-	-	-	0.14
	-	-	-	80 50
(*i+2*) residue **Type I/II' β-turn** **-90/-80** **0**	-	0.28	-	-
	-	-70 0	-	-
(*i+2*) residue **Type I'/II β-turn** **90/80** **0**	-	-	-	-
	-	-	-	-
$\gamma_{inverse}$ **-80** **70**	0.04	-	-	-
	-75 60	-	-	-
$\gamma_{classic}$ **75** **-65**	0.04	-	-	-
	70 -60	-	-	-
Upper **right** **quadrant**	-	0.20	-	-
	-	55 170	-	-
Lower **right** **quadrant**	-	0.06	-	-
	-	55 -170	-	-
$^3J(H^NH^\alpha)$	7.03	7.48	6.63	6.69
	6.99±0.02	7.44±0.02	6.60±0.02	6.66±0.01
$^3J(H^N,C')$	1.13	1.25	-	1.14
	0.94±0.11	1.19±0.09	-	1.00±0.08
$^3J(H^\alpha,C')$	2.12	3.34	-	2.59
	2.07±0.16	3.24±0.24	-	2.47±0.05
$^3J(H^N,C^\beta)$	1.88	1.41	-	1.89
	1.59±0.24	1.25±0.28	-	1.79±0.05
$^1J(N,C^\alpha)$	11.00	11.61	-	11.03
	11.24±0.03	11.89±0.05	-	11.02±0.13
$^2J(N,C^\alpha)$	8.24	7.95	-	8.14
	8.38±0.04	7.91±0.04	-	8.11±0.10

Table 2: The parameters used for the simulations of the vibrational spectra and reproduction of the J coupling constants for GXG peptides where X=A, V, I, L, M, F and Y (top) and X=E, D, K and R (bottom). These parameters include the mole fractions of the considered sub-ensembles as well as the centers of their (ϕ,ψ) distributions, listed in the upper and lower part of split cells, respectively in the top half of the table. Simulated (upper sub-cells) and experimental (lower sub-cells) NMR J coupling constants are listed in the bottom half of the table.

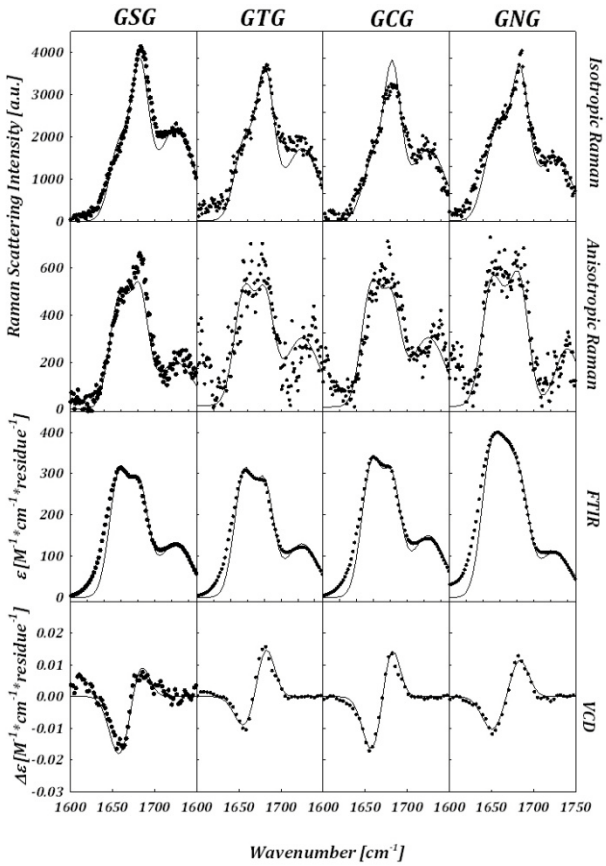


Figure 14: Experimental (dotted line) and simulated (solid line) isotropic (top row) and anisotropic (second row) Raman, IR (third row) and VCD (bottom row) spectra of GSG, GTG, GCG, and GNG (as indicated at the top of each column) from 1600-1750cm^{-1}, encompassing the amide I' region. The fitting parameters and methods of the simulations are described in the text.

	S	T	C	N
PPII **-70** **150**	0.45	0.35	0.51	0.46
	-79 160	-89 170	-84 170	-83 166
Aβt **-110** **150**	0.30	0.35	0.20	0.18
	-103 160	-103 170	-99 170	-110 166
aβ **-140** **140**	-	-	-	0.12
	-	-	-	-145 120
(*i+1*) residue **Type I β-turn**	-	-	-	-
	-	-	-	-
(*i+1*) residue **Type I' β-turn**	-	-	-	-
	-	-	-	-
(*i+2*) residue **Type I/II' β-turn** **-90/-80** **0**	0.10	0.15	0.12	0.04
	-50 0	-60 0	-65 0	-60 0
(*i+2*) residue **Type I'/II β-turn** **90/80** **0**	0.15	0.15	0.07	-
	70 0	60 0	60 0	-
γ_1	-	-	-	-
	-	-	-	-
γ_2	-	-	-	-
	-	-	-	-
Upper right quadrant	-	-	0.10	0.20
	-	-	60 160	75 75
Lower right quadrant	-	-	-	-
	-	-	-	-
$^3J(H^N H^\alpha)$	7.01	7.72	7.29	7.54
	6.99±0.07	7.73±0.02	7.29±0.01	7.53±0.01
$^3J(H^N,C')$	1.13	0.98	0.86	1.12
	0.87±0.18	0.84±0.08	0.87±0.08	0.99±0.09
$^3J(H^\alpha,C')$	2.71	2.91	2.85	3.02
	2.77±0.32	2.74±0.05	2.79±0.03	2.88±0.10
$^3J(H^N,C^\beta)$	1.76	1.62	1.80	1.65
	1.71±0.12	1.40±0.12	1.89±0.08	1.39±0.07
$^1J(N,C^\alpha)$	11.04	11.53	11.73	11.24
	11.73±0.05	11.63±0.04	11.81±0.04	11.21±0.09
$^2J(N,C^\alpha)$	7.89	7.78	8.12	8.29
	7.86±0.06	7.69±0.02	8.04±0.01	8.28±0.09

Table 3: The parameters used for the simulations of the vibrational spectra and reproduction of the J coupling constants for GXG peptides where X=S, T, C and N. These parameters include the mole fractions of the considered sub-ensembles as well as the centers of their (ϕ,ψ) distributions, listed in the upper and lower part of split cells, respectively in the top half of the table. Simulated (upper sub-cells) and experimental (lower sub-cells) NMR J coupling constants are listed in the bottom half of the table.

The amide I' profiles shown in Figure 13 and Figure 14 look similar with respect to their intensity distributions, but some differences are noteworthy. Most of the investigated residues exhibit a non-coincidence between the isotropic Raman and IR profiles where the high frequency band in the former is more intense than the latter and *vice versa*. This non-coincidence is indicative of a dominant sampling of extended conformations associated with the upper left quadrant of the Ramachandran plot (*122*). However, there are some notable differences. In particular, the isotropic Raman and IR band profiles for GDG exhibit a coincidence with no redistribution of intensity. The same is seen, although to a somewhat lesser extent, for GIG. This coincidence is typical for the central residue sampling regions of the Ramachandran space outside of the extended region. The negative VCD couplets are all indicative of a significant PPII fraction being populated by the respective residues, but the amplitude of these couplets are clearly different (*140*), thus suggesting different conformational distributions. The nearly non-existent VCD couplet for GDG is an exception to this, underscoring the previous notion that D predominantly samples structures other than PPII.

The CD spectrum for PPII is characteristically distinct from that of other secondary structures such as α-helices and β-sheets (strands). The CD signal at ~215nm is positive and at ~195nm is negative for PPII, which are both opposite in that regard from what is observed for α-helices and β-sheets (strands). The room temperature UVCD spectra shown in Figure 15 are all indicative of a substantial sampling of PPII conformations. However, some significant differences are seen in the spectra for GSG, GCG, GTG, GNG and GDG. Particularly, the positive maximum which is typical for PPII signals at 215nm

becomes a minimum for these peptides. Considering the relative decrease in magnitude of the negative maximum at 195nm as well in these spectra, the overall differences indicate the propensity for these residues to sample additional conformations.

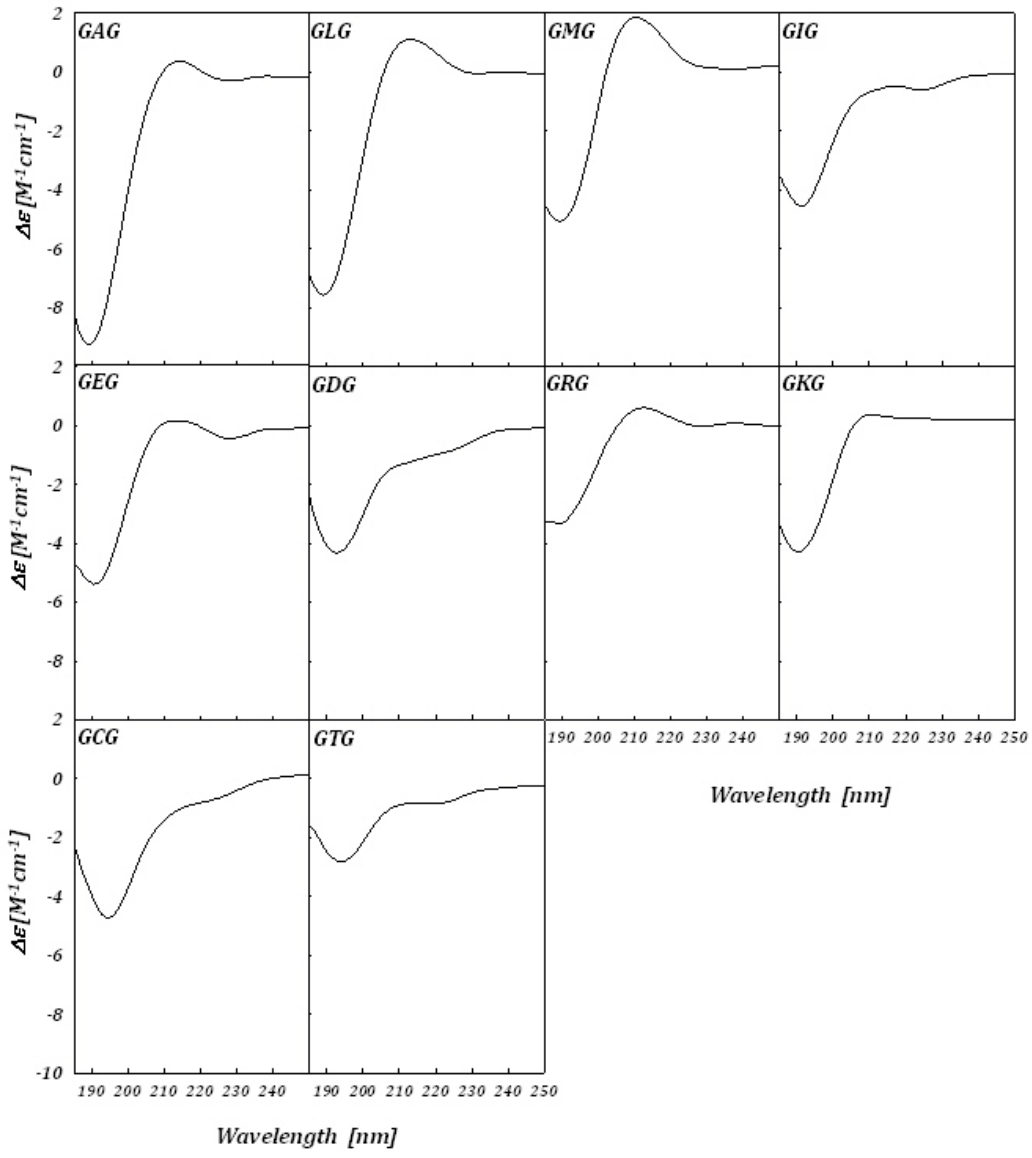


Figure 15: Room temperature UVCD spectra of GAG, GLG, GMG, GIG, GEG, GDG, GRG, GKG, GCG and GTG taken from 185-250nm.

5.2 Aliphatic, Aromatic and Charged Residues

Our analysis revealed a PPII fraction of 0.79 for alanine, which is only slightly lower than the PPII propensity obtained for the central residue of trialanine (0.84) (*86*) using the same method of analysis, thus suggesting that the energy of nearest neighbor interactions between alanine residues is small, in agreement with Chen et al. (*141*). The remaining fraction of the GAG ensemble is almost evenly distributed over Ramachandran plot regions associated with aβ and turn conformations (Table 2), in agreement with findings for trialanine (*86*). The total distribution function is plotted in Figure 16. As one can infer from Table 2, the set of J-coupling constants obtained from simulations are sufficiently close to the corresponding experimental values.

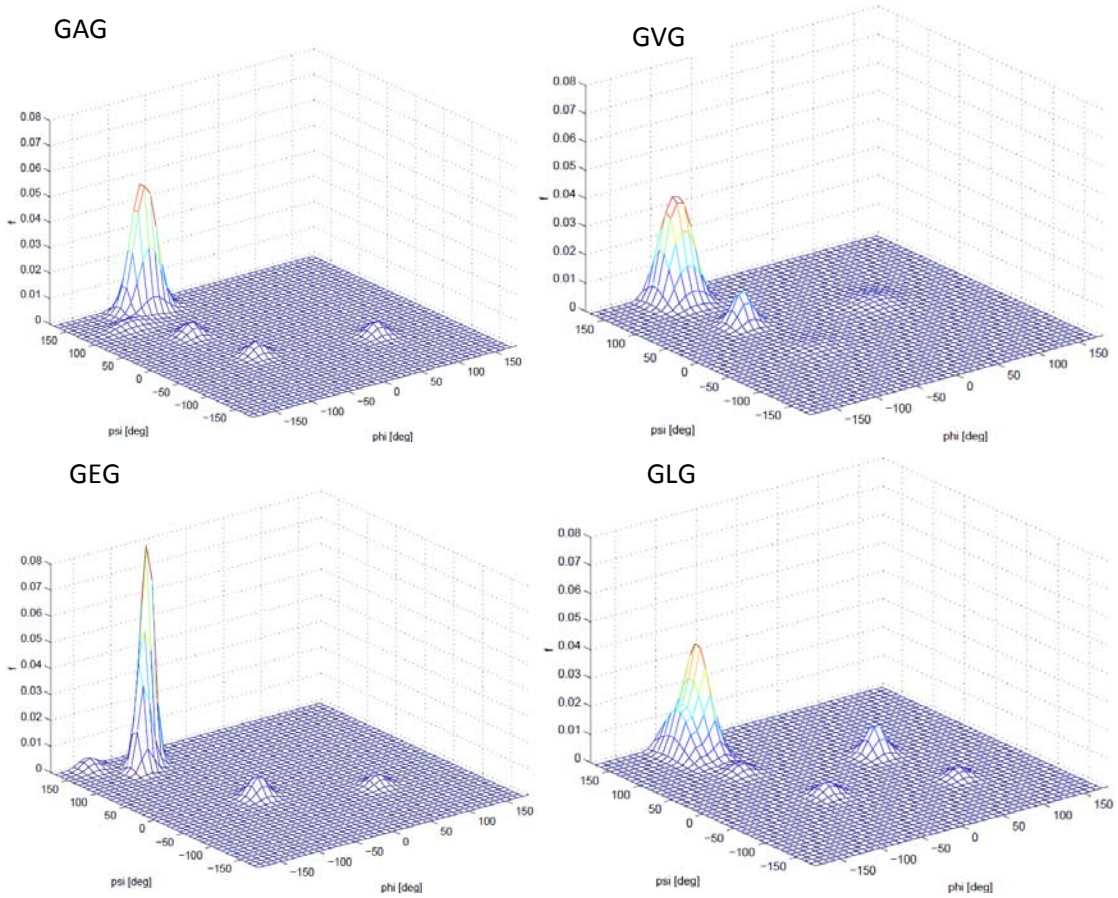


Figure 16: Taken from ref (*48*) and modified. Conformational distributions of the central residue in GAG (top left), GVG (top right), GLG (bottom right) and GEG (bottom left) obtained from the analysis of amide I' profiles and J-coupling constants as described in the text.

The VCD (Figure 13) and the ϕ-dependent $^{3}J(H^{N},H^{\alpha})$ coupling constants of GVG and GFG (Table 2) indicate that the PPII fractions of both guest residues are substantially lower than that of alanine. A comparison of the GAG and GVG SRCD spectra in Figure 17 points in the same direction.

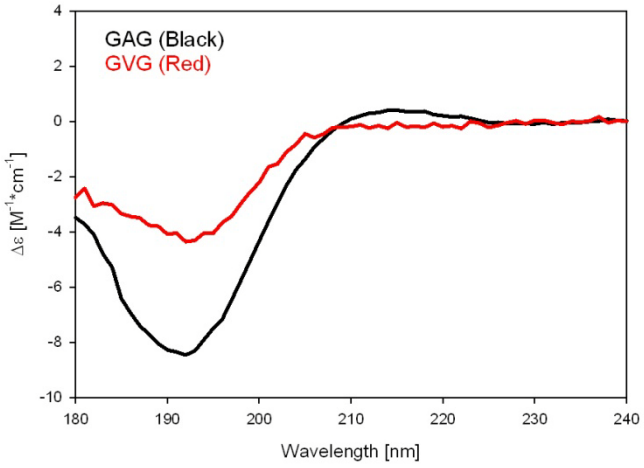


Figure 17: Taken from ref (*48*) and modified. Room temperature synchrotron radiation circular dichroism (SRCD) spectra of GAG (black line) and GVG (red line).

The PPII conformation yields a couplet in the UVCD spectra with a characteristic strong negative component at ~198 nm and weaker positive component at ~215 nm. The reduced intensity of the couplet for GVG at both of these wavelengths is indicative of a lower PPII propensity. It was recently shown that the intensity of these components can be directly related to PPII content (*42*). The conformational ensembles of both amino acid residues sample the aβt-region (0.38 and 0.40 for V and F, respectively) and PPII region (0.40 and 0.42 for V and F, respectively) evenly. The distributions of valine and phenylalanine exhibit small propensities turn conformations (Table 2). The total distribution function of GVG is plotted in Figure 16 for the purpose of illustration. A recent computational study by Xu et al. using a Hamiltonian replica exchange molecular

dynamics (HREMD) approach corroborates the notion that valine and phenylalanine have an increased preference for β-like structures relative to alanine (*142*).

Interestingly, the β-strand propensity of V is rather different in GVG and VVV. The latter exhibits a substantially higher β-like propensity (0.68) than the former, a lower value for PPII (0.16) and a higher value for the right-handed helical state (0.16) (*86*). Moreover, the distribution of the β-strand conformation found by Schweitzer-Stenner is located in aβ-region ((ϕ,ψ) = (-130°, 135°)). This discrepancy is somewhat less pronounced, though still significant if one compares GVG with the trivaline (VVV) analysis of Graf et al., who obtained fractions of 0.29, 0.52, and 0.19 for PPII, β-strand and right-handed helix, respectively (*85*). It can, therefore, be concluded that adjacent valine residues in a polypeptide chain shift their mutual propensities towards β-strand, increase the propensity for sampling the right-handed helix region and decrease the PPII propensity. Based on our propensity values, the enhancement of the β-strand propensity of the central valine residue in trivaline would correspond to a Gibbs energy value of 1.2 kJ/mol, at room temperature.

Analogous to Val and Phe are Ile and Tyr. Ile is more branched than valine and thus we expect for the β-content to be increased. This is, in fact, not the case. The PPII content for V (0.40) and I (0.42) is similar. Both residues sample the aβt region, although the sampling of this region by isoleucine is somewhat less compared to valine (Table 2). The propensity to sample regions associated with the *i+1* residue in type I and I' β-turns is similar within the uncertainty of the analysis. The main difference occurs in the sampling of turn structures. Ile samples a region centered at (ϕ,ψ) = (-85°, 0°) with a

mole fraction of 0.10. This region is consistent with the dihedral angles found in the *i+2* residue of type I and II' β-turns. Tyr differs from Phe in that a hydrogen on the aromatic side chain ring is substituted by an OH group, thus making it somewhat polar as well. Our analysis indicates that Y samples the PPII and aβt-region similarly to that of F. Both residues sample γ-turns (classic and inverse) similarly as well. Like isoleucine, tyrosine samples the dihedral angles of the *i+2* residue in type I and II' beta turns with a mole fraction of 0.10.

Leucine has a rather long aliphatic side chain, which is less sterically demanding, concerning the backbone hydration, than that of valine or isoleucine. Earlier results obtained for the short peptides ALA and AL indicated a comparable population of the PPII and β-strand region (*71, 106*). Our analysis of GLG revealed a slightly different picture. Indeed, we found a propensity of 0.56 for PPII, in agreement with Eker et al. (*71*). However, a substantial fraction (24%) of the ensemble is found in the aβt-region with the center of the corresponding distribution located at ((ϕ,ψ) = (-98°, 160°) (c.f. Table 2). The corresponding total distribution function therefore suggests that this residue significantly samples the region between the canonical PPII and β-strand conformations, which is in good agreement with recent theoretical predictions (*73*).

Charged amino acid residues such as glutamic acid have attracted interest because of their abundance in intrinsically disordered proteins and their alleged propensity for the PPII conformation (*43, 143*). The values for the fractions obtained from our data set for GEG are, in principle, consistent with this notion, but a PPII fraction of 0.54 indicates that glutamic acid is substantially less inclined to sample the PPII trough than alanine.

Sampling of the aβt-region is substantial (0.26) with a small fraction of antiparallel β-strand. The turn fractions are again small (Table 2). The total distribution function is shown in Figure 16. The obtained χ-values of GEG reflect the propensity of the protonated residue, which does not resemble its state at physiological pH. Unfortunately, fast H↔D exchange at near neutral pH prevents a measurement of the $^3J(H^N,H^\alpha)$ coupling constant. To check whether there is any influence of the residue charge on the propensity of glutamic acid, we additionally measured and analyzed a full set of amide I' profiles of GEG at pD = 5.1. We achieved a satisfactory simulation of the amide I' profiles (data shown in SI ref. (*48*)) by using the exact fractions, half-widths and centers of distributions obtained for cationic GEG. This reproducibility indicates that the conformational propensity of glutamic acid does not significantly depend on its protonation state. Generally, this finding underscores the notion that the terminal charges have a limited influence on the conformation of the central residue of tripeptides, most likely because the water molecules in the hydration shell attenuate electrostatic interactions (*103*).

An acidic residue similar to glutamic acid, only with a shorter side chain (by one CH_2) is aspartic acid. One would naively assume that this residue would behave somewhat similarly to Glu due to the acidic nature of the side chain. A first inspection of the experimental results indicates that this is not the case. The NMR coupling constants vary between the two residues in every instance. Also, the vibrational spectra, especially the magnitude of the VCD signal for asp, indicate large differences in the sampling ensembles of these two acidic residues. This is indeed what our analysis

yielded. The PPII sampling of D is minimal (0.04). The aβt and region associated with the *i+2* residue in type I and II' β-turns are populated the most (0.28 each). Of all investigated residues the aβ region is the most populated (0.14) for Asp. Interestingly, this residue significantly populates a region in the upper and lower right side of the Ramachandran plot with the distributions centered at (ϕ,ψ) = (55°, ±170°). The populations of these regions are 0.20 and 0.06, respectively. The necessity to consider sampling from these regions comes from the fact that there is a nearly nonexistent VCD signal, yet the ψ-value coupling constants indicate populations of sub-ensembles with large (in magnitude) ψ backbone angles. In these regions our simulations produce an intense VCD couplet with an opposite signal than PPII. Also, the $^{3}J(H^{\alpha},C')$ coupling constant is larger than that of the rest of the investigated residues. This condition is satisfied with sampling of regions with positive ϕ-angles. A comparison with asparagine, which is similar to aspartic acid, will be detailed below in the polar residues section.

A first observation of the vibrational and NMR data indicates that arginine predominantly samples the PPII region. This is indeed the case. This residue has a PPII propensity of 62%, which is one of the highest for the investigated residues, yet is still significantly less than alanine. The remainder of the ensemble is populated by aβt (0.10), pβ (0.06) and the regions associated with the *i+1* residue in type I (0.08) and type I' (0.14) β-turns.

The above findings indicate that a residue like lysine shows a similar behavior as glutamic acid or arginine, irrespective of the character of the terminal group of the side

chain. This hypothesis prompted us to explore the conformational ensemble of GKG, based on the analysis of its amide I' profiles and $^{3}J(H^{N},H^{\alpha})$ coupling constants. We added GMG to check whether a modification of the hydrophobic chain (S instead of CH_2) makes any substantial difference. The corresponding vibrational spectra are shown in Figure 13, with the results of our simulations in Table 2. We determined, indeed, that lysine in GKG (Table 2) behaves like glutamic acid, in that it exhibits a PPII fraction of 0.50, and an aβt-fraction of 0.41. Interestingly, methionine has a somewhat higher PPII propensity with a fraction of 0.64, which, together with its aβt-region fraction of 0.36 makes it somewhat comparable with glutamic acid, though the positions of the respective β-strand distributions lie in different sub-sections of the Ramachandran plot.

5.3 Polar Residues

Serine and threonine are particularly interesting to investigate because the side chains are similar to that of alanine and valine, respectively, except for the replacement of a methyl hydrogen in the former and a methyl group in the latter to a hydroxyl group. However, contrary to alanine and valine (*106*), these amino acids are polar residues assumed to adopt mostly a β conformation, as has been reported for threonine (*73*). Our results suggest a substantial sampling of the aβt-region (0.30), but the propensity for PPII (0.45) is dominant for serine. Threonine samples these regions equally (0.35 each), however, the PPII distribution is shifted toward β-structures by 10° compared to serine. Interestingly, the remainder of the ensemble samples the *i+2* residue dihedral angles of type I and II' (0.10 for serine and 0.15 for threonine) and type I' and II (0.15 for both) β-turns. Cysteine is a similar residue to serine except that the hydroxyl side chain

of serine is replaced by a thiol. Considering, the somewhat minimal difference between these two residues, the sampling of cysteine should match that of serine. This happens to not strictly be the case. Although the NMR coupling constants for these two residues are very similar, the VCD couplet for GCG is more intense than that of GSG by a factor of ~1.5. Our analysis yields a dominant PPII sampling for cysteine (0.51). The aβt region makes up 20% of the conformational ensemble for this residue. The *i+2* residue dihedral angles of type I and II' (0.12) and type I' and II (0.07) β-turns are also sampled by this polar residue, similar to that of other polar residues, Ser and Thr. Interestingly, there is also an appreciable sampling (0.10) of a distribution centered at (ϕ,ψ) = (60°, 160°) for cysteine.

A polar residue, which is sterically similar to protonated aspartic acid, is asparagine, except the hydroxyl group on the side chain of the former is replaced by a NH_2 group on the latter. Apparently, this substitution induces considerable changes in the sampling of Asn. The somewhat intense VCD couplet indicates a predominant sampling for PPII. Our analysis reveals just that, with a mole fraction of 0.46 in the PPII region. The aβt (0.18) and aβ (0.12) -regions are moderately sampled by this residue. A small sampling of the *i+2* residue dihedral angles in type I/I' and II/II' β-turns occurs also, though not to the extent of the other polar residues. A region outside of these distributions with dihedral angles of (ϕ,ψ) = (75°, 75°) is significantly (0.20) populated as well.

Chapter 6 Discussion

6.1 Side Chain Correlation to Propensities

Our investigation of the conformational ensemble sampled by a set of amino acid residues representing aliphatic (A), branched aliphatic (V and I), extended aliphatic (L and M), aromatic (F and Y), polar (S, T, C and N), acidic (E and D) and basic (R and K) side chains yields the following picture. The preference of residues with aliphatic side chains for the PPII conformational state is highest for alanine, which has the shortest and least sterically demanding side chain in its class. Aliphatic and sterically demanding residues (V, I and L) significantly populate the aβt-region in addition to the PPII region. Leucine has a clear preference for PPII, whereas valine and isoleucine have more of an equal preference for both sub-ensembles. Interestingly protonated glutamic acid behaves like an aliphatic residue such as leucine. The fact that the propensity of glutamic acid does not depend on its protonation state suggests that the hydrophilic carboxylate group does not matter significantly in this regard. Similarly, protonated arginine behaves like an aliphatic residue in that it preferentially samples the PPII region. Aspartic acid exhibits a peculiar ensemble of sampling with most of its propensity located in the aβt, aβ, type I and II' (*i+2* residue) β-turn as well as a region in the upper right quadrant of the Ramachandran space. Interestingly, residues with hydroxyl (polar) groups on their side chains behaved somewhat similarly. These residues showed a low propensity for PPII and β-like structures, however, the propensity for these residues to populate the distributions for the dihedral angles of type I/I' and type II/II' (*i+2* residue) β-turns. This effect was similar for cysteine, whose thiol capped side chain is comparable to that of

serine, although the PPII propensity was stabilized at the expense of the aβt-propensity. In a more general context, the sampling of the β-strand region is mostly confined to the aβ and aβt regions. A comparison of GVG and VVV suggests that only nearest neighbor interactions can shift the β propensity to the aβ-region (*85, 86*).

The β-like propensities of V and F, and their possible modulation by nearest neighbors, is important for understanding the role of V_n and F_n segments in proteins as well as in synthetic and natural peptides. It deserves to be mentioned that many lines of experimental evidence suggest that β-propensity of specific residues, i.e. V and F, is relevant under certain circumstances where aggregation occurs (*144-146*). We have to emphasize, however, that recent investigations by the Hecht group cast some doubt on the notion that propensities of amino acid residues matter for the self-aggregation of peptides (*147, 148*). Thus, one has again to differentiate between the propensities of amino acid residues before and after the nucleation process. The relationship between both is unclear for β-sheet formations.

The parameters determining the propensities of amino acid residues in unfolded peptides are still a matter of debate. With respect to the stabilization of the PPII conformation, an early proposal was made based on DFT calculations for an alanine dipeptide, which suggested that a bridge comprised of two hydrogen bonded water molecules connecting the CO and NH groups of neighboring peptide units stabilizes PPII conformations (*49*). Drozdov et al. argue to the contrary (*103*). They analyzed the hydration structure and thermodynamics of different low-energy conformers of the model alanine dipeptide using Monte Carlo simulations. The authors concluded that

solvation does not solely stabilize PPII over helical conformations, rather its effect was found to be indirect, neutralizing electrostatic interactions between non-bonded pairs of atoms. Steric interactions favoring PPII therefore become dominant. These results and those of Garcia led to the conclusion that PPII structures allow for an optimal hydration of the peptides (*94, 103*). It is therefore expected that an increasingly bulky side chain will perturb the optimal hydration and thus destabilize PPII propensity. This is what we observe for residues V, I, L and M. The fact that M populates PPII the most out of these aliphatic, bulky residues leads to the notion that not only the chemical nature, but the length of the side chain plays a role in modulating the solvation shell and thus the conformational sampling of amino acids. Also, F and Y, with rather bulky, sterically demanding side chains, fit consistently into this picture, although the polar group on the Y side chain modulates the solvation shell differently than that of F. In this context, the results of Drozdov et al. indicate that the equilibrium between PPII and β is, in fact, directly affected by solvation. The charged residues E, R and K correlate with these notions as well. The side chains of these residues are rather bulky, but extend far enough from the backbone so as to not significantly disrupt the solvation of the backbone.

By shortening the side chain of E, i.e. that of D, the solvation of the backbone is significantly modulated. This can be seen in Table 2. The fact that D and N are reported to form unique turns which are stabilized by adjacent seven membered rings, such as those in the N-glycosylation sequence in proteins, termed Asx-turns (*149*), further indicates that the length of the side chain can induce stable structures other than those

generally considered for unfolded peptides. Asp and Asn behave similarly according to length of their side chain. The presence of γ-carbonyl group on their side chains allows for these residues to form the so called Asx turns which are ten membered rings stabilized by a hydrogen bond by the side chain carbonyl oxygen to the i+1 amide proton. This side chain-residue turn is analogous to the residue-residue interaction in certain β-turns, and thus Duddy et al. hypothesized an Asx and β-turn mimicry (*150*). Our results offer limited information into the sampling of various Asx turns for Asp and Asn. Nevertheless, a comparison warrants mentioning. Our ψ-angle distributions can be related to the angle ψ_e in these turns, where $\psi_e = \psi + 120°$. The majority of our ψ-angle distribution is centered on 0° and 170°. These would correspond to 120° and -110° in Asx-turns. Interestingly, a large fraction of the distribution for the ψ_e angle in these turns is located at these values. Additional information could be gained by estimating the distributions of the χ_1 rotamer angles, which relate to the ϕ_e angles in these turns.

The residues that do not fit into the above categories of steric bulk are S, C and to a lesser extent T, whose composition is close to some of the aliphatic residues but for the replacement of a β-hydrogen or methyl group by a hydroxyl or thiol group. However, a hydroxyl or thiol group can act as an acceptor for hydrogen bonding itself, which, in turn, can disrupt the solvent (H_2O) network, which stabilizes PPII as discussed before. In addition to destabilizing PPII, regions according to the dihedral angles adopted by the i+2 residue in types I/I' and II/II' β-turns. Also, recent investigations of the PDB reveal that the side chains of these polar residues (S and T) are involved in a unique type of

turn, termed an ST-turn (*151*). These turns involve a side chain-backbone interaction that is stabilized by an OH of the ser/thr side chain hydrogen bonded to the amide proton of the i+1 residue creating a somewhat stable nine member ring. It is unclear whether cysteine exhibits this same effect in stabilizing side chain-residue turn structures, although we speculate that it would. Concerning the ST-turn population, our results offer limited information. The relationship between the ψ backbone angle and the ψ angle found in these turns (ψ_e) is $\psi_e \approx \psi + 120°$. The ψ-distributions for these residues occur at 170° and 0°, which would correspond to ψ_e values of -110° and 120°. These values match with the distributions for type II' and II ST-turns, respectively. Additional information could be gained by estimating the preference of the χ_1 rotamer angles, which would relate to the ϕ_e angle of these turns.

6.2 Comparison to Coil Libraries

Coil libraries are generally considered as suitable tools to explore the conformational propensities of amino acid residues in unfolded peptides and proteins (*61, 152-154*). They are used, e.g., to calculate expectation values for J-coupling constants which can then be used with experimental values of residues in unfolded peptides and proteins to identify deviations from a statistical coil behavior (*155, 156*). However, owing to the existence of nearest neighbor and second neighbor interactions, the thus obtained results do not necessarily reflect the intrinsic propensities of amino acid residues, as they were inferred from the present study, because the values extracted from coil libraries reflect the fact that nucleation has already occurred. A comparison to these values is useful nevertheless to compare the extent of the nucleation process in the

regions that are supposedly representative of the unfolded state. The most noticeable difference between the intrinsic distributions inferred from our data and coil libraries is the fact that the latter exhibit much larger fractions of helical conformations (*66, 152*). Even the most restricted data set of Jha et al., for which helices, sheets, turns and pre-proline residues were omitted, shows helical fractions varying between ca. 0.18 for valine and nearly 0.4 for aspartic acid and threonine (if glycine and proline are neglected). We must therefore conclude that this effective helical propensity is induced by nearest neighbor or even non-nearest neighbor interactions. With respect to PPII, the values obtained from the restricted coil library, referred to as $c_{\alpha\beta t}^{intern}$ in Jha et al., are well below (by 0.10-0.20) the propensities obtained from our study of A, R, S, C, T, K, E, and M, whereas both agree rather well for V, I, F, Y, N and L. Concerning the β propensities, the values obtained from the restricted coil library lie well below our values for E and K, while the respective values for V, L, F, Y, T, S, D and M are comparable. For A, the value obtained from the coil library (0.25) is much higher than our value (0.06) obtained for the β-strand propensity. The same can be said for R, N, C and I. This comparison seems to suggest that the propensities of residues with long and bulky side chains are less affected by their respective contexts in a peptide or protein than residues with short (A), polar (S, T) or charged (E, K) side chains, which is a somewhat surprising observation since bulky residues are considered as more effective in modulating the conformational propensities of their neighbors (*61*). It is noteworthy that the centers of our PPII distributions are generally rather close to those obtained from coil libraries. Most of our β-like distributions deviate from the coordinates found in

coil libraries. These differences can be viewed, e.g. in the β distributions of V, F and L, which were found to be located in the aβt-region (ϕ,ψ) = (-99°, 170°), (-100°, 160°) and (-98°, 160°), respectively. These coordinates are rather different from the canonical values obtained in parallel and antiparallel β-sheets. For the restricted coil library, the centers of the β distributions for V and F are also not represented by canonical β-sheet values, though the differences are not as pronounced as they are for our values.

6.3 Comparison to Theoretical Studies

Computational studies of conformational propensities have thus far focused mostly on alanine. The work of Tran et al. (*73*) is an exception in this regard, in that these authors used Monte-Carlo simulations to study the propensity of all 20 amino acids and their context dependencies. Here we focus on the results that they obtained for the residues investigated in the present study in a glycine rich context. Figure 18 depicts a comparison of PPII propensities discussed here with that reported by Tran et al. It should be noted that in Figure 18, the values in the histogram of Tran et al. include the fractions found in the sub-ensembles of PII and P_{hyp}, which is defined by the region (-100° < ϕ < -60°, 100° < ψ < 140°). As seen in Figure 18, our results correlate well with those of Tran et al., except for alanine, leucine, arginine, threonine and aspartic acid. Interestingly, these residues are close to the maximum or minimum of our PPII propensity hierarchy. The results for residues with intermediate PPII propensity agree well with Tran et al. in this context. With respect to PPII, they obtained very similar propensity values between 0.4 and 0.5 (these numbers are the sum of what Tran et al. called P_{II} and P_{hyp}). Leucine again departs from this pattern in that the authors found its distribution dominated by a

conformational sub-ensemble located in the transition region between PPII and the antiparallel β-strand coordinate. As indicated above, the respective distribution we obtained for GLG attributes 24% of the total ensemble to a region, which heavily overlaps with the $P_{II}/P_{hyp} \Leftrightarrow \beta_{A/P}$ transition region of Tran et al. Interestingly, Tran et al. also obtained a substantial population for V and I in this region, which also agrees with our results. If we consider our extended helical region propensities, there is an agreement with the helical propensities reported for the simulations of L, C, T and D by Tran et al. The rest of the considered residues have overestimated helical propensities, mostly at the expense of the β-strand content, which is underestimated.

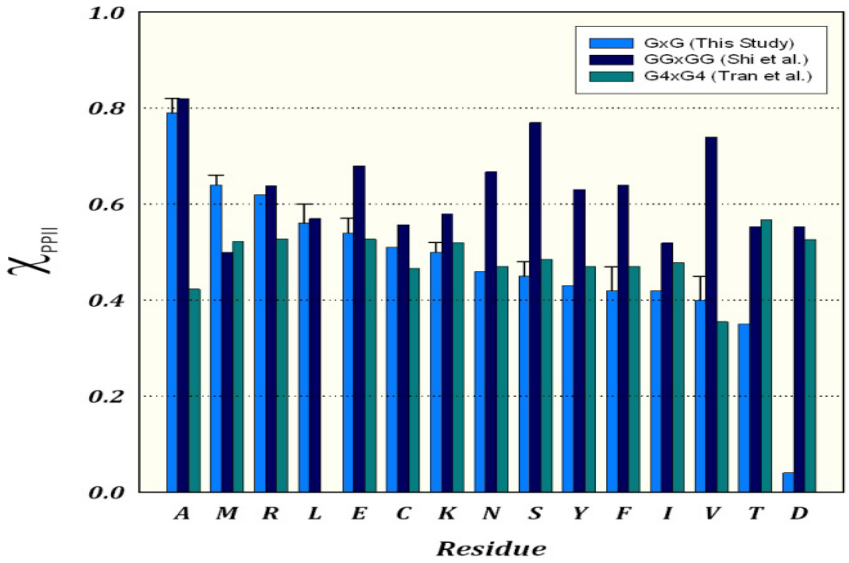


Figure 18: Histograms comparing results for the PPII propensity (χ_{PPII}) of individual amino acids in GxG (this study) (left side: light blue), GGxGG (Shi et al. (*72*)) (middle: dark blue) and G_4xG_4 (Tran et al. (*73*)) (right side: green). Amino acids are indicated on the x-axis by one letter codes.

More recent work of the Pappu group is noteworthy in this context. Vitalis and Pappu employed a new model for implicit solvation together with various force fields to explore the conformational manifold sampled by Ac-x-Nme peptides in water (x again represents different amino acids) (*157*). They judged the validity of their simulations by

comparing them with the respective $^{3}J(H^{N},H^{\alpha})$ coupling constants of x residues. Generally, they found a good agreement between simulation and experiment for OPLS-AA and Amber force fields, but not for alanine for which the $^{3}J(H^{N},H^{\alpha})$ was constantly overestimated. This underscores that the peculiarity of alanine is still not fully understood.

6.4 Comparison to Experimental Studies

Kallenbach and coworkers reported a comprehensive list of experimentally determined intrinsic amino acid propensities for all 20 naturally occurring amino acids (*158*). Their values were obtained by analyzing the corresponding $^{3}J(H^{N},H^{\alpha})$ constants of the guest residues in Ac-GGXGG-NH_2 in terms of a mixture of two different (representative) conformations of the PPII and β-strand sub-ensembles. Figure 18 shows a histogram comparing our estimated PPII propensities to those obtained by Shi et al. for the investigated amino acids. Our results agree well with the residues we determined to have a PPII propensity larger than 50% (A, M, R, L, E, C and K). The most significant discrepancies were obtained for S, V and D. These discrepancies are partially due to differences between the respective $^{3}J(H^{N},H^{\alpha})$ coupling constants, in that our values are much higher (by >0.4) for these three residues. The higher coupling constants translate to a larger average negative ϕ angle, corresponding to a larger fraction of β conformations. As a consequence, our PPII propensities are much lower, particularly for V and S, than those reported by Shi et al. (i.e. 0.74 and 0.77, respectively) (*158*). The differences between the data provided by Kallenbach and our data may also arise, In part, from differences in measuring the $^{3}J(H^{N},H^{\alpha})$ coupling constants. Also, the fact that

authors chose only two representative conformations to analyze the data is problematic. Other conformations with ϕ-angles that correspond to coupling constants in the range of the values reported by Shi et al. have not been considered as they have been for the current study. The main differences in comparing the results reported here and the results of Shi et al. are seen in the polar residues as well as for D, V, F and Y. Similarly to what we report here, deviations from canonical PPII UVCD spectra in the region >200nm are seen for the aforementioned residues where our results differ. These deviations imply that sampling of conformations other than PPII occur for these residues. As previously detailed in this section, this is indeed what we see for these residues.

It is noteworthy in this context that Avbelj and coworkers used $^{3}J(H^{N},H^{\alpha})$ NMR coupling constants, in combination with IR and Raman spectroscopy, to determine conformational preferences of amino acids in a dipeptide context, and reported PPII fractions of 0.68 and 0.53 and β fractions of 0.17 and 0.43 for alanine and valine, respectively (*159*). The authors considered three representative conformations, and report only minor helical fractions. Qualitatively, their results point into the same direction as ours. Some earlier papers report propensities for residues in a non-glycine context. Creamer and coworkers used a P_3xP_3 host-guest system to derive PPII propensities for A, V, L, I, N and M (*160*). The respective values are listed in Table 4. Their PPII propensities, compared with our results, are lower for alanine and higher for valine, isoleucine and asparagine.

AA	$^3J(H^N,H^\alpha)$ X(PPII)[a]	$^3J(H^N,H^\alpha)$ X(PPII)[b]	$^3J(H^N,H^\alpha)$ X(PPII)[c]	$^3J(H^N,H^\alpha)$[d]	$^3J(H^N,H^\alpha)$[e]	X(PPII)[f]
A	6.11 (0.79)	5.73 (0.82)	6.02 (0.63)	6.08	6.1	0.63
V	7.46 (0.40)	7.05 (0.74)	7.32 (0.53)	7.55	7.2	0.49
I	7.47 (0.42)	7.52 (0.52)	7.37	7.50	7.1	0.50
L	6.78 (0.56)	7.15 (0.57)	6.84	6.99	6.8	0.58
F	7.45 (0.42)	6.97 (0.64)	7.17	7.35	7.3	-
Y	7.37 (0.43)	6.91 (0.63)	7.12	7.31	7.8	-
E	6.99 (0.54)	6.78 (0.68)	7.02	6.50	6.7	-
D	7.44 (0.04)	7.50 (0.55)	7.51	6.93	7.8	-
R	6.66 (0.62)	7.20 (0.64)	6.91	6.92	6.9	-
K	6.60 (0.50)	7.10 (0.58)	6.85	6.92	7.0	-
M	7.08 (0.64)	7.70 (0.50)	7.09	6.97	7.1	0.57
S	6.99 (0.45)	6.30 (0.77)	7.05	6.62	7.0	-
T	7.73 (0.35)	7.65 (0.55)	7.32	7.64	7.9	-
C	7.29 (0.51)	7.60 (0.56)	7.30	7.10	7.3	-
N	7.53 (0.46)	7.57 (0.67)	7.50	7.29	7.7	0.55

Table 4: Comparisons of $^3J(H^NH^\alpha)$ NMR coupling constants, and in parentheses, PPII fractions, for individual amino acids indicated for different studies. [a]Results found in this study; [b]Results from Shi et al. in a GGXGG context (*72*); [c] NMR and vibrational spectroscopic results obtained from Avbelj and coworkers for amino acid dipeptides (*79, 159*); [d] ^3J-coupling constants from a GGXGG study using 6M GdmHCl by Plaxco et al. (*161*); [e] ^3J-Coupling constants obtained from distributions of ϕ values from coil libraries (*79*); [f] PPII propensities estimated by Creamer and coworkers for the X residue in P_3XP_3 peptides (*97*).

Eker et al. used average conformations of several AxA peptides to illustrate the different propensities of guests residues (*71*). The authors then interpreted their results qualitatively in terms of a two conformer (PPII and β-strand) model. In agreement with our findings, their results suggest rather mixed populations for S, M, F and L. As in other earlier studies (*162, 163*), their obtained PPII propensity of alanine is underestimated. Although some of our previous work estimated PPII content of terminal amino acids, a

comparison is nevertheless useful (*106*). For the terminal amino acids in alanine based dipeptides, A, K, L, I, S, V and T, Hagarman et al. reported PPII fractions of 0.63, 0.48, 0.48, 0.39, 0.36, 0.35 and 0.19, respectively. These values compare surprisingly well with what we obtained in the current study, except for alanine and threonine, in view of the limited method and nature of the considered residues.

6.5 Significance of Intrinsic Propensities

Taken together, new insights emerge from the present study. In contrast to what is indicated by coil libraries (*62, 69*) and MD (*93*) simulations, all the investigated amino acid residues have a very limited (< 0.15) propensity for sets of dihedral angles found within those of helical conformations. It is intriguing to transfer our results into the parameters of the frequently used Zimm-Bragg theory. In this approach, the product, σs, of the nucleation parameter, σ, and the statistical weight of a helical conformation, s, reflects the probability of the formation of a peptide/protein segment comprising 3 helical residues. We switch back to the notation of α-helix in this explanation of helix→coil parameters. Hence, we can estimate, e.g., for alanine $\sigma s = (\chi_\alpha/(\chi_{PPII}+\chi_\beta+\chi_\gamma))^3$ which yields 1.5×10^{-4}. This value is considerably lower than those obtained from host-guest experiments ($\sim 8 \times 10^{-4}$). However, one has to consider a rather large uncertainty for any of the mole fractions obtained for the minor species of the conformational ensemble. A χ_α value of 0.085 would already be consistent with the experimental σs-value, which is well in the range of experimental uncertainty. Moreover, computational evidence suggests that a helical conformation does not tolerate a β-strand as its neighbor (*65*). It is reasonable to assume that α_L and γ-turns can also be excluded so

that $\sigma s = \chi_{\alpha r}{}^{3} \Big/ \left[(1-\chi_{\alpha r})\chi_{PPII}^{2} \right]$. Thus, a value of 0.077 already reproduces the experimental σs value. Hence, the low helix propensity derived from our data makes perfect sense with respect to known helix⇔coil transition parameters. Based on this estimation, one can further conclude that the higher propensities for right-handed helical structures indicated by coil libraries suggest that the initiation of helix formation is generally more likely in larger proteins than in short peptides, which is in accordance with the fact that many helix forming segments of proteins are disordered once they are separated from their protein context (*164*).

It should be noted in this context that MD simulations generally overestimate the nucleation parameters of polyalanine peptides (*157*), which reflects an overestimation of the population of (right-handed) helical conformers in the unfolded state by many force fields (*157, 165*). Exceptions from the rule are the modified Amber force field A94/MOD, OPLS/AA/L (*166*) and the force field of Cornell et al. (*167*) modified by Garcia and Sanbonmatsu (*168*). It is interesting to note that the modified Amber force field was also found to account for the high PPII propensity of alanine (*107*).

The studies of Tran et al. (*73*) and Shi et al. (*158*) led to the conclusion that PPII is the default conformation for most amino acid residues. Our results support this notion only for alanine, methionine and arginine and to a lesser extent for leucine, glutamic acid, cysteine and lysine, but not for serine, threonine, valine, isoleucine, phenylalanine and tyrosine, for which an equal PPII and β-like preference was determined. Alanine is a special case in that its PPII propensity exceeds the corresponding propensities of other residues even more than suggested by coil libraries (*61, 152, 169*). While the

distributions of the PPII sub-ensembles are all centered close to the canonical value, the respective β-strand distributions differ in terms of their positions, and some of them depart substantially from the coordinates of parallel and antiparallel β-sheets, in that they populate the region between PPII and canonical β-strand troughs of the Ramachandran plot.

Chapter 7 Conclusion

The results presented here constitute our attempt to provide a sound basis for exploring the conformational manifolds of unfolded peptides and proteins and how they relate to intrinsic propensities of amino acids in a polypeptide chain. The values reported herein reflect a valuable intrinsic propensity scale for a representative subset of amino acids. Furthermore, this work emphasizes the necessity to have an indispensible intrinsic propensity scale for individual amino acids as a basis for comparison of further context dependent studies and the calibration of molecular mechanics force fields for computational biochemistry.

Section 2: The Conformational Heterogeneity of Cytochrome c

Chapter 1 Background

1.1 Cytochrome c

Electron transport chains are pivotal in natural biological processes such as photosynthesis in plants and respiration in animals. The mitochondrial electron transport chain involved in respiration transports electrons from an electron donor, e.g. NADH, to an electron acceptor (O_2) in the terminal part of the chain through a series of redox reactions. A proton pump gradient is established in connection with these redox reactions, which results in the production of ATP and is essential to sustain life. Involved in the electron transport chain are complexes that act as proton pumps and complexes that act as electron transporters. Most of the involved complexes are trans-membrane proteins, although cytochrome c is loosely associated to the inner membrane of the mitochondria. The function of cytochrome c in the electron transport chain is to shuttle electrons from complex III to complex IV (*170*) (Figure 19).

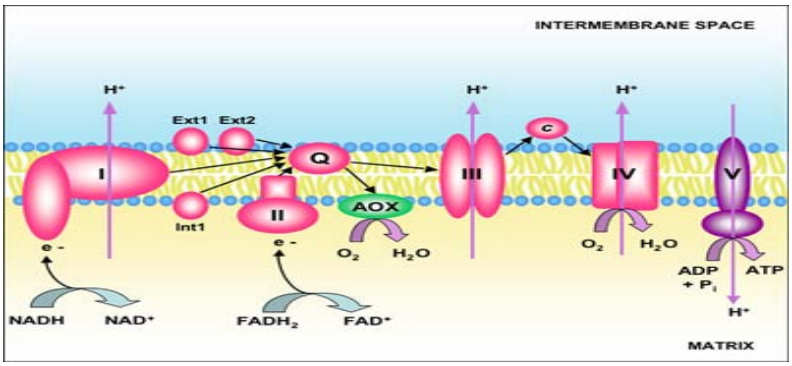


Figure 19: Schematic of the mitochondrial electron transport chain

This relatively small (~12.4kDa) and highly soluble heme protein plays a role in many, yet mechanistically undetermined, biological processes in addition to its role in electron

transport. It has been classified as an enzyme in some cases and was shown to exhibit peroxidase activity (*171, 172*). Also, cytochrome c has been shown to be an essential part of programmed cell death or apoptosis (*173, 174*). Upon its release from the mitochondria, cytochrome c activates caspases, which are ultimately responsible for cell death (*175*). It has become increasingly evident that cytochrome c undergoes a conformational change when bound to anionic phospholipids which is considered to be a precursor state to its role in apoptosis, whereupon the aforementioned caspases are set off by the aggregation of Apaf-1 upon the binding of the latter to cytochrome c (*174, 176*). The role cytochrome c plays in apoptosis has been shown to require a conformational change of the protein (*176*). Conformational flexibility of cytochrome c regarding its involvement in biological processes is a well-established phenomenon (*172, 177-179*), though the extent of the details of tertiary and secondary structure changes has yet to be fully assessed.

The structure of the native, fully folded cytochrome c has been solved by x-ray crystallography and confirmed by various other methods (*180-182*) (Figure 20).

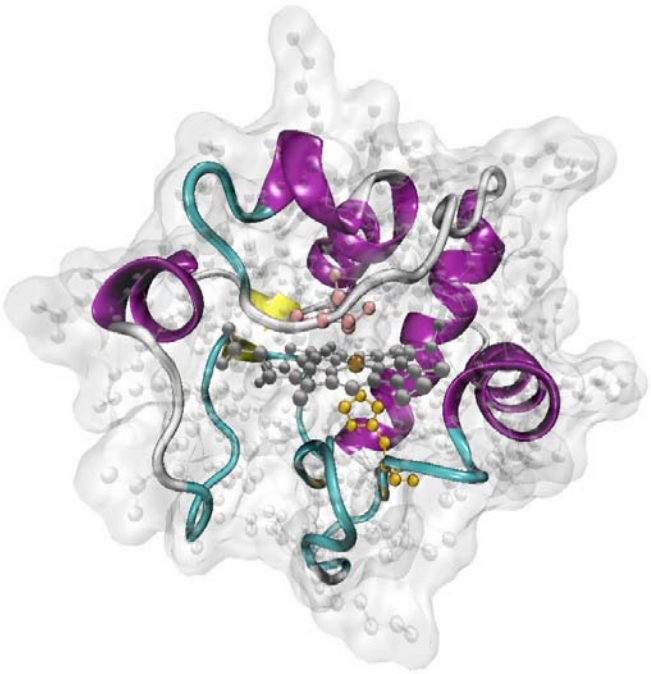


Figure 20: Ribbon structure representation of cytochrome c rendered using VMD software.

The structure of the so-called native state of this protein, obtained at neutral pH and room temperature in solution, is well known (*182, 183*). However, it is not clear whether this state is the most functionally relevant relating to its biological processes. The central iron atom of the heme in cytochrome c is capable of undergoing oxidation and reduction. The reduced state is remarkably stable (*184*) over a wide range of pH and temperature values compared to its oxidized counterpart; it requires temperatures of ~100°C and extreme pH values to unfold. The conformational flexibility of ferricytochrome c has been the subject of many recent studies, as it is thought that the states adopted at non-native conditions are the ones involved in biological processes. In fact, a conformational change has been monitored upon cytochrome c binding to cytochrome oxidase (*172, 178, 179*).

1.2 Non-native States of Ferricytochrome c in Solution

Conformational flexibility of cytochrome c has not only been observed in biological process, but in solution as well. The oxidized state of cytochrome c adopts five different states - termed I-V by Theorell and Åkesson (*185*) - in aqueous solution over the pH range 1-12. These native and non-native states have been characterized by spectroscopic methods such as resonance Raman, circular dichroism and absorption spectroscopy. During protein unfolding and denaturation, non-native states represent intermediates in the unfolding process (*186*). State I is the acid unfolded state at low ion concentrations, though an increase in ion binding changes it to a molten globule-like state (*187*). State II is a partially folded structure found in the pH range 2-4, with most of its secondary structure intact (*188, 189*). The so-called native state, state III, is adopted at neutral pH (4-9) and room temperature. The heme in the native state is proximally coordinated to the H18 and distally coordinated to the M80 residue of the surrounding protein (Figure 21). The coordination of the proximal histidine is common in heme centered proteins, though cytochrome c differs from more prominent heme proteins such as myoglobin and hemoglobin in that its central iron atom is coordinated at both axial sites by amino acids of the protein matrix and its heme chromophore is covalently linked to the protein through thioether bonds with cysteine residues 14 and 17. The sequence Cys-X-X-Cys-His is a conserved heme attachment motif in c-type cytochromes.

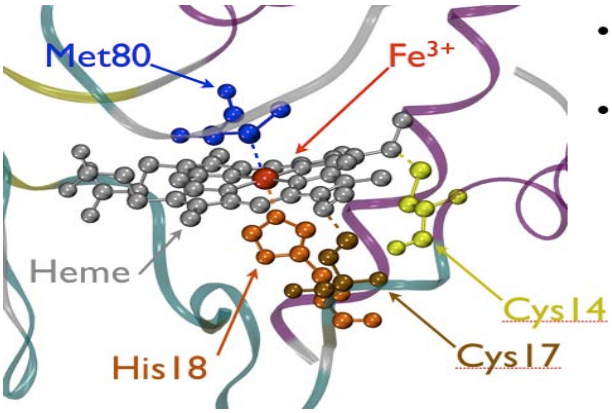


Figure 21: Active site environment of oxidized cytochrome c showing the most relevant amino acids coordinating the protein to the heme, rendered from VMD software.

An increase in pH induces a conformational change to state IV, known as the alkaline transition, which is identified by a misligation of the M80-heme iron linkage to one of two lysine residues, K73 or K79, in close proximity to the heme crevice (*190-193*). States IV_a and IV_b denote the fact that population of these states involve the sixth coordination of different lysine ligands, 73 and 79, respectively (*192, 194*). Assfalg *et al.* showed that the proximal side of the heme maintains its secondary and tertiary structure upon an increase in pH up to 11, though the distal side of the heme experiences an increased solvent exposure along with its misligation from M80 to K73 (*195*). Upon further increase in pH, another alkaline state, state V, is adopted where sixth ligand is unknown (*192*). They argue that the ε-amino acid side chains of the ligated lysines in state IV are replaced by strong field ligands in the pH range where state V is populated. The candidates for this ligation are hydroxides or alkoxides. These can be provided by water or the side chains of, e.g. Tyr or Thr. In fact, they identified three plausible candidates within close proximity to the heme; Tyr 67, Thr 78 and Wat 166. They excluded the Tyr and Thr based on similarities between the wild-type and Tyr67Phe and Thr78Ala

mutants' resonance Raman spectra for state V. They concluded that the most probable ligand is deprotonated water (Wat 166).

1.3 695nm Charge Transfer Band

A prominent spectral marker band used to probe the Fe-M80 ligation is the frequently investigated 695nm charge transfer (CT) band in the absorption and CD spectra, making it an optimal candidate to mirror the alkaline transition due to the replacement of M80 as a ligand during this transition (*196, 197*). Theorell and Åkesson first described this band (*198*), and later detailed its disappearance at pH values <2.5 and >9.35 (*185*). Schejter and George showed that this band disappears or loses intensity at high temperatures and in the presence of denaturing conditions (*196*). The 695nm band thus serves as a tool to probe chemically and thermally denatured states where M80 has been replaced as a sixth ligand, which also occurs in the binding of cytochrome c to cardiolipin containing phospholipids (*199*). Previously, Eaton and Hochstrasser assigned this band to as an A_{2u} porphyrin→ A_{1g} (d_z^2) iron CT transition (*200*). More recently this band has been attributed to an S(p_z) → Fe^{3+}(d_{xy}) CT transition (*201*), which is consistent with its disappearance upon misligation of the methionine distal ligation. Hence, the 695nm charge transfer band is commonly used to probe the integrity of the sixth coordination of the heme to M80.

1.4 Thermal Intermediate States

The thermal stability of cytochrome c depends on the redox state of the heme iron. Ferricytochrome c begins unfolding at relatively low temperatures (40° C) compared to that of the ferrous state, which requires much higher temperatures (~100° C) to unfold.

The thermal transition of the oxidized protein involves the population of a thermodynamic intermediate state, termed state III_h, in the temperature range from 50-60° C, before the final step of unfolding into state U occurs at temperatures above 70° C (*197*). The transition into the thermally populated intermediate state can again be probed by the aforementioned 695nm CT band, which loses significant intensity due to the weakening of the Fe-M80 ligation as the protein unfolds. It was argued that the M80 coordination is replaced upon population of the thermal intermediate (*186*). It is known that both transitions cause a decrease in the redox potential (*202*). A drastic decrease in the effective pK-value of the alkaline transition upon increasing temperature was observed, leading the authors to state that an appreciable fraction of state IV could be populated at neutral pH in the intermediate temperature range (*186*). Using this argument, Taler et al. rationalized that the states III_h and IV were identical based on a detailed NMR study. A decrease in this pK-value has certainly been observed elsewhere, though several findings contradict the correspondence of the alkaline and thermally populated states (*191, 202*). Battistuzzi et al. showed that the van't Hoff plot of the apparent pK-value of the alkaline transition is biphasic, which these authors interpreted as indicative of the existence of a neutral high temperature state which is neither identical with the alkaline nor with the native state (*191*). Questions have arisen regarding the comparability of temperature and pH induced non-native states of cytochrome c, which indicate the necessity to thermodynamically and structurally characterize the former and the latter in more detail.

1.5 Solution Conditions

Many studies aimed at elucidating the structure-function relationship of cytochrome c have been performed using different solution conditions. For instance, an array of buffers, buffer concentrations and salt concentrations as well as various stabilizing agents are used throughout different studies. The surface of cytochrome c is highly charged (+8) at neutral pH values which makes the protein susceptible to electrostatic interactions with the ions in solution. The Sola group has shown that anions like Cl^-, HPO_4^{2-}, HCO_3^-, NO_3^-, and SO_4^{2-} as well as others bind to the positive patches of the protein (*203*). It has been accepted in the past that anion binding to these patches has a negligible effect on the structure and heme environment of the protein. Some studies have confirmed the ionic effects on the solution structure of cytochrome c. Feng and Englander reported salt-dependent chemical-shift changes in the area of residues 83-89 (*204*). Moench et al. reported that the distance between F82 and the heme decreases upon an increase in ionic strength (*205*). According to Trewhella et al., who used small angle X-ray scattering to estimate the solution structures of reduced and oxidized horse heart cytochrome c, the protein expands at low and compresses at high ionic strength (*206*). Liu et al. used resonance Raman spectroscopy to detect that Tyr 48 and Trp 59 are in fact affected by changes in ionic strength (*207*).

1.6 Heme-Protein Interactions

The functions of heme proteins are vast, though a majority of these proteins have structurally similar active sites. Functional properties of heme proteins have been related to the interaction of the prosthetic heme group and its surrounding protein matrix, in particular, the side chains of the residues in the surrounding heme cavity. For

instance, in hemoglobin, a doming deformation involving an out-of-plane displacement of the iron atom in the heme which is induced by the proximal ligation allows for an oxygen molecule to bind, which then triggers the onset of a relaxed conformation of the remaining subunits, inducing a higher oxygen affinity, which gives this protein its essential biological function. Cytochrome c belongs to an important class of redox active proteins, where the comparatively high reduction potential of the former is predominantly due to the sulfur of the distal coordination and the influence of the internal electric field. In this context, the perturbations of the heme imposed by its surroundings, i.e. the protein matrix and peripheral substituents are particularly relevant in destabilizing its redox potential. Although it is still a matter of debate, point mutations of yeast cytochrome c which make the protein more analogous to the horse heart protein are known to decrease the redox potential of the former. It is thought that this effect is due to a decreased ruffling deformation based on the fact that studies on Zn-porphyrins suggest that out-of-plane deformations decrease the redox potential substantially (*208*). INDO/s calculations indicate that reduction potentials of the conjugated ring and metal are affected by ruffling (*208*). Heme-protein interactions have been probed by various spectroscopic techniques such as resonance Raman, absorption, circular dichroism, EPR, NMR and photon echo spectroscopy, although X-ray crystallography is considered to be the most accurate method to probe heme distortions. It has become increasingly evident that even small distortions, undetectable in X-ray structures, are functionally relevant. Taking this into

consideration, the exploration of these distortions and their functional significance is of the utmost importance.

1.7 Introduction

For the studies described in this thesis, we have used polarized resonance Raman, absorption and circular dichroism spectroscopy to probe functionally relevant distortions of the heme as well as tertiary and secondary structure changes of the surrounding protein environment. We also identify thermodynamic intermediates in the pH and temperature induced unfolding landscape of ferricytochrome c. In what follows, a brief background of absorption and CD spectroscopy will be given. After that, the origin of distinct optical bands in the absorption spectrum of metal porphyrins, known as the Gouterman four-orbital model will be discussed. Electronic and vibronic perturbations and their influences on absorption and CD spectra will be highlighted. Then, resonance Raman scattering of metal porphyrins and its usability for probing in-plane and out-of-plane distortions of the prosthetic heme group will be explained. The results section is then organized as follows. The results are presented where CD and absorption spectroscopy were used as tools to resolve the splitting of the optical bands into their individual components, as well as to decompose the 695nm CT band into individual subbands. We then use the heterogeneity of the 695nm band to probe the thermal unfolding of oxidized cytochrome c. We use this band combined with observed changes in a different heme absorption band, namely the Soret band, to identify the role of acetate and monohydrogen phosphate anions in solution. The CT band was subsequently used to identify an intermediate involved in the alkaline transition, which

is populated at low ionic strength conditions. After that, the results of a study using CD and absorption spectroscopy of the Soret and CT bands to elucidate the structural and thermodynamic distinction of states III_h and IV, as well as structurally and thermodynamically characterize the alkaline states, are presented. Finally, we used the depolarization ratios and normalized intensities of a number of Raman bands from Raman active in-plane and out-of-plane modes to identify and quantitatively compare in-plane and out-of-plane distortions of the heme group in three different ferrocytochrome c isoforms.

Chapter 2 Theoretical Background

2.1 Theory of absorption and circular dichroism

Electronic circular dichroism has been used frequently in the past to estimate secondary structure fractions and tertiary structure changes of peptides and proteins. It was not until recently that the growth in understanding of the theory and improvement in instrumentation allowed for an accurate quantitative determination of secondary structure for proteins and peptides and changes in tertiary structure for proteins with chromophores. The combination of visible absorption and CD spectroscopies has recently been used to explore specific heme protein interactions transduced through space by the internal electric field in the heme pocket (*209-212*). Visible circular dichroism and absorption spectroscopy probe the heme environment in cytochrome c. Specifically, CD probes the chirality of the heme environment. We use the two techniques in combination to probe functionally relevant structural changes in the heme vicinity upon populations of non-native conformational states of oxidized cytochrome c.

The basic theory of absorption is well known and the theory of circular dichroism has been described elsewhere (*213*), so this section will be confined to a brief description of both. Absorption at a given wavenumber can be explained by the well known Beer-Lambert law:

$$A(\tilde{\nu}) = \varepsilon(\tilde{\nu})lc \tag{14}$$

Where ε is the molar extinction coefficient at the respective wavenumber, *l* is the pathlength in cm that light travels through the sample and *c* is the concentration expressed in molar units. The former is typically expressed in units of $M^{-1}*cm^{-1}$ and reflects the capability of the medium to interact with the incident radiation field and can be represented by a superposition of Voigtian profiles for electronically allowed transitions:

$$\varepsilon(\tilde{\nu}) = \sum_{l} \frac{1}{\sigma_l \sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{f_l \Gamma_l / 2\pi}{((\tilde{\nu}_l) - (\tilde{\nu}))^2 + \Gamma_l^2} e^{-(\tilde{\nu} - \tilde{\nu}_l)/2\sigma_l^2} d\tilde{\nu} \tag{15}$$

where

$$f_l = A_\varepsilon \tilde{\nu}_l |\langle l|\vec{\mu}|g\rangle|^2 = A_\varepsilon \tilde{\nu}_l \left| \int \psi_l^* \vec{\mu}\, \psi_g d^3 r \right|^2 \tag{16}$$

is the oscillator strength associated with the dipole $\vec{\mu}$ induced by the radiation field. $\psi_{l,g}$ are the electronic wavefunctions of the states $|g\rangle$ and $|l\rangle$. A_ε= 108.92 D^{-1} M^{-1} cm^{-1} (D: Debye=10^{-18} esu·cm). The transition dipole moment is defined as:

$$\vec{\mu} = \sum_i e\,\vec{r_i}$$

(17)

where *e* is the elementary charge and $\vec{r_i}$ is the coordinate of the *i-th* electron with respect to an arbitrary coordinate system. The transition dipole moment couples the ground state |*g*> to the excited state |*l*> by absorbing a photon from the radiation field. $\tilde{\nu}_l$ is the wavenumber corresponding to the energy difference between |*l*> and |*g*>. Γ_l and σ_l are the halfwidths at half maximum of the Lorentzian and Gaussian profiles, respectively associated with the excitation of state |*l*>. The halfwidth of a Lorentzian profile reflects the lifetime of the excited state, while the halfwidth of a Gaussian profile reflects the inhomogeneous broadening.

Circular dichroism can be considered as a measure of chirality because it is defined as the differential absorption of right- and left-handed circular polarized light:

$$\Delta\tilde{\varepsilon}(\tilde{\nu}_l) = \tilde{\varepsilon}_L(\tilde{\nu}_l) - \tilde{\varepsilon}_R(\tilde{\nu}_l)$$

(18)

This difference cannot be accounted for by solely considering the electronic dipole moment, because they would be equal for both field rotations. However, the radiation field also induces a magnetic moment:

$$\vec{m} = \frac{e}{2m_e}\sum_i \vec{L}_i$$

(19)

which may couple the ground and excited states, where m_e is the mass of a resting electron and $\vec{L}_i$ is the angular momentum operator associated with the *i-th* electron. If one considers contributions from both the electronic and magnetic transition dipole moments, the effective oscillator strength for each field rotation (L, R) is written as:

$$f_l^{L,R} = A_\varepsilon \tilde{\nu}_l(\langle l|\vec{\mu}|g\rangle + Im\langle l|\vec{m}|g\rangle)(\langle l|\vec{\mu}|g\rangle - Im\langle l|\vec{m}|g\rangle) = f_l + G_l \pm 2R_l$$

(20)

The magnetic moment is imaginary and the term

$$R_l = Im(\langle l|\vec{\mu}|g\rangle\langle l|\vec{m}|g\rangle)$$

(21)

is the rotational strength of the transition. The term G_l is the oscillator strength of the magnetic transition. The magnetic dipole moment exhibits rotational symmetry so that the respective transition requires circular polarized light. By substituting f_l with $f_l^{L,R}$, the extinction profiles for left and right handed circular polarized light can be calculated. The CD spectrum is the difference of the former and the latter and can be written again as the superposition of Voigtian profiles:

$$\Delta\varepsilon(\tilde{\nu}) = \sum_l \frac{1}{\sigma_l\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{R_l\Gamma_l/2\pi}{((\tilde{\nu}_l) - (\tilde{\nu}))^2 + \Gamma_l^2} e^{-(\tilde{\nu}-\tilde{\nu}_l)/2\sigma_l^2} d\tilde{\nu}$$

(22)

R_l can be either positive or negative, thus giving rise to positive and negative Cotton bands. One would expect the shapes of corresponding CD and absorption band profiles to be identical in the case where the profiles from different transitions do not overlap. In molecules, however, this is not the case, where absorption bands are comprised of a series of subbands arising from multiple transitions into excited vibrational states coupled to the excited electronic state. The shapes of corresponding CD and absorption profiles associated with the same electronic transition should still be identical irrespective of the underlying vibronic structure in the case where strong non-Condon coupling is absent (*214, 215*). This is because the electronic and magnetic transition dipole moments and their relative orientations are independent on the involved vibrational states. If the profiles of different electronic transitions substantially overlap, the situation changes and the band shapes of the CD and absorption profiles cannot be assumed to be identical. Similarly, if the rotational strengths of the overlapping individual transitions carry opposite signs, the previous statement holds true. The latter is exactly the case for the native state of cytochrome c. These facts along with experimental values such as the difference between peak positions and halfwidths of individual profiles can be used to reproduce the CD and absorption spectra of cytochrome c.

2.2 Four-Orbital Model

The absorption spectrum of cytochrome c stems from the canonical absorption spectrum of a metal porphyrin, seen in Figure 22. In this figure, two absorption bands dominate the spectrum, namely the B (Soret) and Q-bands.

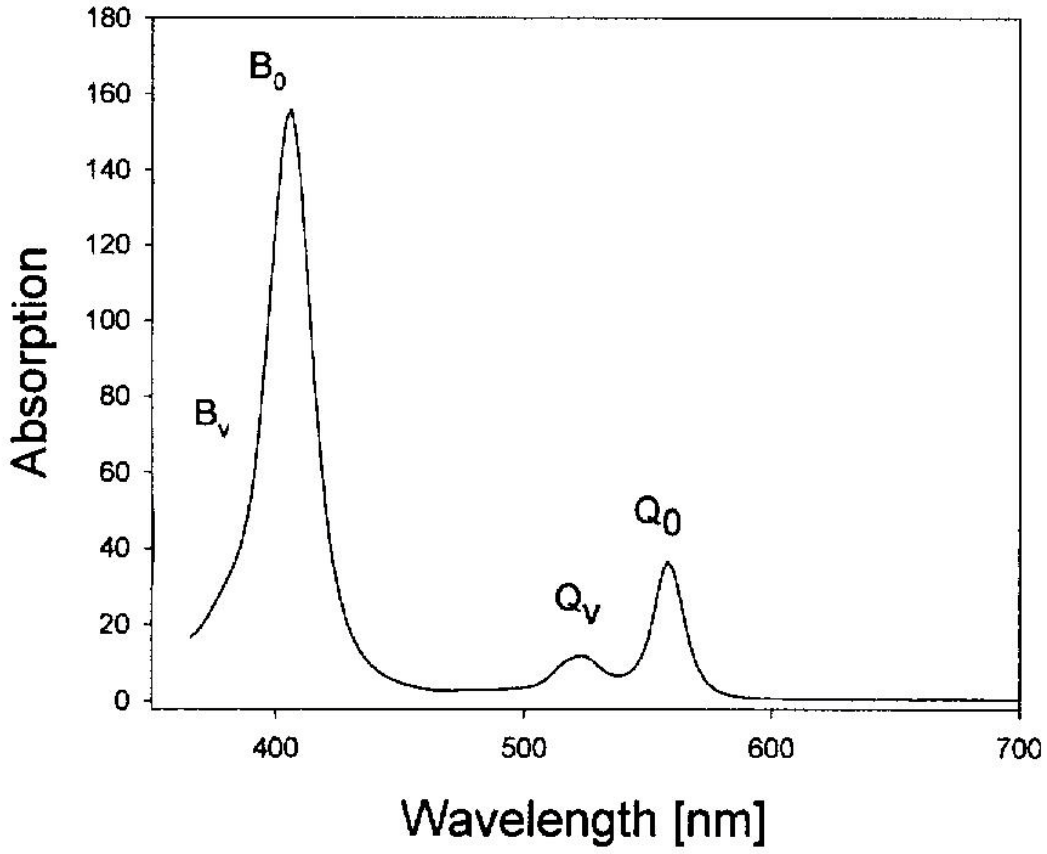


Figure 22: Taken from ref (*216*) and modified. Visible absorption spectrum of NIckeloctaethylporphyrin in CS_2.

The heme chromophore in cytochrome c is a macrocyclic molecule called a metal porphyrin. Unsubstituted Planar metal porphyrins have D_{4h} symmetry. Simpson analogized a porphyrin to an 18-membered cyclic polyene (*217*). Upon interacting with electromagnetic radiation the delocalization of the π-electrons on the ring gives rise to $\pi \rightarrow \pi^*$ transitions in the visible regime. The π-molecular orbitals are antisymmetric with respect to the inversion center constituted by the metal ion and thus transform like the irreducible representations A_{1u}, A_{2u}, $B_{1u,}$ B_{2u} and E_g (*218*). The HOMOs transform like A_{1u} and A_{2u} and the LUMOs transform like E_g. The electronic transitions from A_{1u} and A_{2u} to E_g do not account for the characteristics of the absorption spectrum of porphyrins like

the frequency position of two of the observed bands, the difference in intensity of the experimental B_0 and Q_0 bands and the presence of the B_v and Q_v bands.

The theoretical basis of the absorption spectra of porphyrins has been developed in considerable detail by Gouterman and coworkers (*219*) and later by others (*220*). Consequently this theory is known as the Gouterman four-orbital model. In what follows a condensed version of this will be presented. The cyclic polyene theory undoubtedly accounts for the appearance of two bands, namely the near-UV and visible bands. This theory does not consider electron-electron interaction. To account for this interaction, the Schrödinger equation was modified to include a two-electron operator which accounts for the Coulomb interaction and couples the electronic configurations related to the same symmetry, which is known as configuration interaction. The representations, A_{1u}, A_{2u}, E_{gx}, and E_{gy} characterize the orbitals denoted by Gouterman as the two HOMOs and two LUMOs, respectively, which resulted from linear combinations of atomic orbitals-molecular orbital (LCAO-MO) calculations. The electronic transitions from the two HOMOs to the LUMO give rise to four different electronic configurations which are consequently described by the following wavefunctions:

$$\begin{vmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{vmatrix} = \begin{vmatrix} A_{1u}E_{gx} \\ A_{1u}E_{gy} \\ A_{2u}E_{gx} \\ A_{2u}E_{gy} \end{vmatrix} \qquad (23a)$$

The wavefunction $|A_{1u}E_{gx}\rangle$ signifies that one electron has been transferred from the HOMO (A_{1u}) to the LUMO (E_{gx}). $|\mu_i\rangle$ provides the basis for which to formulate the electronic wavefunctions:

$$|\Psi_n\rangle = \sum_{i=1}^{4} b_{ni}|\mu_i\rangle$$

(23b)

The coefficients, b_{ni} are found by diagonalizing the electronic Hamiltonian:

$$\hat{H}_{el} = \begin{bmatrix} E_{A_{1u}E_{gx}} & 0 & 0 & \delta/2 \\ 0 & E_{A_{1u}E_{gy}} & -\delta/2 & 0 \\ 0 & -\delta/2 & E_{A_{2u}E_{gx}} & 0 \\ \delta/2 & 0 & 0 & E_{A_{2u}E_{gy}} \end{bmatrix}$$

(24)

Where $\delta = 2\langle A_{2u}E_{g\rho}|\sum_{i<j} U_{ee}(r_i, r_j)|A_{1u}E_{g\sigma}\rangle$ and ρ≠σ=x,y. $U_{ee}(r_i, r_j)$ is the two electron operator which accounts for configuration interaction, (r_i, r_j) are electronic coordinates and $E_{A_{1u}E_{g\rho}}$ and $E_{A_{2u}E_{g\sigma}}$ are the energies of the individual configurations.

In the case of configuration interaction $\left|A_{1u}E_{gx}\right\rangle$ and $\left|A_{1u}E_{gy}\right\rangle$ couple with $\left|A_{2u}E_{gx}\right\rangle$ and $\left|A_{2u}E_{gy}\right\rangle$. The total wavefunction can be written as a linear combination of these wavefunctions. It is convenient to start by considering that the unperturbed molecular orbitals are accidentally degenerate so that $E_{A2u} = E_{A1u}$. The electronic states are then

given by a 50:50 mixing of the state vectors of the above configurations in the nomenclature of Schweitzer-Stenner (*220*) as:

$$\left|B_x{}^0\right\rangle = 2^{(-1/2)}\left[\left|A_{2u}E_{gx}\right\rangle - \left|A_{1u}E_{gy}\right\rangle\right] \quad (25a)$$

$$\left|B_y{}^0\right\rangle = 2^{(-1/2)}\left[\left|A_{2u}E_{gy}\right\rangle + \left|A_{1u}E_{gx}\right\rangle\right] \quad (25b)$$

$$\left|Q_x{}^0\right\rangle = 2^{(-1/2)}\left[\left|A_{2u}E_{gx}\right\rangle + \left|A_{1u}E_{gy}\right\rangle\right] \quad (25c)$$

$$\left|Q_y{}^0\right\rangle = 2^{(-1/2)}\left[\left|A_{2u}E_{gy}\right\rangle - \left|A_{1u}E_{gx}\right\rangle\right] \quad (25d)$$

Two-fold degenerate electronic eigenstates $\left|B_{xy}{}^0\right\rangle$ and $\left|Q_{xy}{}^0\right\rangle$ are obtained which transform like the irreducible representation E_u in D_{4h} symmetry. Configuration interaction leads to an energy separation between the two electronic states:

$$\Delta E = E_B^0 - E_Q^0 = \delta \quad (26)$$

Using equations 25a-d, the electronic transition dipole moments are calculated as:

$$R_{B\sigma}^0 = 2^{(-1/2)}\left\langle \mathrm{A_{2u}E_{g\sigma}}\middle|\mathrm{R_\sigma}\middle|\mathrm{g}\right\rangle + \left\langle \mathrm{A_{1u}E_{g\sigma}}\middle|\mathrm{R_\sigma}\middle|\mathrm{g}\right\rangle \quad (27a)$$

$$R_{Q\sigma}^{0} = 2^{(-1/2)} \langle \mathrm{A_{2u}E_{g\sigma}}|\mathrm{R_{\sigma}}|\mathrm{g}\rangle - \langle \mathrm{A_{1u}E_{g\sigma}}|\mathrm{R_{\sigma}}|\mathrm{g}\rangle$$

(27b)

Where σ=x,y and $|g\rangle$ is the ground state. The matrix elements $\langle \mathrm{A_{2u}E_{g\sigma}}|\mathrm{R_{\sigma}}|\mathrm{g}\rangle$ and $\langle \mathrm{A_{1u}E_{g\sigma}}|\mathrm{R_{\sigma}}|\mathrm{g}\rangle$ turned out to be nearly identical for the molecular orbitals of Longuet-Higgins et al. (*218*) which makes the transition from the ground into the Q^0 state nearly forbidden. In contrast to the assumed ideal state, the energies $E_{A_{1u}E_{g\rho}}$ and $E_{A_{2u}E_{g\sigma}}$ are not accidentally degenerate due to a one electron interaction between $\left|B_{xy}{}^{0}\right\rangle$ and $\left|Q_{xy}{}^{0}\right\rangle$ which in turn causes the 50:50 mixed states to unmix and thus makes the transition into the excited Q^0-state allowed. The corresponding interaction matrix is written as:

$$H'_{el} = \begin{bmatrix} E_{Qx}^{0} & 0 & \delta_{A1g} & 0 \\ 0 & E_{Qy}^{0} & 0 & \delta_{A1g} \\ \delta_{A1g} & 0 & E_{Bx}^{0} & 0 \\ 0 & \delta_{A1g} & 0 & E_{By}^{0} \end{bmatrix}$$

(28)

Where $\delta_{A1g} = \langle Q_{xy}^{0}| \sum_i \partial H_{\mathrm{el}}/\partial q_i^{A1g} \left|B_{xy}^{0}\right\rangle \partial \overline{q}_i^{A1g}$ reflects the coupling of the 50:50 mixed states due to changes in electronic energy along coordinates q_i^{A1g}. Changes along these coordinates do not change the symmetry of the porphyrin. A_{1g} is the totally symmetric representation of the point group D_{4h}. A deformation of this type can be achieved by changing the oxidation state of the central metal ion, where in response, the ionic

radius changes inducing a totally symmetric perturbation of the porphyrin macrocycle. Solving the Schrödinger equation using the electronic Hamiltonian in eq. 28 results in:

$$|Q_\sigma\rangle = cosv|Q_\sigma^0\rangle - sinv|B_\sigma^0\rangle \tag{29a}$$

$$|B_\sigma\rangle = sinv|Q_\sigma^0\rangle + cosv|B_\sigma^0\rangle \tag{29b}$$

Where σ=x,y and the unmixing parameter v is written as:

$$tanv = \frac{\delta_{A1g}}{E_B^0 - E_Q^0} \tag{30}$$

The one-electron interaction δ_{A1g} separates the B_0 and Q_0 bands further apart and redistributes intensity from the B_0 to the Q_0 band.

This description of the electronic states coupling to nuclear motions has thus far neglected. Vibrational side bands are present in both the B- and Q-bands, denoted with a subscript, v. A vibronic approach is necessary to describe these side bands which reflect the transitions, 0→1, from the electronic and vibrational ground state into the first excited state of various vibrational modes in the heme group.

Vibronic sidebands arise from inter and intra-state vibronic coupling of the two lowest energy electronic states. Vibronic coupling can arise from asymmetric electronic and

vibronic perturbations imposed by peripheral substituents, the surrounding protein matrix and the resulting internal electric field, but this is not strictly the case. Even in the absence of these perturbations, vibronic coupling can arise from inter and intra manifold coupling. The vibronic side bands of the B and Q bands can be described a Herzberg-Teller perturbation approach explained in the next section. The B_v band, which was shown to be dominated by the A_{1g} mode, v_4, for iron porphyrins in heme proteins (*211*), and is known to be comprised mostly of modes from intrastate Frank-Condon coupling. Vibronic coupling between the B- and Q-states pumps intensity into the Q_v band. The presence of the Q_v band is mostly due to interstate B_{1g} and A_{2g} Herzberg-Teller coupling.

2.2.1 Electronic and Vibronic Perturbations

Electronic and vibronic perturbations are capable of splitting optical bands of heme proteins. The splitting of the Q-band has been resolved before by using cryogenic absorption measurements (*209, 221*). Dragomir et al. observed the individual components of the B band in the CD spectra which was subsequently used to estimate the splitting of the B and Q bands for the different redox states of various cytochrome c isoforms (*222*). Changes in electronic and vibronic perturbations which occur upon populating non-native states induce changes in the CD spectra. For example, structural rearrangements in the vicinity of the heme environment induce changes in the internal electric field, which has been determined by Manas et al. (*223*) and confirmed by Schweitzer-Stenner (*212*), to contribute to the B-band splitting. In what follows, these

perturbations and their influence on the Soret band CD spectra of heme proteins will be explained.

Assuming a metalloporphyrin (heme) is subject to an external potential field, the electronic Hamiltonian of the system is written as:

$$\widehat{H}_{el} = \widehat{H}_{el,0} + V \tag{31}$$

The eigenfunctions corresponding to the unperturbed Hamiltonian can be attributed to the 50:50 mixed states, i.e. the singlet ground state $|g\rangle$ and the excited states $\left|B^0_{x,y}\right\rangle$ and $\left|Q^0_{x,y}\right\rangle$. The excited states are two-fold degenerate having E_u symmetry (*219*). Thus, the electronic interactions between the excited states require that the symmetry representation of the potential *V* contains contributions from A_{1g}, B_{1g}, B_{2g} or A_{2g}. For coupling between ground and excited states, the symmetry of the perturbing potential must have E_u symmetry. The interaction matrix elements, V^{Γ}_{lm}, of the electronic Hamiltonian are thus written as $V^{\Gamma}_{lm} = \langle m|V^{\Gamma}|l\rangle$, where *l* and *m* represent the excited states B and Q and Γ is the symmetry representation of the perturbing potential of the D_{4h} point group. The perturbations of B_{1g}, B_{2g} and A_{2g} symmetry are generally assumed to be small, though this is not the case for $V^{A_{1g}}_{BQ}$. The electronic Hamiltonian can then be diagonalized with respect to this term to get the so-called unmixed states. Subsequently, Rayliegh-Schrödinger time independent perturbation theory can be used to find the non-diagonal elements of the electronic Hamiltonian matrix. Further more, a quadratic stark effect imposed by the internal electric field can induce splitting as well

(*212*). If the splitting of optical bands were solely due to electronic perturbations then the splitting of the B^0 and Q^0 bands would be equal in the framework of the four-orbital model. Vibronic perturbations must therefore be considered to account for the fact that the B and Q band splittings are in fact unequal. Vibronic perturbations reflect the variation of the perturbing potential, V^{Γ}, on the nuclear coordinates where:

$$\chi_{lm}^{\tilde{\Gamma}} = \left\langle m \left| \frac{\delta V^{\Gamma}}{\delta q_{j'}^{\Gamma'}} \right| l \right\rangle q_j^{\Gamma'} \tag{32}$$

Where $q_j^{\Gamma'}$ is the q-th normal coordinate exhibiting the symmetry Γ' in D_{4h}. The electronic states identified with *l* and *m* are the 50:50 mixed states $\left|B_{x,y}^0\right\rangle$ and $\left|Q_{x,y}^0\right\rangle$. Group theory thus dictates that for $\chi_{lm}^{\tilde{\Gamma}} \neq 0$, $\tilde{\Gamma} = \Gamma \otimes \Gamma'$ = A_{1g}, B_{1g}, B_{2g} or A_{2g}. Vibronic perturbations change the eigenenergies of the vibronic B and Q states (*224*). The splitting of optical bands is particularly relevant for heme proteins where one component of the excited state transition may carry more rotational strength or even an opposite sign of rotational strength which are the cases for the B-bands of deoxymyoglobin and native ferricytochrome c. In the former, one component of the B-state transition carries more rotational strength so that the splitting of the individual components is reflected by the noncoincidence between the absorption and CD profiles (*211*). In the latter, an observed couplet situated at the ε_{max} position comes from the fact that the components of the excited B state carry opposite rotational strengths and a sufficient splitting between them produces a couplet. When the protein undergoes a

transition into its non-native temperature or alkaline induced states, the splitting of the B band decreases (practically zero) and a cotton band situated at the ε_{max} position becomes apparent.

2.3 Resonance Raman Scattering of Porphyrins

Raman scattering results in an induced dipole moment by the excitation radiation electric field vector which can be written as $\vec{\mu} = \hat{\alpha}\vec{E}$, where $\hat{\alpha}$ is a polarizability tensor. The distortion of a molecule in an electric field is measured by its polarizability. In the case where the excitation frequency can be tuned within the absorption band of a chromophore, the vibrational modes associated with chromophore become resonance enhanced. This is the cases for porphyrins so that vibrational modes coupled to the macrocycle appear as strong bands when the excitation frequency is tuned to the resonance frequency of the porphyrin absorption bands. The tensor elements for the Raman tensor in porphyrins in the weak coupling limit (small displacements of the excited states) that solely accounts for single mode contributions have been previously reported as (*225*):

$$\alpha_{\rho\sigma} = \sum_{l,m} \frac{M_{gl}^{\rho} H_{lm}^{r} M_{mg}^{\sigma} Q_{0,1}^{r}}{\left(\overline{E_{0,0}^{l,r,\lambda}} - \Omega_{\mathrm{L}} - i\Gamma_{0,0}^{l,r,\lambda}\right)\left(\mathrm{E}_{0,0}^{l,r,\lambda} - \mathrm{E}_{0,1}^{m,r,\lambda}\right)} + \sum_{l,m} \frac{M_{gl}^{\sigma} H_{lm}^{r} M_{mg}^{\rho} Q_{0,1}^{r}}{\left(\overline{E_{1,0}^{l,r,\lambda}} - \Omega_{\mathrm{L}} - i\Gamma_{1,0}^{l,r,\lambda}\right)\left(\mathrm{E}_{1,0}^{l,r,\lambda} - \mathrm{E}_{0,1}^{m,r,\lambda}\right)} \tag{33}$$

Where (resonance) Raman activity depends on the vibronic coupling matrix element $c_{lm}^{\Gamma_r}$:

$$c_{lm}^{\Gamma_r} = H_{lm}^{r} Q_{0,1}^{r} = \langle l|\partial H_{\mathrm{el}}(q,Q)/\partial Q_r|m\rangle\langle 1|\delta Q_r|0\rangle \tag{34}$$

Q and *q* are the nuclear and electronic coordinates, respectively and Γ_r is the symmetry of a given Raman active vibration in the D_{4h} point group. Since ideal D_{4h} symmetric porphyrins exhibit two-fold degenerate, E_u symmetry in the two lowest excited electronic states, B and Q (*219*), group theory dictates that the vibronic coupling operator $\partial H_{\mathrm{el}}(q,Q)/\partial Q_r$ must transform like the product representation $E_u \otimes E_u$ which can be decomposed into the one dimensional irreducible representations, A_{1g}, A_{2g}, B_{1g} and B_{2g}. Consequently, only A_{1g}, A_{2g}, B_{1g} and B_{2g} modes are resonance Raman active with B- and Q-band excitation. The modes denoted with a subscript g (*gerade*) correspond to deformations that preserve the inversion symmetry of the porphyrin. Specifically, A_{1g} modes are totally symmetric and preserve the overall symmetry of the molecule. B_{1g} and B_{2g} type vibrations induce an asymmetric in-plane deformation of the porphyrin and lower the symmetry from four to two-fold, i.e. $D_{4h} \rightarrow C_{2v}$. The lowest frequency modes corresponding to B_{1g} type vibrations induce a rhombic distortion along the N-M-N line and B_{2g} type vibrations distort the methine bridge connections line, inducing a triclinic type distortion (Figure 23).

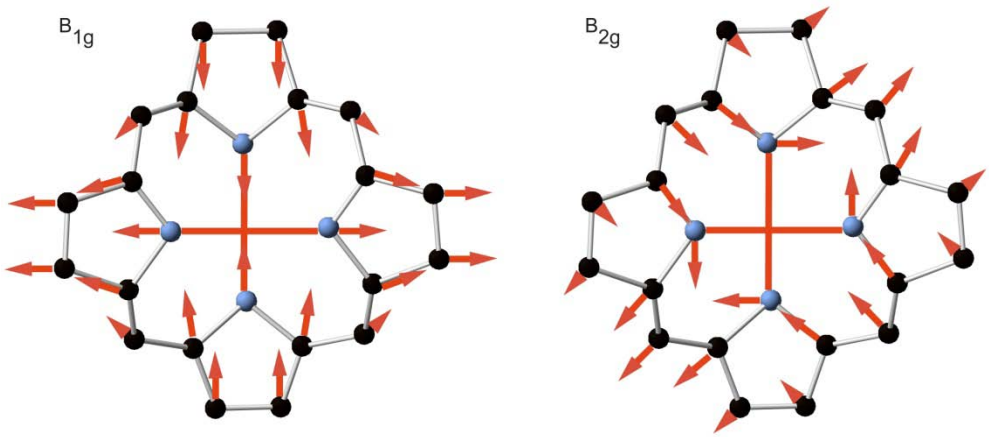


Figure 23: Normal mode compositions of B_{1g} and B_{2g} deformations represented by the respective modes of lowest frequency.

A_{2g} type vibrations induce an antisymmetric in-plane deformation of the macrocycle and can be considered as a combination of clockwise and anti-clockwise rotation, which preserves the four-fold symmetry but lowers it from $D_{4h} \rightarrow C_{4h}$. These symmetry modes are involved in different types of vibronic coupling.

A_{1g} modes may induce intrastate (Frank-Condon) and interstate (Herzberg-Teller) coupling which is described by the following vibronic coupling matrix elements:

$$c_{BB}^{A_{1g}} = c_{B_xB_x}^{A_{1g}} = c_{B_yB_y}^{A_{1g}}$$

$$c_{QQ}^{A_{1g}} = c_{Q_xQ_x}^{A_{1g}} = c_{Q_yQ_y}^{A_{1g}} \qquad (35)$$

$$c_{BQ}^{A_{1g}} = c_{QB}^{A_{1g}} = c_{B_xQ_x}^{A_{1g}} = c_{Q_yB_y}^{A_{1g}}$$

Within the framework of the four-orbital model, the parameters, c_{QQ} and c_{BB} depend on the mixing parameter, v, and the coupling parameter, c_{QB}.

B_{1g}-type vibronic coupling is interstate (HT) and intrastate Jahn-Teller coupling, where the latter is associated with distortions along the x and y coordinates with opposite sign so that the corresponding matrix elements are:

$$c_{BB}^{B_{1g}} = c_{B_xB_x}^{B_{1g}} = -c_{B_yB_y}^{B_{1g}}$$

$$c_{QQ}^{B_{1g}} = c_{Q_xQ_x}^{B_{1g}} = -c_{Q_yQ_y}^{B_{1g}} \qquad (36)$$

$$c_{QB}^{B_{1g}} = c_{QB}^{B_{1g}} = c_{Q_xB_x}^{B_{1g}} = -c_{Q_yB_y}^{B_{1g}}$$

Vibronic coupling due to B_{2g}-type coupling is interstate Jahn-Teller and Herzberg-Teller coupling and the corresponding matrix elements are:

$$c_{BB}^{B_{2g}} = c_{B_xB_y}^{B_{2g}} = c_{B_yB_x}^{B_{2g}}$$

$$c_{QQ}^{B_{2g}} = c_{Q_xQ_y}^{B_{2g}} = c_{Q_yQ_x}^{B_{2g}} \qquad (37)$$

$$c_{QB}^{B_{2g}} = c_{QB}^{B_{2g}} = c_{Q_xB_y}^{B_{2g}} = c_{Q_yB_x}^{B_{2g}}$$

In the framework of the four-orbital model for B_{1g} and B_{2g} type coupling $c_{BB} = -c_{QQ}$.

A_{2g}-type coupling is solely due to interstate HT coupling and the matrix elements are:

$$c_{QB}^{A_{2g}} = -c_{QB}^{A_{2g}} = c_{Q_xB_y}^{A_{2g}} = -c_{Q_yB_x}^{A_{2g}} \qquad (38)$$

The Raman tensors for these modes are written in terms of McClain tensor representations:

$$
\begin{aligned}
T(l,m)^{A_{1g}} &= \begin{pmatrix} c_{lm}^{A_{1g}} & 0 \\ 0 & c_{lm}^{A_{1g}} \end{pmatrix} \\
T(l,m)^{B_{1g}} &= \begin{pmatrix} c_{lm}^{B_{1g}} & 0 \\ 0 & -c_{lm}^{B_{1g}} \end{pmatrix} \\
T(l,m)^{B_{2g}} &= \begin{pmatrix} 0 & c_{lm}^{B_{2g}} \\ c_{lm}^{B_{2g}} & 0 \end{pmatrix} \\
T(l,m)^{A_{2g}} &= \begin{pmatrix} 0 & c_{lm}^{A_{2g}} \\ -c_{lm}^{A_{2g}} & 0 \end{pmatrix}
\end{aligned}
\qquad (39)
$$

Electronic transitions polarized perpendicular to the heme plane are of no relevance for Raman scattering at porphyrin modes in eq. 39, so all z-components in the Raman tensors can be omitted. The complete first-order contribution to the Raman tensor is written as a sum of McClain tensors multiplied by the appropriate frequency denominator and transition dipole strength. In the most general case where the porphyrin is in a low symmetry, admixtures of A_{1g}, B_{1g}, B_{2g} and A_{2g} tensors occur. Consequently, a linear combination of the aforementioned individual tensors describe the Raman tensor of the representation Γ_r (*226*):

$$
\hat{\alpha} = \begin{bmatrix} a_{1g}(\tilde{\nu}_L) + b_{1g}(\tilde{\nu}_L) & b_{2g}(\tilde{\nu}_L) + a_{2g}(\tilde{\nu}_L) \\ b_{1g}(\tilde{\nu}_L) - a_{2g}(\tilde{\nu}_L) & a_{1g}(\tilde{\nu}_L) - b_{1g}(\tilde{\nu}_L) \end{bmatrix}
\qquad (40)
$$

Depolarization ratios can be determined from eq. 40 by using invariants of the isotropic, anisotropic and antisymmetric part of the Raman tensor (*226*):

$$\rho = \frac{3}{4} \cdot \frac{a_{1g}^2(\tilde{\nu}_L) + 5a_{2g}^2(\tilde{\nu}_L) + 2[b_{1g}^2(\tilde{\nu}_L) + b_{2g}^2(\tilde{\nu}_L)]}{6a_{1g}^2(\tilde{\nu}_L) + 2[b_{1g}^2(\tilde{\nu}_L) + b_{2g}^2(\tilde{\nu}_L)]}$$

(41)

Thus, the depolarization ratios of the following modes in an unperturbed D_{4h} symmetry according to eq. 41 are as follows: A_{1g}=0.125, B_{1g}=0.75, B_{2g}=0.75 and A_{2g}=∞. Experimentally the depolarization ratios can be measured as $\rho = I_y/I_x$, where I_y is the scattered intensity perpendicular and I_x is the scattered intensity parallel to the scattering plane for a given Raman band.

2.4 Static Normal Coordinate Deformations

Heme-protein interactions and substituents on the periphery of the macrocycle give rise to symmetry-lowering deformations. Any symmetry-lowering deformation, ∂Q_T , of the macrocycle can be described as a superposition of static normal coordinate deformations (SNCD), $\partial \bar{Q}_i^{\Gamma_i}$ (*227*):

$$\partial Q_T = \sum_{\Gamma_i, i} \gamma_i^{\Gamma_i} \partial \bar{Q}_i^{\Gamma_i} \qquad (42)$$

where $\gamma_i^{\Gamma_i}$ is the amplitude of the distortion along the normal coordinate of the *i*-th vibration exhibiting D_{4h} symmetry, Γ_i . The restoration energy is proportional to the square of the wavenumber, making the set of $\partial \bar{Q}_i^{\Gamma_i}$ dominated by deformations along the normal coordinates of lowest wavenumber modes of the respective symmetry representations (*227*).

Macrocycle distortions can affect vibronic coupling properties of Raman active modes. This affect can be accounted for by expanding the vibronic coupling operator, $\partial \hat{H}_{el}(q,Q)/\partial Q_r^{\Gamma_r'}$ of a Raman active porphyrin macrocycle vibration, $Q_r^{\Gamma_r}$, where Γ_r' is the representation of the lowered symmetry group. In first order, this expansion yields:

$$\frac{\partial \hat{H}_{el,0}(q,Q)}{\partial Q_r^{\Gamma_r'}} = \frac{\partial \hat{H}_{el,0}(q,Q)}{\partial Q_r^{\Gamma_r}} + \sum_{\Gamma_i}\sum_{i} \frac{\partial^2 \hat{H}_{el,0}(q,Q)}{\partial Q_r^{\Gamma_r}\partial Q_i^{\Gamma_i}} \partial \bar{Q}_i^{\Gamma_i} \qquad (43)$$

$\partial \bar{Q}_i^{\Gamma_i}$ is the *i*-th SNCD of symmetry Γ and $\hat{H}_{el,0}$ is the unperturbed electronic Hamiltonian of D_{4h} symmetry. In the framework of Goutermans four-orbital model, the vibronic coupling operator accounts for coupling of electronic states having E_u-symmetry. Thus, the representations Γ_r and $\Gamma_r \otimes \Gamma_i$ in the first and second terms of eq. 43 respectively must transform like $\Gamma_r' = A_{1g}$, B_{1g}, B_{2g} or A_{2g}. The first term on the right side of eq. 43 represents vibronic coupling of Raman active modes in unperturbed symmetry. The second term on the right side of eq. 43 accounts for symmetry lowering distortions. This term describes contributions from in-plane normal distortions as well as accounts for vibronic coupling of out-of-plane modes when out-of-plane distortions are operative because the product of their representations yields either A_{1g}, B_{1g}, B_{2g} or A_{2g}. The description of this effect will be detailed later.

Chapter 3 Methods of Identifying Conformational Changes and Heme Deformations of Cytochrome c

3.1 Conventional Methods

Perturbations of the heme group are caused through space by an electric field produced by polar and ionizable side chains of the surrounding protein (*212, 223, 228, 229*) and through bond by the covalent and non-covalent interactions between heme and protein environment. The perturbations induced through bond cause in-plane (ip) and out-of-plane (oop) deformations of the heme, which affect functional properties such as ligand binding affinities (*230*), electron-transfer rates (*208*) and redox potentials (*231*). Cytochrome c heme chromophores in crystal structures have been shown by Berghuis *et al. (183)* to exhibit larger non-planar heme deformations than other heme proteins such as myoglobin and hemoglobin, mostly due to the additional covalent bonds between heme and protein (*232*). A method used to quantify non-planar deformations was developed by Shelnutt, Jentzen and coworkers (*227, 231, 233, 234*). Their algorithm characterizes the structure of porphyrins in terms of the deformations along the normal coordinates of lowest wavenumber modes of the macrocycle, termed static normal coordinate deformations (SNCDs), which is the minimal basis set and is typically sufficient. This method is referred to as normal coordinate structural decomposition (NSD). NSD describes the non-planarity of porphyrins in terms of SNCDs classified by the irreducible representations, A_{1u}, A_{2u}, B_{1u}, B_{2u} and E_g, of the point group D_{4h}. The irreducible representations denote the following normal deformations; propellering (A_{1u}), doming (A_{2u}), ruffling (B_{1u}), saddling (B_{2u}) and waving (E_g) (Figure 24).

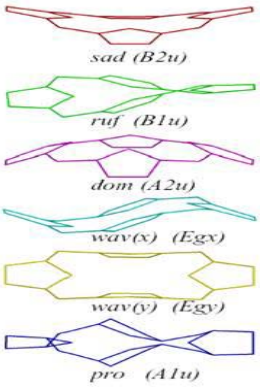


Figure 24: Taken from ref (*231*) and modified. Schematic representation of out-of-plane distortions of porphyrin macrocycles.

The lowest wavenumber normal coordinates of each type of symmetry, or minimal basis set, are needed to simulate distortions of the heme macrocycle. The six lowest wavenumber out-of-plane (oop) normal coordinates have been shown to be generally sufficient for characterizing non-planar heme deformations determined from crystal structures (*233*). This simplicity is due to the fact that the energy to be invested for a given SNCD is proportional to the square of the wavenumber of the respective normal mode. Hence, a linear combination of these six distortions can classify and quantify out-of-plane heme distortions.

Non-planar deformations of heme groups have been quantified by analyzing the crystal structures of the respective proteins, with a comprehensive list of proteins published by Shelnutt (*235*). However, crystal structures are not available for all relevant ligation, spin and redox states of a given heme protein so heme structures extracted from molecular dynamics (MD) simulations have accordingly been subject to NSD analysis (e.g. for horseradish peroxidase) (*236*). Moreover, the modification of heme deformations by the change of solution conditions (pH, ionic strength, viscosity) can

generally not be explored by X-ray crystallography, which is a particularly important issue for ferricytochrome c, for which the structure of the heme cavity is rather pH dependent (*237, 238*). Finally, it is not yet clear whether the X-ray structure, affected by crystal packing effects or crystallization conditions, really represents the dynamic average structure in solution. Most recently performed MD simulations as well as optical spectroscopy data for horse heart and yeast cytochrome c seem to argue to the contrary (*221*).

3.2 Resonance Raman Depolarization Ratios

Even though it is clear from elementary theory that non-planar deformations can make the oop modes of porphyrins and heme groups Raman active, they have not been frequently used for a quantitative analysis of heme deformations. Some recent resonance Raman studies on non-native states of cytochrome c (*239*) and nitric-oxide synthase (*240*) utilized the Raman intensities of some specific oop modes to qualitatively infer changes of heme planarity upon oleic acid and ligand binding, respectively. It has recently been shown for myoglobin cyanide (*241*), horseradish peroxidase (*242, 243*) and various nickel porphyrins (*244*) that polarized resonance Raman spectroscopy can be utilized to probe SNCDs by exploiting the fact that the Raman cross section of oop modes is proportional to the square of SNCDs (*245*). Earlier, it was demonstrated that the depolarization ratio dispersion of A_{1g}- and A_{2g}-modes can be employed to probe the combined presence of ruffling (B_{1u}) and saddling (B_{2u})

deformations (*225*) which can be inferred by the not yet introduced term in the third order expansion of the vibronic coupling operator, $\frac{\partial^3 \hat{H}_{el}}{\partial Q_r^{\Gamma_r} \partial Q_i^{\Gamma_i} \partial Q_j^{\Gamma_j}} Q_{01}^{\Gamma_r} \delta Q_i^{\Gamma_i} \delta Q_j^{\Gamma_j}$.

Deviations from the canonical depolarization ratios indicate contributions from other tensor elements in the right hand side of eq. 43. Planar and non-planar distortions of the heme chromophore can be inferred directly from the departure from canonical ρ values. For instance, when the depolarization ratio ($\rho = I_y/I_x$) of a B_{1g} mode is less than 0.75, it indicates an increased A_{1g} contribution to the Raman tensor. This contribution can be inferred from the second term on the right hand side of eq. 43 as follows. In order for this term to transform like A_{1g} the operative deformation must have B_{1g} symmetry so that the representation of the vibronic coupling operator reads as $B_{1g} \otimes B_{1g} = A_{1g}$. This line of logic translates as an increased rhombic, B_{1g} deformation. The same concept can be invoked to infer the influence of out-of-plane (oop) modes in the Raman tensor. For example, modes of B_{1u}-symmetry, which are inactive in D_{4h}, can become Raman active by a macrocycle deformation of the same type of symmetry so that $\Gamma_r^{'} = B_{1u} \otimes B_{1u} = A_{1g}$. The oop-modes become Franck-Condon active yielding polarized (ρ=0.125) Raman bands with B band excitation. The deviation from canonical depolarization ratios for oop modes indicates asymmetric or antisymmetric contributions to the Raman tensor. For example, we consider a depolarization ratio of 0.25 obtained for an A_{2u} mode of horse heart cytochrome c, to identify possible admixtures to the Raman tensor that would give rise to this result. If the macrocycle deformation transforms like B_{1u} (ruffling), the 1st order vibronic coupling operator in eq.

43 transforms like $A_{2u} \otimes B_{1u} = B_{2g}$. If admixed to the (dominant) totally symmetric contributions, $A_{2u} \otimes A_{2u} = A_{1g}$, the depolarization ratio increases. However, the same effect can be caused by saddling (B_{2u}) deformations, which would produce vibronic coupling of $A_{2u} \otimes B_{2u} = B_{1g}$ symmetry. The presence of dominant oop deformations can be verified with data from NSD analysis using crystal and MD structures. Upon comparison, it is apparent that the increase in the depolarization ratio is due to a dominant ruffling distortion in the heme. The structure of horse heart cytochrome c is well known from NMR and crystallographic data, though heme distortions of isoforms and site directed mutants require an abundance of time to classify with these techniques. Resonance Raman spectroscopy offers an alternative useful technique for these purposes as well as those previously discussed.

3.3 Absorption and Circular Dichroism Spectroscopy

Heme protein interactions have been explored by utilizing the combination of visible circular dichroism and absorption spectroscopy as well. We focus, in the current studies, on the interpretation of the CD and absorption spectra of cytochrome c. Not surprisingly, due to the necessity to apply a complex theoretical approach, a limited attempt to understand the characteristics of experimental CD spectra of heme proteins have been made in the past. Optical Rotary Dispersion (ORD) was used prior to CD as a tool to monitor heme protein interactions. For the Soret band Ulmer first reported a positive cotton effect for ferro- and a negative effect for ferricytochrome c (*246*). Urry and Doty later examined the ORD spectra upon denaturation of the protein (*247*). Myer then reported ORD of pH, temperature and ligand dependence on the conformation of

cytochrome c (*248*). Drucker et al. eventually measured the B and Q-band CD spectra of this protein. The authors reported marked differences between the spectra of the two different oxidation states (*249*). To that end, no physical rationale was ever offered for the results. The first theoretical attempt to understand the visible CD spectra of heme proteins was made by Woody and coworkers. They rationalized the observed Cotton effect from myoglobin and hemoglobin Soret band CD spectra in terms of a coupled oscillator model involving $\pi \rightarrow \pi^*$ transition dipole coupling of the heme and aromatic side chains from the nearby protein matrix (*250*). Another, more recent study from Woody and coworkers used the heme undecapeptide of cytochrome c, namely microperoxidase-11, to model the Soret band CD spectra based on a structure obtained from MD simulations (*251*). They found that the positive Cotton signal was due significantly to an intrinsic heme contribution, which arose from out-of-plane distortions of the latter. The authors, in fact, also identified contributions from the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions of the protein matrix. Even more recently, it was reported that out-of-plane distortions such as ruffling of the heme can correlate to the observed rotational strength in the Soret band CD spectra (*252*).

Schweitzer-Stenner combined CD and absorption spectra to provide a clearer picture of the effects of the internal electric field and its gradient on CD spectra (*212*). He was able to consistently model these spectra for three different cytochrome c isoforms for both oxidation states for the B and Q-bands by invoking electronic and vibronic contributions. The obtained splitting of the B-band was interpreted as a contribution from the internal electric field from a quadratic Stark effect, whereas that of the Q-band was assigned to

the gradient of the latter. In this detailed analysis, the author was able to quantify the electric field strength for oxidized and reduced states and concluded that the oxidized state internal electric field was enthalpically stabilized compared to its reduced counterpart.

3.4 Synchrotron Radiation Ultra Violet Circular Dichroism

For the past 50 years, far UV circular dichroism spectroscopy has been established as a diagnostic tool to determine the structure and dynamics of proteins in solution (*253*). Since CD is diagnostic of chirality, it is a particularly suitable tool to explore the secondary structure of proteins (*254*). Secondary structures of proteins such as α-helices, β-sheets, polyproline II, turns, etc., exhibit characteristic backbone dihedral angles (ϕ, ψ), which are indicated in the Ramachandran plot shown in Figure 25.

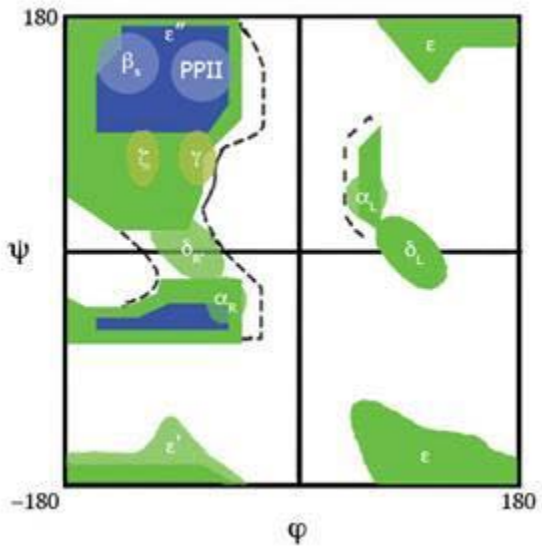


Figure 25: Taken from ref (*210*) and modified. Representative Ramachandran plot showing sterically allowed regions.

The amide backbone $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ electronic transitions occur in the ultra-violet (UV) regime. These transitions give rise to circular dichroism, owing to excitonic coupling between $\pi \rightarrow \pi^*$, as well as $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions in different peptide groups of a polypeptide chain (*255*). Polarization effects have recently been shown to contribute as well (*256*). This coupling depends on the dihedral angles so that different secondary structures yield significantly different CD spectra (*254*). Proteins containing large fractions of α-helices, such as myoglobin, yield a UV-CD spectrum with minima at 222 nm and 208 nm and an intense maximum at 192 nm (Figure 26). Proteins such as Concavlin A, which contains a substantial amount of β-sheet, yield a UV-CD spectrum with a minimum at 210 nm and a maximum at 190 nm, the relative intensities of which are less than those from α-helices (Figure 26). Collagen, a protein with substantial fractions of poly-proline II helix yields a UV-CD spectrum with opposite characteristics of α-helices and β-sheets in that it shows a positive maximum at ~220nm and a negative one at ~200nm.

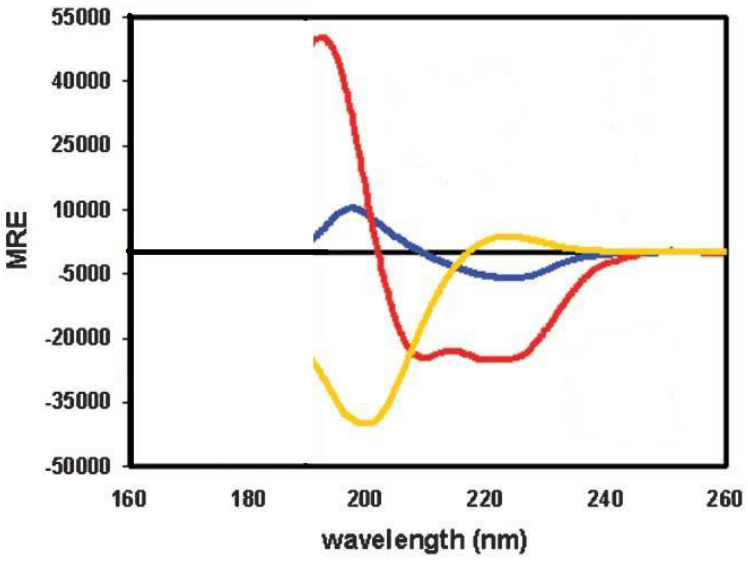


Figure 26: Taken from ref (*257*) and modified. Representative UVCD spectra of myoglobin (red) concavlin A (blue) and collagen (yellow) structures taken from 190-260nm.

The amount of α-helix present in a given protein can be roughly estimated by comparing the ellipticity at 222 nm (*258*). Secondary structure prediction software, such as the neural network analysis (K2D) (*259-261*), rigid regression (Contin program) (*262*), variable selection (VARSLC program) (*263*) and the self-consistent method (SELCON program) (*254, 264, 265*) have been developed for a more quantitative analysis of secondary structure composition, though the results varied (*266*). Spectra obtained from conventional instruments are limited in this regard, due to the confines of the Xenon arc lamp light source, which significantly decreases in intensity below 190 nm.

Though being developed almost 30 years ago, (*267, 268*) only over the last 10 years (*257*) synchrotron radiation has been used to measure vacuum ultra-violet (VUV) CD spectra, due to the fact that this light source is still very intense at wavelengths below 190nm and thus allows for higher energy transitions such as $n \rightarrow \sigma^*$ and $n' \rightarrow \pi^*$ to be

probed. Valuable spectral information is gained by extending the recording of CD spectra into the VUV (< 190nm) region (Figure 27).

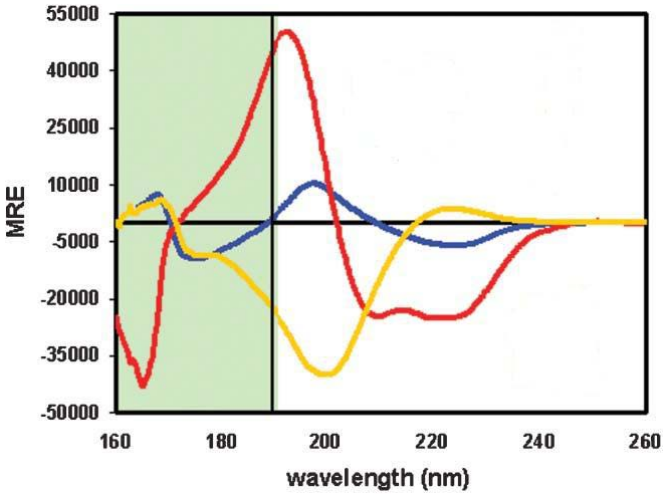


Figure 27: Taken from ref (*257*) and modified. Representative SRCD spectra from myoglobin (red) concavlin A (blue) and collagen (yellow) structures from 160-260nm.

For instance, spectra from α-helices show an apparent shoulder on the positive maximum, located at 180nm, which is difficult to measure properly using conventional instruments (Figure 27). α-helices and β-sheets show signals with opposite rotational strengths in the region below 170nm as well. The characteristics of CD spectra observed in the VUV enhance the capability of a quantitative secondary structure analysis using prediction software. Most recently, Wallace and coworkers developed a comprehensive secondary structure analysis program known as Dichroweb (*269, 270*), which is available online. Dichroweb allows the use of many of the most popular structural determination algorithms and several reference sets encompassing all ranges of wavelengths. This online server allows the user to input a variety of different (a) file formats from various instrument manufacturers, (b) units and (c) wavelength scales, with matching outputs, making it very user friendly. Along with secondary structure fractions, the universal

goodness of fit parameters is provided with the output file, which allows a judgment of their reliability. To appropriately quantify secondary structure using UVCD, exact protein concentrations and pathlengths of cells must be known, and the spectra must be background corrected for solvent effects. The reader is herewith referred to the work of Greenfield, and Wallace and associates (*257, 258, 269, 270*), which include more details about Dichroweb and analyses using UVCD for secondary structure determination.

Chapter 4 Materials and Methods

4.1 Non-native States of Ferricytochrome c Project

4.1.1 Materials

Cytochrome c from bovine heart and horse heart was purchased from Sigma-Aldrich (St. Louis, MO) and dissolved in (a) a 0.1 M Tris·HCl (Trizma) buffer (pH 7.2, I=0.1 M), from Sigma-Aldrich (St. Louis, MO) and (b) in a 1 mM MOPS (3-(N-Morpholino) propanesulfonic acid) buffer (pH 7 and 8.5, I =0.001M), from Sigma-Aldrich (St. Louis, MO). An alkaline solution was obtained with a 50mM Bis/Tris buffer (pH9.6) titrated with small amounts of 0.1M NaOH solution to achieve pH values of 10.5 and 11.4. Complete oxidation was achieved by adding a small amount of potassium ferricyanide obtained from Fisher Scientific (Pittsburgh, PA) to the sample. Subsequently, the excess oxidant was removed from the sample by dialysis against the respective buffer solution.

4.1.2 Visible and Far-UV CD and Absorption Measurements

We prepared 50 µM protein solutions for visible CD and absorption measurements. The B band spectra were measured as a function of temperature in digital form with a J-810 spectrapolarimeter (Jasco Inc.) purged with N_2, in the spectral region from 350-550nm. Far-UV CD spectra were measured in the spectral region from 185-240nm with a protein concentration of 0.5mM in Tris HCl buffer, measured with the same instrument and the following parameters. The temperature was controlled by a Peltier heating/cooling system from 5^{o}–90^{o} C (±1.0 ^{o}C) (278-363K). Between ten and twenty accumulations were averaged using a 5nm band width, a 500 nm/min scanning speed, and a 0.1 or 0.5nm data pitch. Additionally, a background subtraction was carried out for all the

spectra using similar parameters. A 1.0 mm quartz cell (Helma) was used for all measurements. The pH was measured before and after each measurement.

Synchrotron radiation circular dichroism (SRCD) measurements were performed at Brookhaven National Laboratories. Beamline U11 was used to measure the SRCD spectra in the spectral region from 170-260nm of horse and bovine heart cytochrome c. The proteins were dissolved in 0.1M TrisHCl buffer at a concentration of 6mg/ml (~0.5mM). Three accumulations were averaged at each indicated temperature, controlled by a peltier heating/cooling system, with similar parameters used for the background subtraction.

4.2 Out-of-plane Deformations of Ferrocytochrome c Mutants Project

4.2.1 Sample Preparation

Horse heart cytochrome c (hhc) and chicken cytochrome c (chc) were purchased from Sigma Aldrich Co. (St. Louis, MO) with minimum 99%purity, and were used without further purification. Yeast C102T cytochrome c (yc) was prepared by expression of plasmid-borne mutant yeast iso-1-cytochrome c (cyc-1) genes in the host *S. cerevisae* GM3C2 (*271*), which lacks the wild type gene. Yeast C102T cytochrome c was purified after yeast cell lysis and extraction or precipitation of most other soluble components by cation-exchange chromatography (*272*). For resonance Raman experiments, all cytochromes c were dissolved in a 0.01M tris buffered saline solution (pH 8) to final concentrations of 0.5mM (horse heart and chicken) and 0.673mM (yeast) and were then reduced with a small amount of sodium dithionite obtained from Fluka. We used the

934 cm^{-1} line of 0.5 M sodium perchlorate as internal intensity band wavenumber standard for resonance Raman experiments.

4.2.2 Polarized Resonance Raman Spectroscopy

We used a Renishaw Ramascope, a confocal Raman microscope, equipped with a CCD camera and 2400l/mm grating to record Raman scattering parallel (I_x) and perpendicular (I_y) polarized to the polarization of the excitation laser beam in the wavenumber range <1000cm^{-1}. We used a 441.6nm excitation wavelength from a Kimmon (Englewood, CO) IK series He-Cd laser. The laser was focused onto the sample with a 50X microscope objective. The power at the sample was 5mW. A 50µm slit width was used to focus the scattered light into the spectrometer. Multiple scans were averaged to increase the signal to noise ratio. We used the spectral decomposition program MULTIFIT (*135*) to complete the spectral analysis. Depolarization ratios (ρ) were calculated as the ratio of the intensity of the perpendicular (I_y) polarization and the intensity of the parallel (I_x) polarization, i.e. $\rho=I_y/I_x$.

4.2.3 Absorption Spectroscopy

We used a Jasco J810 spectrapolarimeter to measure the Soret band absorption and subtract the background of the three proteins. We used a final concentration of 0.5mM in 0.01M tris buffered saline solution (pH 8) with a 0.1mm pathlength, 500nm/min scan speed, 0.5nm data pitch, and 10 accumulations. Room temperature variations (20.0^{o}C) were controlled by a Peltier temperature controller ($\pm 1^{o}$C).

4.2.4 NSD Calculations

The non-planar deformations of the heme groups in the crystal structure of horse heart (1hrc)(*181*) and yeast isoenzyme-1 (1ycc) (*180*) cytochromes c have already been published by Shelnutt and co-workers (*227, 231*). For comparative purposes, we also performed NSD calculations on the horse heart, yeast, and chicken hemes extracted from MD simulations, as fully described elsewhere (*221*), using the average structure extracted from the production runs.

Chapter 5 Results and Discussions

5.1 Using Absorption and CD Spectroscopy to Reconcile Issues which Prevent a Detailed Structural Characterization of Cytochrome c[1]

Some of the questions and uncertainties that have arisen, such as the necessity for a better understanding of CD spectra of heme proteins and the affect of varying solution conditions on the structure of cytochrome c have already been assessed in our group. In what follows, a brief overview of the reported results will be presented. The specific questions that were addressed lead, in part, to the structural characterization of native and partially unfolded states of ferricytochrome c.

Dragomir et al. reported the Soret and Q-band CD and absorption spectra for oxidized and reduced horse, bovine, chicken and yeast cytochrome c isoforms (*222*). The observed Soret band couplet unambiguously revealed splitting of the excited B-state. The center of the couplet is situated right at the frequency position of the 0→0 B band transition. The positive and negative components are thus assigned to different components of the excited B-state. An empirical simulation of the CD spectra determined that this splitting was around 300 cm^{-1} for oxidized horse heart cyt c. The magnitude of Soret band Kuhn anisotropy plots ($\Delta\varepsilon/\varepsilon$) (*273*) corroborates this band's vibronic origin. The observed Soret band CD spectra for the reduced state indicate that the splitting of the excited B-state is substantially reduced (~20 cm^{-1}). CD spectra of the Q-band resolve the vibronic structure of the Q_v band, which was only previously attainable with cryogenic absorption measurements (*209, 221*). This resolution allowed for an assignment of the observed positive components of the vibronic side band.

[1] This section encompasses the results for studies to which I contributed less than 50%. Hence, the details will be briefly explained.

Based on the frequency shifts compared to the Q_0 position these components were attributed to A_{2g} modes ν_{26}, ν_{22} and ν_{19}. Also, based on a simulation of the Q-band and comparison to low temperature absorption measurements, the authors concluded that CD was a useful tool to estimate the excited Q-state splitting, even at room temperature.

Most importantly, Dragomir et al. observed an asymmetry of the 695nm charge transfer band for horse and yeast ferricytochrome c. They self consistently decomposed the 695nm CD and absorption bands into sub-bands carrying the same halfwidths and frequency positions. The 695nm CT band was decomposed into a minimum of three subbands at neutral pH which led Dragomir et al. to identify the existence of at least three conformational substates, which can be seen in Figure 28.

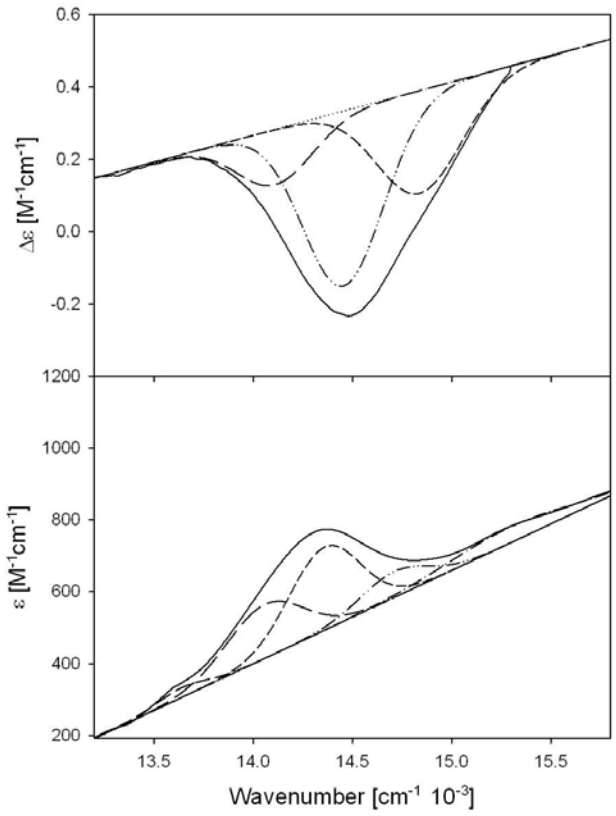


Figure 28: Taken from ref (*210*) and modified. Decomposition of the absorption (bottom) and CD (top) band profile of the 695 nm transition of bovine ferricytochrome c into three subbands. The spectra were measured at pH 7 with a 0.1 M Tris HCl buffer.

These substates have been previously classified in myoglobin and have been used to probe the dynamics of ligand rebinding. Conformational substates occur in local free energy minima of the global free energy minimum conformation of the protein and can be classified by small structural variation which may correspond to thermodynamic or kinetic distinction from other substates. Conformational substates are defined as performing the same function with different rates yet being thermodynamically distinct. A detailed study of the 695nm CD band by Cupane and coworkers confirmed the existence of at least three subbands based on the fine structure of the band at cryogenic temperatures (*274*). These authors speculate that these substates in cytochrome c

perform specific functions in biological processes. In cytochrome c, structural heterogeneity of, for example, the distal Fe-M80 ligation is likely to give rise to the observed subbands. Based on the influence that the Fe-M80 linkage has on the redox potential of the protein, Spilotros et al. believes that these conformational substates may become functionally relevant in the myriad processes of which cytochrome c is a part (*274*).

Subsequently, this CT band was further used to probe ferricytochrome c unfolding to detect whether the detected subbands arose from structural, opposed to electronic heterogeneity. Schweitzer-Stenner et al. investigated the changes of the 695nm absorption band as a function of increasing temperature (*275*). The authors discovered that the decrease in intensity of the respective band profiles is wavenumber dependent and biphasic seen in Figure 29.

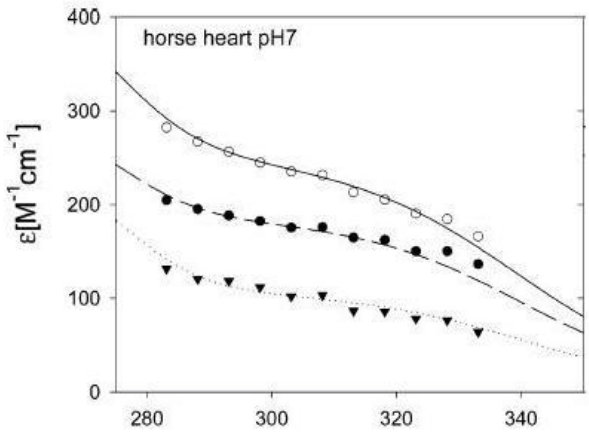


Figure 29: Taken from ref (*275*) and modified. Molar Extinction coefficients, *ε*, at 14 370 (open circles), 14 000 (closed circles), and 14 700 (triangles) cm-1 as a function of temperature obtained from absorption spectra taken at the indicated pH.

The former observation is an indication that the aforementioned conformational substates are thermodynamically distinct. The biphasic decrease in intensity as a function of temperature indicates that the thermal unfolding from room temperature of cytochrome c involves an intermediate state. The authors also found the existence of another state, which is populated at low temperatures (<263 K) and slightly acidic pH values.

To test whether there is an influence of ion binding to the surface of cytochrome c, Shah and Schweitzer-Stenner probed the changes of the Soret and 695nm CT band as a function of dihydrogen phosphate and sodium acetate concentration (*276*). The authors found that an ionic strength increase (via an increase in sodium acetate concentration) *per se* did not effect the heme environment, whereas anion binding induced a stabilization of the Fe-M80 linkage. This stabilization was inferred from the increase in oscillator strength of the 695nm band upon increasing the $H_2PO_4^-$ concentration. The ramifications for this finding include the necessity to reassess the affect of anion binding on the function of cytochrome c.

The above mentioned results led to a reinvestigation of the 695nm CT band's disappearance at alkaline pH which was discovered by Schejter and coworkers (*196, 277*). The discovery of Shah and Schweitzer-Stenner pointed to the necessity to check whether the ionic strength (anion binding) conditions influence the disappearance of the 695nm band. Verbaro et al. performed a titration, with the 695 nm band as a spectral marker of the Fe-M80 linkage, from pH 7-10.5 and plotted the decrease in

effective oscillator strength (Figure 30) of two of the aforementioned subbands at high (50mM) and low (0.5mM) phosphate ion concentrations (*278*).

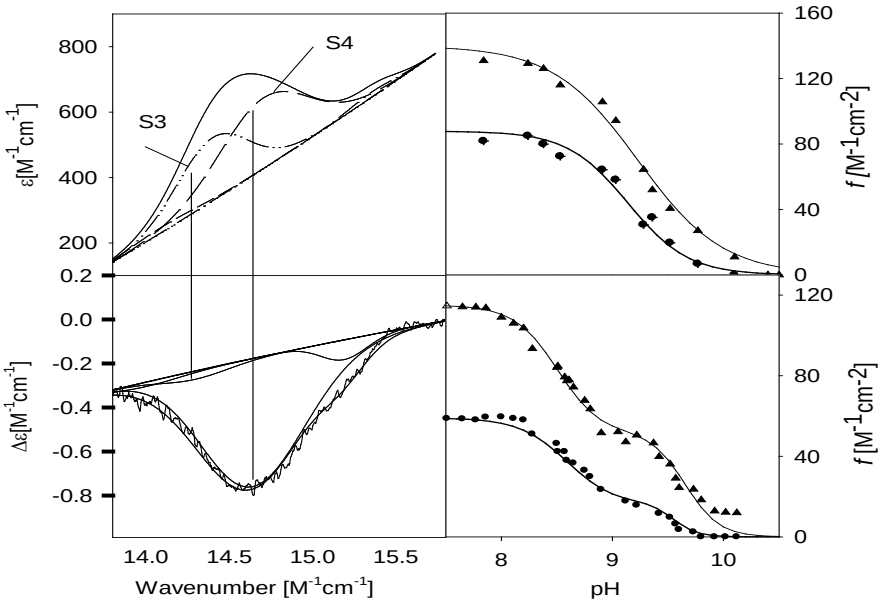


Figure 30: Taken from ref (*278*) and modified. Left Column: Absorption and CD band profiles of the 695 nm band of horse heart ferri-cytochrome c measured at pH 7 with a 50 mM phosphate buffer. The profiles were decomposed with a procedure described in Materials and Methods. Right Column: Integrated (absorption) intensities of the sub-bands S3 (▲) and S4 (●) as a function of pH measured with a 50 mM (upper panel) and a 0.5mM phosphate buffer (lower panel).

The observed biphasic titration curve at low ionic strength was rationalized by attributing a population of an intermediate of the alkaline transition, termed state III*, before the replacement of M80 by a lysine in state IV. The apparent pK-value of this transition is 8.5. The population of this intermediate is indeed effected by and ultimately undetectable by this method as ion binding increases. It is noteworthy that a thermodynamic intermediate of the alkaline transition, termed state III.5, was discovered by the Romesberg group (*279*). This intermediate was discovered at high ionic strength and it was reported that population of this intermediate state is

characterized by a M80 misligation, so that the 695nm CT band would disappear. The intermediate state III* is characterized by a weakening of the 695nm CT band, which emphasizes the structural difference of these two alkaline transition intermediate states.

5.2 Non-Native Alkaline States of Ferricytochrome c Probed by CD and Absorption Spectroscopy[2]

In view of the contradictions that arose in characterizing the thermally populated and alkaline states of ferricytochrome c, specifically, the results of Taler et al. (*186*), it is desirable to reinvestigate these states under well defined conditions. Also, despite numerous attempts to thermodynamically and electrochemically characterize the transition from native to alkaline states (*186, 191, 192, 194, 202*), information about the structural properties of these states is limited. Concomitant changes in the secondary structure have been qualitatively estimated using UVCD (*248*), but a quantitative assessment of these changes was never reported.

This section is organized as follows. First, we introduce a protocol which enables us to unambiguously discriminate, thermodynamically, between the states III_h and IV. Secondly, we detail temperature dependent far-UV CD spectra to elucidate the secondary structure of III_h. Next, we analyze representative CD and absorption spectra by means of a recently developed algorithm to estimate the splitting of the B-band, owing to electronic and vibronic perturbations (*211*). Finally, we characterize differences between alkaline states as well as thermally populated intermediate states.

[2] The results presented hereforth are from studies to which I contributed more than 50%

In what follows we do not discriminate between the substates IV_1 and IV_2 and V_1 and V_2 discovered by Döpner et al. (*192*).

5.2.1 Thermal Unfolding of the Native State III

To separate the thermal unfolding of the native state from the alkaline III→IV transition, we had to select experimental conditions which allow the protein to sample the native state at ambient temperatures and avoid a detectable population of the alkaline state at high temperatures. To this end, we used a 0.1 M Tris HCl buffer to ensure that the aforementioned requirements were met. The pH value of this buffer decreases by 0.03 pH units per ^{o}C increase of temperature. The pH of this buffer was 7.6 at the lowest temperature (278 K). With rising temperature, the pH decreased to reach 6.3 at 333 K and 5.6 at 363 K. This downshift of the pH value is comparable with the temperature dependence of the effective pK value of the III→IV transition, for which Battistuzzi et al. obtained 9.4 at 278 K and 7.9 at 340 K for the ionic strength of 0.1 M used in our experiment (*191, 202*). Thus, we kept the effective difference between the actual pH value of our experiment and the pK value of the III→IV transition approximately constant, so that a detectable occupation of state IV at high temperatures can be ruled out. Figure 31 exhibits the corresponding CD and absorption spectra of bhc and hhc in the aforementioned buffer measured as function of temperature between 278K and 363K in increments of 5K.

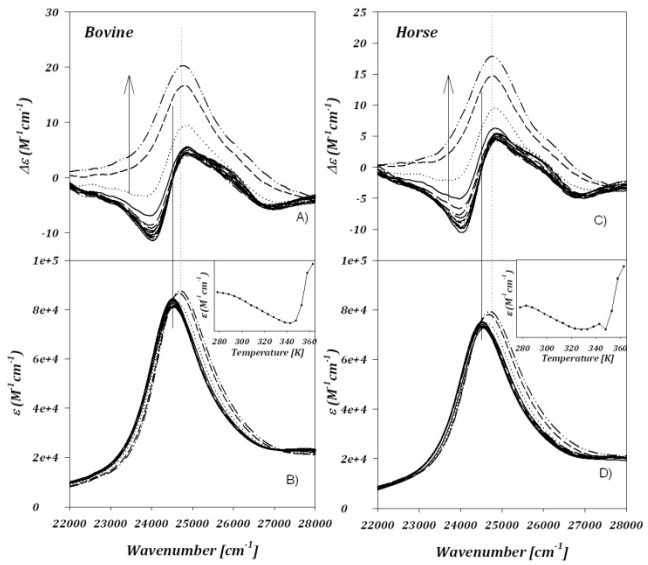


Figure 31: Taken from ref (*280*) and modified. Visible circular dichroism (A and C) and absorption (B and D) spectra of bovine (left column) and horse heart ferricytochrome c (right column) taken at pH 7 in a Tris HCl buffer between 278 and 363 K in increments of 5 K. The arrows in these and all subsequent figures indicate the spectral changes with increasing temperature. The solid vertical line indicates maximum absorption frequency position at low temperature and dotted line indicates maximum absorption frequency at high temperature. Inset: ε vs temperature at the maximum absorption frequency. Spectra taken at temperatures between 278 and 348 K are plotted as solid lines, spectra taken at 353, 358 and 363 K are plotted as dotted, dashed, dashed-dotted line.

The rather high ionic strength, which ensures the binding of 2 (3) Cl^- ions to the surface of bhc (hhc) (*202*), resembles the conditions frequently used for exploring the temperature dependence of ferricytochrome c (*184, 196, 197, 281*). The CD spectra display a pronounced, slightly negatively biased couplet at room temperature, which gradually changes into a more symmetric couplet with increasing temperature between 278 and 338K. The insets of Figure 31 (B and D) display the respective extinction value, ε_{max}, at the peak position as a function of temperature. ε_{max} decreases with rising temperatures between 278 and 340 K and subsequently increases between 333 and 363K. This second phase additionally involves a blueshift of the absorption spectrum. At higher temperatures the CD couplet disappears and is replaced by a slightly blueshifted, positive Cotton band. The temperature dependence of the Δε values at the

maximum ($\Delta\varepsilon_{max}$), central ($\Delta\varepsilon_{cent}$) and minimum ($\Delta\varepsilon_{min}$) position of the couplet are displayed in Figure 32.

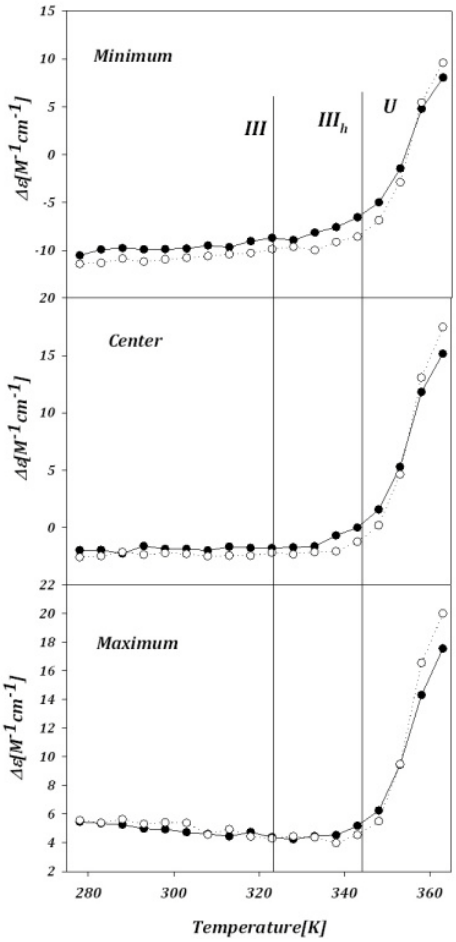


Figure 32: Taken from ref (*280*) and modified. Δε vs temperature of bovine heart (open circles) and horse heart (closed circles) ferricytochrome c between 278 and 363K measured in pH7 Tris HCl buffer. The upper panel is at the 24010cm^{-1}, the middle is at 24450cm^{-1}, and the lower is 24876 cm^{-1}, corresponding to the minimum, the center and the maximum position of the B-band couplet at room temperature. The temperature regimes of the three states III, III$_h$ and U (unfolded) are indicated.

As expected, corresponding values of hhc and bhc are very similar, but hhc systematically exhibits slightly more negative values at the minimum and the central position of the couplet between 278 and 343K. Interestingly, the three Δε values show different temperature dependences. $\Delta\varepsilon_{min}$ increases and $\Delta\varepsilon_{max}$ decreases with rising temperature below 343K, which yields the more symmetric couplet at higher

temperatures. All values sharply increase above 343K, reflecting the cooperativity of the unfolding process. Thus, our data suggest a similar melting temperature for both proteins, but the slope of Δε(T) is clearly steeper for hhc, indicating a higher degree of cooperativity. This result seems to be at variance with results from recent IR spectroscopic investigations of the thermal unfolding of horse heart and bovine ferricytochrome c, which suggest different unfolding mechanisms for the two proteins (*184*). For horse heart cytochrome c, Filosa et al. proposed a sequential unfolding of protein domains (in agreement with work from the Englander group (*282*)), whereas the data for bovine cytochrome c were interpreted as indicating a more global, cooperative unfolding (*184, 197, 281*). However, our data likely reflect the specific secondary and tertiary structure of the heme environment rather than global changes of the protein.

The data in Figure 31 and Figure 32 indicate only modest spectral (and thus also structural) changes in the temperature region between 320 and 340 K, with which the population of the thermal intermediate, III_h, is generally associated. However, the temperature dependencies of ε and Δε values exhibited in Figure 31 and Figure 32, respectively, clearly indicate the population of an intermediate in the temperature region where III_h is generally populated. To check whether this state really resembles the well known characteristics of the III_h state, we also measured the temperature dependence of absorption and CD profile of the charge transfer band at 695 nm between 278 and 348K, which are exhibited in Figure 33.

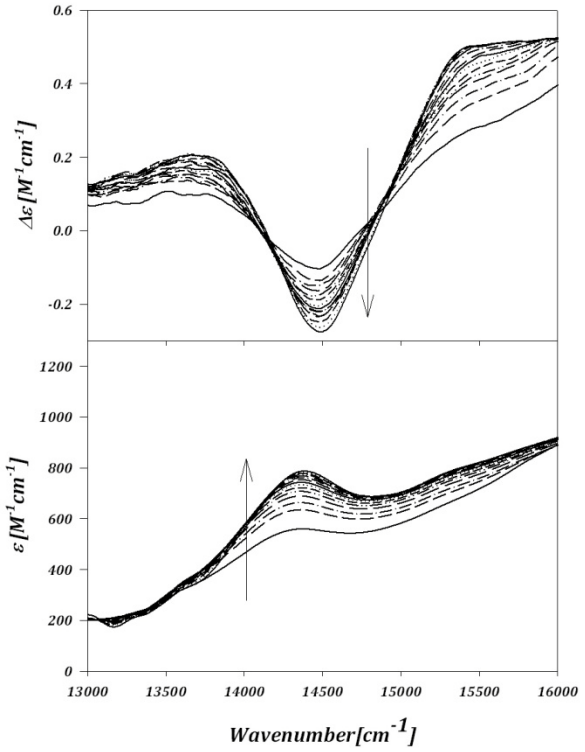


Figure 33: Taken from ref (*280*) and modified. Temperature-dependent CD (top) and absorption (bottom) spectra of bovine heart cytochrome c measured in the charge-transfer band region from 13000-16000cm^{-1}. The spectra were measured in a pH7 Tris HCl buffer with a 5mM concentration.

The data clearly reveal the decrease of this band's intensity above 310 K (Figure 33 inset) as first observed by Schejter and George (*196*), which is generally viewed as characteristic for a III→III$_h$ transition. Altogether, these findings suggest III$_h$ is thermodynamically distinct from state IV, since the latter cannot be significantly populated at the chosen experimental conditions.

Most of the investigations of the III→III$_h$ transition have thus far been performed at neutral pH with a temperature-dependent buffer (*184, 186, 197, 281*). The question arises whether state III$_h$ or IV is populated at higher temperatures, if a temperature-independent buffer with a low ionic strength is used. To address this issue, we measured the CD and absorption spectra of the B-band in a buffer that remained at

(nearly) constant pH as a function of temperature. This experiment was achieved by using a 1 mM MOPS (3-(N-morpholino)propanesulfonic acid) buffer. This buffer has a temperature coefficient of $-1.3 \cdot 10^{-3}$ pH units per ^{o}C. The low ionic strength of the corresponding solution ensured a rather low effective pK-value of the alkaline transition (~9 at room temperature and 8 in III_h region). The corresponding spectra are shown in Figure 34.

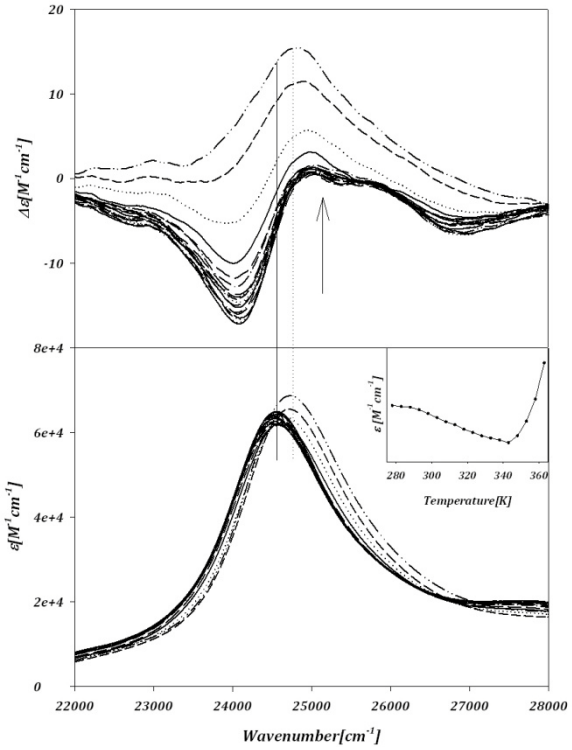


Figure 34: Taken from ref (*280*) and modified. Temperature dependant CD (top) and absorption (bottom) spectra of bovine heart cytochrome c measured in the Soret band region from 22000-28000cm^{-1}. The spectra were measured in a pH7 MOPS buffer with a 0.05mM concentration. Arrow indicates changes as a function of temperature. Solid line indicates maximum absorption frequency position at low temperature and dotted line indicates maximum absorption frequency at high temperature. Inset: ε vs temperature at the frequency of maximum absorption.

Apparently, the CD spectrum and its temperature dependence are qualitatively very similar to what we obtained by using a Tris HCl buffer to decrease the pH while increasing the temperature. This notion is corroborated by the respective Δε(T) plots in

Figure 35. Quantitatively, the two sets of spectra (Figure 31, left side and Figure 34) differ slightly, which can be attributed to the anion concentrations used for the two experiments (*276*). Our data thus far indicate that the non-native states populated at pH 7 and 6 in the III_h temperature region are identical.

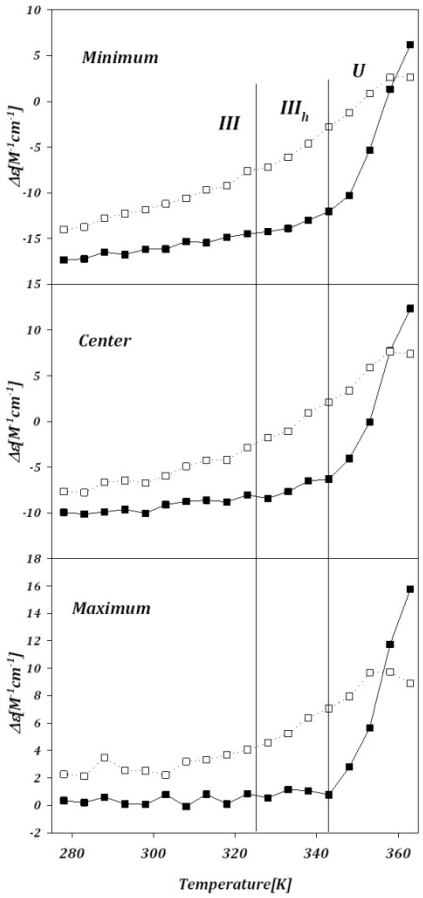


Figure 35: Taken from ref (*280*) and modified. Δε vs temperature of bovine heart pH7.0 (closed squares) and bovine heart pH8.5 (open squares) ferricytochrome c between 278 and 363K measured in MOPS buffer. The upper panel is at the 24010cm^{-1}, the middle is at 24450cm^{-1}, and the lower is 24876 cm^{-1}, corresponding to the minimum, the center and the maximum position of the B-band couplet at room temperature. The temperature regimes of the three states III, III_h and U (unfolded) are indicated.

An earlier attempt to probe the thermal transitions of oxidized horse heart cytochrome c by CD deserves to be mentioned in this context. Myer reported a change of Soret band ellipticity to occur at 313K, which is not reproduced by any of our data (neither for bovine nor for horse heart) (*248*). An inspection of his visible CD spectra shows a

disappearance of the couplet already at 333K. An explanation for the observed data was not provided. We have no explanation for this discrepancy but it should be noted that Myer did not use a buffer for his protein solution.

5.2.2 Secondary Structure Analysis of III_h by UV-CD Measurements

We used far-UV CD spectroscopy to probe secondary structure changes associated with the transition between different III-states. Far-UV CD spectra recorded in the temperature range used for the visible-CD measurements are consistent with the two melting phases reported for horse heart ferri-cytochrome c by Myer (*248*). Figure 36 shows the Δε value of bhc and hhc measured at 218 nm (44964 cm^{-1}) in 0.1 M Tris HCl buffer as a function of temperature.

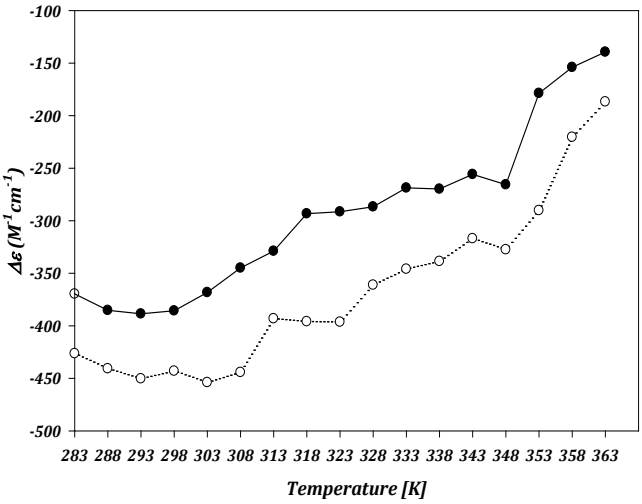


Figure 36: Taken from ref (*280*) and modified. Δε of bovine (closed circles) and horse heart (open circles) ferricytochrome c as a function of temperature at 44964cm^{-1} in the far UV region with a pH7 Tris HCl buffer.

The onset of the first phase is at 323K, while the second one starts at 343K. The first phase can be assigned to the $III \rightarrow III_h$ transition and the second one to the melting of most of the secondary structure. These data add further support to the notion that the

III→III$_h$ transition takes place at pH values well below the pK value of the alkaline transition.

To quantitatively probe secondary structure changes associated with the III→III$_h$ transition we performed far-UV CD experiments using the synchrotron radiation (SR) source at Brookhaven National Laboratories. Thus, we obtained CD spectra in a wavelength range between 170 and 260 nm, which allow for a more reliable secondary structure analysis than CD spectra taken with a conventional instrument. SRCD spectra were taken at 296K and 345K (Figure 37).

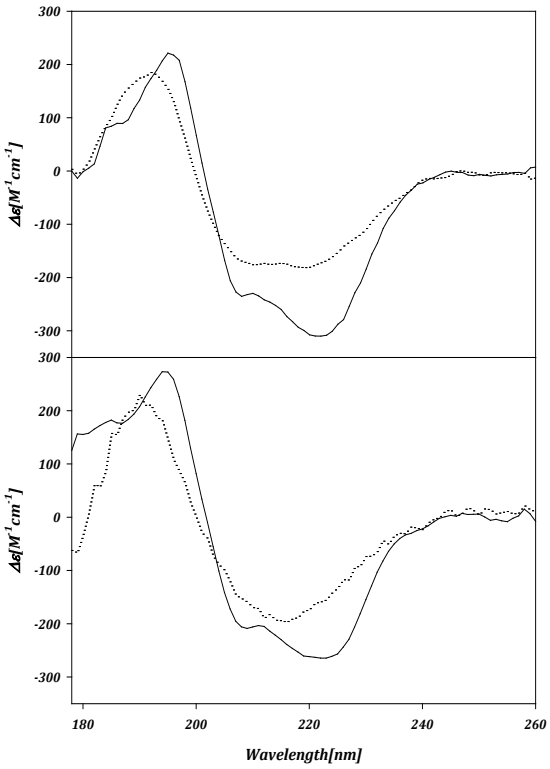


Figure 37: Taken from ref (*280*) and modified. Temperature-dependent far-UV SRCD spectra of bovine (top) and horse (bottom) heart ferricytochrome c. Solid line: 296K, Dotted line: 345K.

The spectra were processed with the spectral processing program CDTools (*283*) and analyzed with the secondary structure from circular dichroism spectroscopic data

program, Dichroweb (*269, 284*), using the program ContinLL and reference set 1. This analysis yielded approximately a 15% loss of α-helical content, redistributed to mostly extended and to a minor extent turn and unordered structures at 345 K (state III_h) compared with 296 K (III). Hence, state III_h is partially unfolded rather than a classical molten globule state.

5.2.3 Interpretation of Visible CD and Absorption Spectra

A more thorough understanding of the above reported CD spectra requires a comparison with the respective absorption spectra. The spectra in Figure 31 and Figure 34 reveal that the peak of B-band absorption is situated between the couplet extrema at temperatures below 343K. At higher temperatures, however, the peak coincides with the maximum of the Cotton band. This result indicates that the B-band is split into its B_x and B_y component below 343K, owing to electronic and vibronic perturbations imposed by the protein environment (*211, 222*).

We used a recently developed vibronic coupling model (*211*) to simulate three CD spectra of bovine cytochrome c (in Tris HCl buffer), recorded at 278, 338 (data not shown) and 363K (simulations shown in Figure 38).

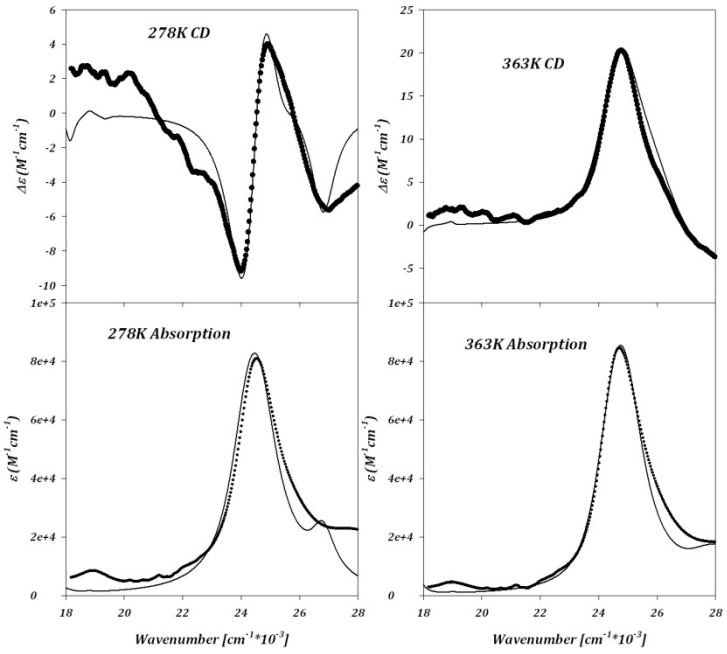


Figure 38: Taken from ref (*280*) and modified. Experimental (dots) and simulated (solid line) CD (upper panel) and absorption spectra (lower panel) of bovine heart ferricytochrome c. The experimental spectra were measured at the indicated temperatures.

Since the oxidation marker band v_4 dominates the resonance Raman spectrum for excitations of the B-band region (data not shown), we considered only the contribution of this mode to the vibronic side band. This model is a simplification, but it is sufficient for exploring the amount of splitting. The general theory of the model is described in ref. (*211*), so we confine ourselves here to presenting the most relevant equations and some modifications. If one neglects coupling between $|B_x\rangle$ and $|B_y\rangle$ states, the energy eigenvalues associated with the $0 \rightarrow 0$ transition are written as:

$$E_{l,0}^{B'} = E_{l,0}^{B} - \sum_{\Gamma''} \frac{\left[c_{B_l B_l}^{\Gamma''}(\nu_4) \right]^2}{\Omega_{\nu_4}^{B}} \qquad (44)$$

where $E_{l,0}^{B'}$ (*l=x,y)* denotes the eigenenergies of the vibronically and electronically perturbed B-state, $E_{l,0}^{B}$ are the energies of the electronically perturbed states $|B_l,0\rangle$. $\Omega_{\nu_4}^{B}$ denotes the vibrational energies of the ν_4 vibration in the excited B-states. The parameter $c_{es}^{\Gamma''}(\nu_4)$ (*e,s=B*$_l$) denotes the corresponding vibronic coupling matrix element. Γ' is the effective symmetry of the respective vibration in the presence of symmetry lowering perturbations. The second term on the right of eq. 44 describes the contribution from intrastate Franck-Condon (*A$_{1g}$*) and Jahn-Teller (*B$_{1g}$, B$_{2g}$*) coupling associated with the ν_4 mode, which is totally symmetric in ideal D_{4h}. The pure electronic eigenenergies, $E_{x,0}^{B}$ and $E_{y,0}^{B}$ are identical in D_{4h}, but are split in a protein environment. The total electronic splitting of the B_0 is written as (*224*):

$$\Delta E^{B} = \Delta_{Stark} + \frac{2\delta_{QB}^{A_{1g}}\delta_{QB}^{B_{1g}}}{E^{B_0} - E^{Q_0}} \qquad (45)$$

where $\delta_{QB}^{A_{1g}}$ and $\delta_{QB}^{B_{1g}}$ describe electronic perturbations of A_{1g} and B_{1g}-symmetry. The latter is assigned to the quadrupole moment of the internal electric field (*212*). These perturbations cause a mixing of Q- and B-states and an equal splitting of their energies. Δ_{Stark} denotes the different shifts of B_x and B_y, owing to a quadratic Stark effect, which mixes the ground state with both components of the excited B-state (*223*). Since the Stark splitting scales with the square of the corresponding transition dipole moment, it is much larger for the B than for the Q-band. The electronic part of the Q-band splitting can be nearly exclusively assigned to the electric field's quadrupole moment (*221, 223*).

As mentioned above, the vibronic side band B_v is modeled as mainly resulting from the vibronic coupling of ν_4. This is a rather crude approximation of the vibronic side band. The respective transition dipole moments (electronic and magnetic) $\vec{\pi}^B_{l,1_{\nu 4}}$ are described by the equation:

$$\vec{\pi}^B_{l,1_{\nu 4}} = \frac{c^{\Gamma''}_{B_l B_l}(\nu_4)}{\Omega^B_{\nu_4}} \vec{\pi}^B_{l,0_{\nu 4}} \qquad (46)$$

Figure 38 compares experimental data and simulation for the spectra recorded at 278 and 363K. The simulation for the 338K spectra is of similar quality. This analysis yielded an identical splitting of 360 cm^{-1} for 278 and 338K (data not shown), whereas the splitting is practically zero for the 363K measurement. Interestingly, our splitting value is only slightly larger than what Manas et al. observed for the B-band in the low temperature spectrum of Zn-substituted horse heart cytochrome c (296 cm^{-1}) (*285*). Compared with the low temperature spectrum, the rotational strength values for B_x and B_y at 338K are reduced by factors of 0.81 and 0.76, respectively. At 363K the amount of the total rotational strength is increased by a factor of 1.6 compared to the corresponding value obtained for 338K. The splitting most likely provides us with a lower limit for the electronic perturbation energy, since vibronic perturbations, which are not fully accounted for at the present level of analysis, generally reduce, rather than increase, the splitting (*224*).

Our experimental data suggest that the B-state electric field strength in the heme plane is not significantly changed by the III$\rightarrow$III$_h$ transition. The concomitant decrease of the rotational strength for B_x and B_y might reflect a reduction of out-of-plane deformations,

which have been shown to cause substantial rotational strength of opposite sign for both B-band transitions (*251*). Additionally, invoked orientational changes of aromatic residues in the heme pocket (e.g. F82) certainly contribute to the rotational strength (*250*). Specific site-directed mutation experiments are necessary to clarify this issue. A final assessment of the B-band's Stark effect has recently been made by virtue of a full vibronic analysis of the Q and the B-band splitting of ferri- and ferrocytochrome c (*212*).

The absence of any detectable splitting in the CD spectra of bhc in the unfolded states and, as shown below, in states IV and V seem to suggest an absence of a quadratic Stark effect. This hypothesis is unlikely for a variety of reasons. First, our spectra indicate that the heme is still bound to the protein in these states. It is therefore likely that the protein still produces an electric field in the heme plane. Moreover, a substantial decrease of the electric field would cause a significant redshift of both Q and B-band (*223*), which is not observed. The absence of splitting can be rationalized in terms of a compensation of vibronic and electronic perturbations, which generally have opposite signs for the Q-band. Alternatively, one may invoke larger protein fluctuations as giving rise to a distribution of different orientations of the electric field with respect to the heme plane. If this distribution is nearly symmetric with respect to one of the C_mFeC_m of the heme macrocycle, the effective electronic splitting could be very small, even though the electric field is still rather strong (*212, 223, 286*).

It is necessary in this context to point out that the interpretations of visible CD spectra of heme proteins are often too simplistic. For hemoglobin and myoglobin (*211, 250*), single positive Cotton bands, were measured for the respective intact molecules. The

absence of a couplet in the B-band regions can have several reasons. Firstly, the band splitting is small and one component (x or y) has more rotational strength than the other one. This is most likely the case for the above discussed thermally unfolded state of cytochrome c. That this is possible has been shown by Blauer et al. (*251*). Secondly, it is possible that only one of the two components carries substantial rotational strength. If the band is split, this can lead to a positive (or negative) Cotton band, the wavenumber position of which does not coincide with that of the respective absorption band. As recently shown, this is the case in deoxymyoglobin and myoglobin cyanide (*211*).

5.2.4 Characterizing the Alkaline States

The results discussed thus far show that III_h and IV are thermodynamically distinct. The $III \rightarrow III_h$ transition state is generally associated with the disappearance of the 695 nm band, which is interpreted as reflecting the rupture of the Fe^{3+}-M80 bond (*186, 196*). This is certainly the case for state IV, but the data displayed in Figure 33 suggest a decrease of the 695 nm band's oscillator strength rather than its disappearance in III_h. This notion is strongly corroborated by the still significant rotational strength of the band in the III_h temperature region. We therefore infer that III_h differs structurally from IV-states in that the Fe^{3+}-M80 bond, even though weakened, is still present in the former. This notion is consistent with the above discussed similarity of the CD spectra of III and III_h. To explain possible structural differences, we measured the temperature dependence of the B band (CD and absorption) of bhc at pH 8.5 and 10.5. The respective pH values were adjusted by adding NaOH to a 1 mM MOPS buffer solution

for the former and to a 50 mM Bis/Tris buffer solution for the latter. Figure 39 depicts the spectra measured at pH 8.5.

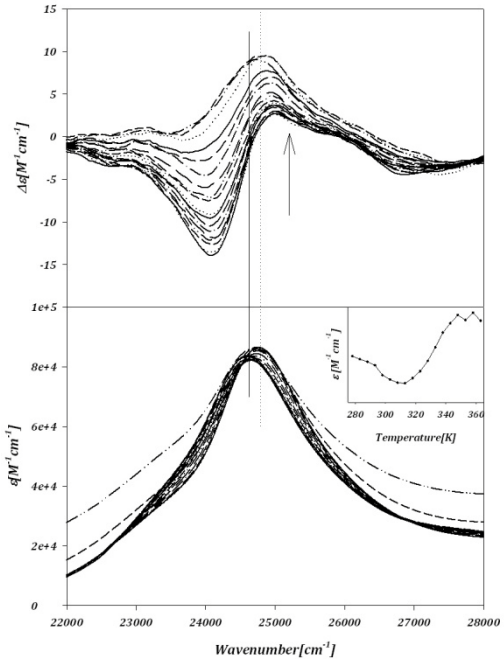


Figure 39: Taken from ref (*280*) and modified. Temperature dependant CD (top) and absorption (bottom) spectra of bovine heart cytochrome c measured in the Soret band region from 22000-28000cm^{-1}. The spectra were measured in a pH8.5 MOPS buffer with a 0.05mM concentration. Arrow indicates changes with increasing temperature. Solid line indicates maximum absorption frequency position at low temperature and dotted line indicates maximum absorption frequency at high temperature. Inset: ε vs temperature at the frequency of maximum absorption.

The corresponding insets exhibit the temperature dependence of ε_{max}. The couplet observed at low temperatures converts to a positive Cotton band at significantly lower temperatures (333K) than the couplets obtained at neutral (Figure 34) and (slightly) acidic pH (Figure 31). Figure 40 compares the temperature dependence of both, $\Delta\varepsilon$ and the Kuhn anisotropy, measured at pH 8.5 with the respective data recorded with solution exhibiting pH 7 at room temperature.

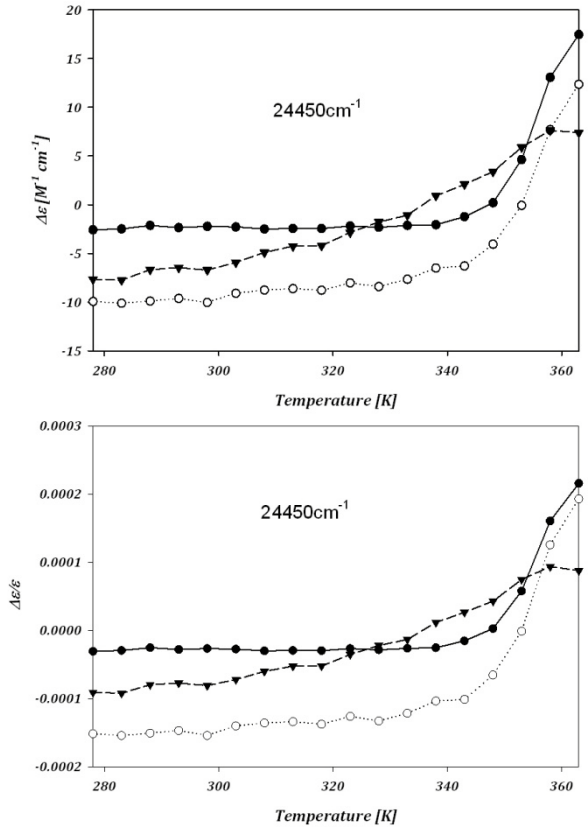


Figure 40: Taken from ref (*280*) and modified. Top) Δε vs temperature and Bottom) Kuhn anisotropy, Δε/ε, vs temperature at the indicated frequency position of pH7 (tris buffer) bhc (open circles), pH7 (MOPS buffer) bhc (closed circles) and pH 8.5 (MOPS buffer) bhc (triangles), between 278 and 363K.

The rather smooth temperature dependencies of Δε and Δε/ε recorded at pH 8.5 stretch over the entire temperature range investigated, thus indicating a non-cooperative transition between a low and high temperature state, which, based on the respective CD spectra, resemble III_h and U. However, a two-state transition is ruled out by the absence of an isodichroic point and by the corresponding non-monotonous temperature-dependence of ε_{max} (inset Figure 39), which exhibits a pronounced minimum at 310 K. At this temperature the effective pK-value of the alkaline transition is nearly identical with the adjusted pH of 8.5 (*191*), so that the minimum coincides with the halfpoint of the III → IV transition. It should be notified that the CD spectra in Figure

39 bear some resemblance with the aforementioned spectra reported by Myer, which might indicate that his actual pH was slightly basic (*248*). In water, cytochrome c can act like a buffer due to the large number of protonatable side chains on the surface of the protein. Thus, a change in pH may occur without a buffer.

More information about state(s) IV can be inferred from the CD (and absorption) profiles measured at pH 10.5, which are exhibited in Figure 41.

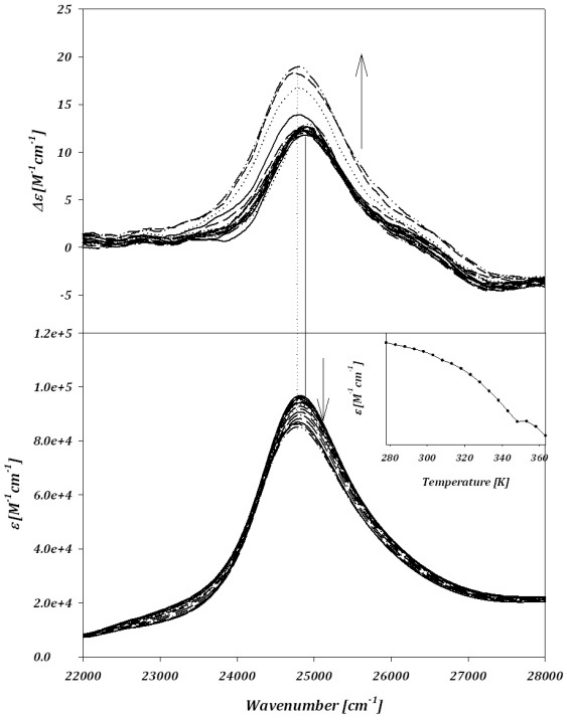


Figure 41: Taken from ref (*280*) and modified. Temperature dependant CD (top) and absorption (bottom) spectra of bovine heart cytochrome c measured in the Soret band region from 22000-28000cm^{-1}. The spectra were measured in a pH10.5 bis/tris buffer with a 50mM concentration. Arrows indicate changes as a function of temperature. Solid line indicates maximum absorption frequency position at low temperature and dotted line indicates maximum absorption frequency at high temperature. Inset: ε vs temperature at the frequency of maximum absorption.

At this pH the alkaline state(s) IV are predominantly populated even at room temperature. The corresponding CD spectra already show a positive Cotton band in the Soret region, which is quite distinct from the symmetric couplet observed for the III$_h$

state. The rotational strength increases with increasing temperature while the absorption peak exhibits a redshift, in contrast to what was observed for the $III_h \rightarrow U$ transition at neutral and modestly acidic pH. The positive Cotton band and its coincidence with the absorption maximum reflect again the drastically reduced splitting between the B_x and B_y bands. Figure 42 exhibits the corresponding temperature dependence of the Kuhn anisotropy and $\Delta\varepsilon_{max}$.

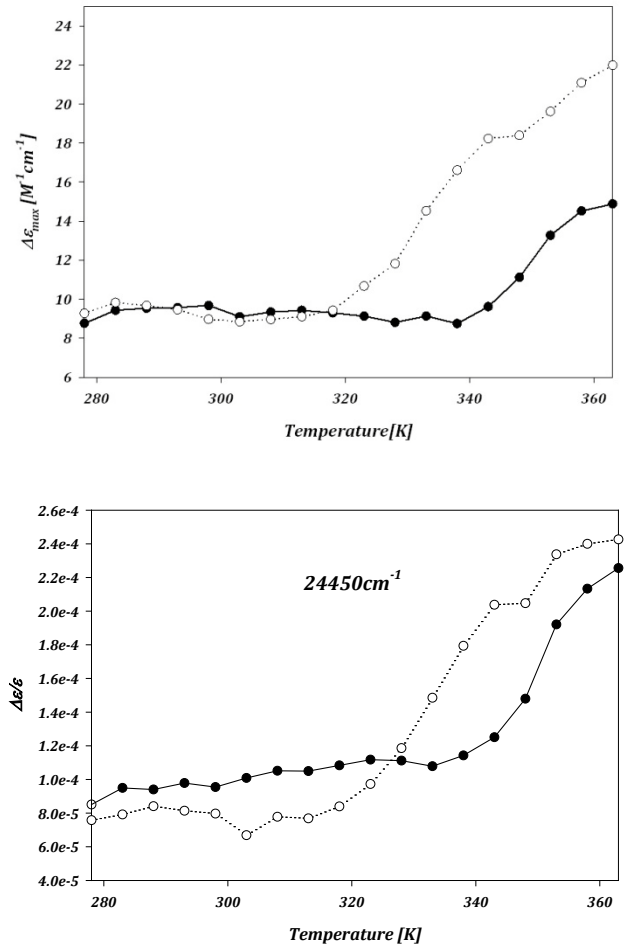


Figure 42: Taken from ref (*280*) and modified. Top) $\Delta\varepsilon_{max}$ vs temperature and Bottom) Kuhn anisotropy, $\Delta\varepsilon/\varepsilon$, vs temperature at the indicated frequency position of pH10.5 (bis/tris buffer) bhc (closed circles) and pH11.5 (bis/tris buffer) bhc (open circles) between 278 and 363K.

While $\Delta\varepsilon_{max}$ seems to be nearly temperature independent between 278 and 338 K within the experimental uncertainty, the less noisy Kuhn anisotropy exhibits a statistically significant though small increase with rising temperature. Both $\Delta\varepsilon_{max}$ and Kuhn anisotropy exhibit a steep increase with increasing temperatures above 348 K. The data indicate the existence of at least two thermal transitions, involving an intermediate IV_h and a certainly unfolded state IV_u. Apparently, the rotational strength is significantly larger at pH 10.5 than at 8.5 and at lower pH values for the entire temperature range. This indicates that at pH 8.5 state IV still coexists with other states (III_h or some yet unidentified state) even at high temperatures. The interpretation of the data at this pH is complicated by the possible coexistence of most likely three alkaline isomers and also by the fact that the (lysine) protonation and the dissociation of M80 are thermodynamically distinct processes (*194*).

To gain a complete picture of the conformational manifold sampled at alkaline pH we also measured the temperature dependence of the CD and absorption spectra of bhc at pH 11.5. The alkaline state, V, can be expected to be populated at room temperature at this pH (*192*). The CD spectra in Figure 43 (pH 11.5) are qualitatively similar to those obtained at pH 10.5 in that they start as a positive Cotton band as well and, like the spectra in Figure 41, increase with increasing temperature. However, the corresponding spectra (Figure 43) are quantitatively different from that observed in Figure 41.

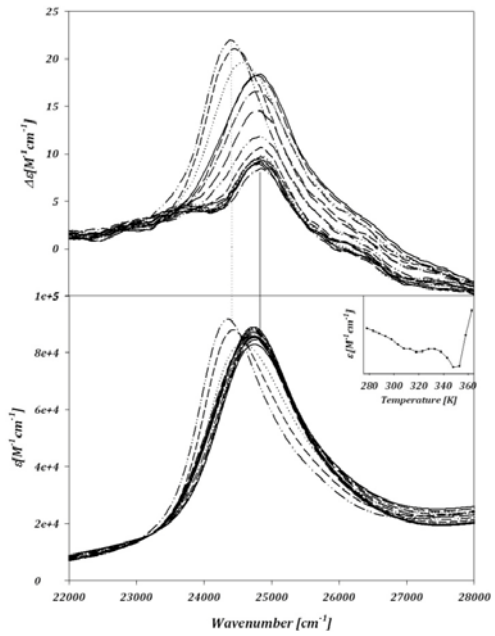


Figure 43: Taken from ref (*280*) and modified. Temperature dependant CD (top) and absorption (bottom) spectra of bovine heart cytochrome c measured in the Soret band region from 22000-28000cm^{-1}. The spectra were measured in a pH11.4 bis/tris buffer with a 50mM concentration. Solid line indicates maximum absorption frequency position at low temperature and dotted line indicates maximum absorption frequency at high temperature. Inset: ε vs temperature at the frequency of maximum absorption.

The magnitude of this difference can be seen by comparing plots in Figure 42, where the associated Kuhn anisotropy and $\Delta\varepsilon_{max}$ are plotted as a function of temperature. The rotational strength of the Cotton band representing state IV is smaller than that of state V upon an increase in temperature above 320K. A comparison of the respective absorption spectra reveals that this predominantly reflects differences between the respective oscillator strengths. The differences between the corresponding Kuhn anisotropies are much less pronounced. $\Delta\varepsilon_{max}$ is practically temperature independent below 340 K for state IV, whereas it exhibits a minimum until 320 K for state V. This minimum also appears in both $\varepsilon_{max}(T)$ (Figure 43) and somewhat less pronounced in $\Delta\varepsilon/\varepsilon$. The melting into the unfolded state occurs at lower temperatures for state V (T >

333 K) than for IV (T > 340 K). The absorption peak extinction decreases with increasing temperature until the band position exhibits a large (~500cm^{-1}) redshift.

For comparative purposes, Figure 44 depicts the differences of the CD and absorption spectra of bhc, which reflect the aforementioned native and non-native conformations.

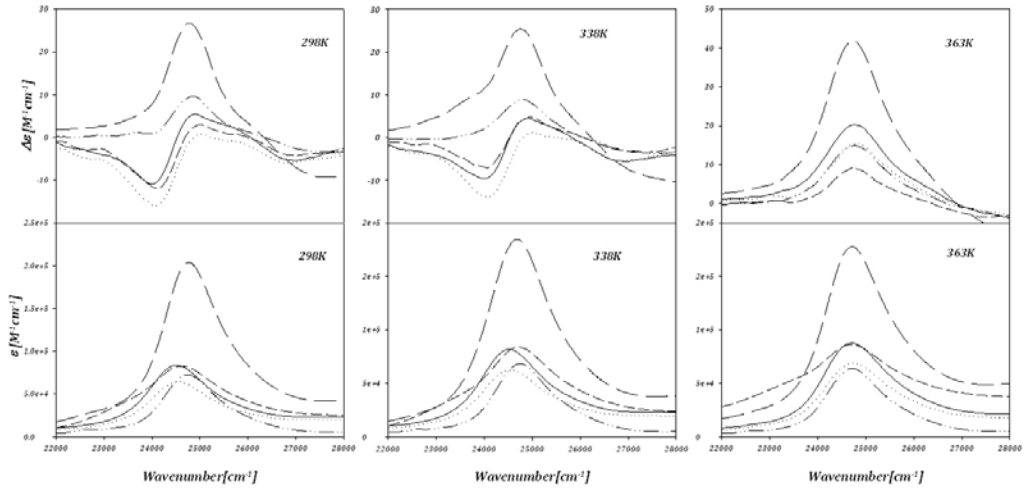


Figure 44: Taken from ref (*280*) and modified. CD (top) and absorption (bottom) of all investigated pH measurements (pH7 Tris HCl buffer: solid line, pH7.0 MOPS buffer: dotted line, pH8.5 MOPS buffer: short dashed line, pH10.5 Bis/Tris buffer: dashed-dot-dot Line and pH11.4 Bis/Tris buffer: long dashed line) of bovine heart cytochrome c at room temperature (left panels), intermediate temperature (middle panels) and high temperature (right panels). Exact temperatures are indicated in individual panels.

The low-temperature spectra represent the states III – V. Spectra taken at the indicated intermediate temperature can at least partially be assigned to the corresponding thermal intermediates, and the spectra taken at high temperature reflect the corresponding unfolded state. At room temperature, the CD spectra obtained in state III (pH7 and 8.5) display couplets in the Soret region with the peak extinctions situated in the middle of the couplet. The higher pH measurements in states IV (pH10.5) and V (pH11.5) display positive Cotton bands at room temperature, though the intensity of the band in state V is much larger than any of the other investigated states. At high

temperature all of the CD profiles turn into positive Cotton bands and become (nearly) coincident with their respective absorption profiles. However, in terms of absolute rotational strength, band shapes and absorptivity, the bands assignable to the unfolded states of III, IV and V are different, thus indicating that these states are not structurally identical in the vicinity of the heme group. In what follows we term these states III_u, IV_u and V_u.

We deduce the following minimal thermodynamic scheme (R1) from our data presented in this section along with the results obtained by Verbaro et al. which characterize an intermediate of the state III→IV alkaline transition:

$$\begin{array}{ccccc} III_u & \Leftrightarrow & III_u^* & \Leftrightarrow & IV_u \Leftrightarrow V_u \\ & & & & \Updownarrow \\ \Updownarrow & & \Updownarrow & & \Updownarrow \quad V_{h2} \\ & & & & \Updownarrow \\ III_h & \Leftrightarrow & III_h^* & \Leftrightarrow & IV_h \Leftrightarrow V_{h2} \\ \Updownarrow & & \Updownarrow & & \Updownarrow \quad \Updownarrow \\ III_n & \Leftrightarrow & III_n^* & \Leftrightarrow & IV_n \Leftrightarrow V_n \end{array} \qquad \text{(R1)}$$

The additional thermal intermediates IV_h and V_h are deduced from the $\Delta\varepsilon_{max}(T)$ in Figure 42. It can be ruled out that the observed minimum between 320 and 330K represent, in fact, a partial population of state V, since the latter is associated with a much larger rotational strength. We wish to reiterate in this context that state IV, and to some extent also state III, encompasses conformational substates with different axial ligands (for IV) (*194*) or different Fe-M80 bonding strengths (III) (*275, 276*). It should be

mentioned that the above scheme might still underestimate the number of intermediates; the temperature dependence of ε_{max} for state V indicates that at least two intermediates are involved. We also propose that not only the alkaline states, but the intermediate, III*, will populate at least one thermal intermediate state before the protein unfolds at higher temperatures.

Finally, we deem it noteworthy that the temperature dependences of circular dichroism and Kuhn anisotropy are very similar at all three investigated alkaline pH values. Since $\Delta\varepsilon/\varepsilon \propto \vec{m}_B \cdot \cos(\vec{\mu}_B, \vec{m}_B)/\vec{\mu}'_B)$, with $\vec{m}_B$ as the magnetic and $\vec{\mu}'_B$ as the electronic matrix element of the transition into the B-states, our data are indicative of a substantial increase of a magnetic dipole component in the heme plane. Since the latter arises from heme-protein interactions (*250*), our data thus indicate an increase of some electronic coupling between the heme and protein chromophores in non-native states, which somewhat contradicts the notion of reduced heme-protein coupling in partially unfolded states.

5.3 Secondary Structure Analysis of pH Induced Non-native States

In order to quantify concomitant changes in secondary structure of the states involved in the alkaline transition, we measured the VUV CD spectra of oxidized cytochrome c was measured in the region between 175 and 240 nm from pH 3 to 12 at room temperature in a 1mM MOPS (3-(*N*-morpholino)propanesulfonic acid) buffer. We included state II in here to see if any major changes in the secondary structure occur upon acidifying the protein. A rather low ionic strength was chosen to avoid any influence of anion binding (*202, 204, 205, 276, 287*). As indicated in the Introduction,

this pH range covers the states II-V of the protein. The SRCD spectra can be seen in Figure 45.

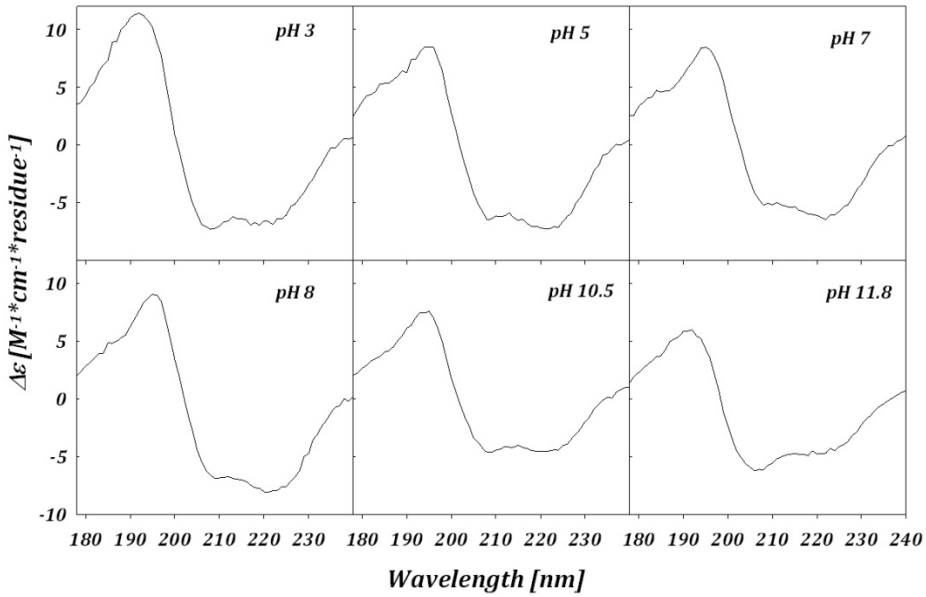


Figure 45: Taken from ref (*210*) and modified. SRCD spectra of oxidized horse heart cytochrome c taken at indicated pH values plotted from 178-240nm in Δε units per residue.

The online Dichroweb software was employed to analyze these spectra. The Contin analysis algorithm (*262*) was utilized with reference set 6. Protein concentrations were determined from the Soret band absorption of the protein (*181*) solution in a 1mm quartz cell (Hellma) by using an extinction coefficient of $10.6 \cdot 10^4$ ($M^{-1} \cdot cm^{-1}$). From the crystal structure of yeast ferri-cytochrome c, Louie and Brayer determined 40% α-helix, 22% β-turn and negligible β-sheet structure (*180*). A practically identical result was obtained for horse heart cytochrome c (*181*). The analysis of the SCRD spectra recorded at pH 7 and 8, at which the native state III is predominantly populated, yielded 44% α-helix and 23% β-turn, which is in very good agreement with crystallographic values. At pH 5, at which the protein is still believed to be in its native state (*187*), the values were almost identical, but indicate a small loss (~3%) of α-helical content. In state II (pH 3)

both the α-helix and β-turn content decreased by ~3% compared to that obtained for the native state measurement. This modest decrease indicates that this state still exhibits a rather intact secondary structure, which would be consistent with the notion that it is a molten globule (*188*). For pH 10.5, at which the protein is almost exclusively in state IV in a low ionic strength solution (*278*), the α-helix content was determined to be 35.5%, while the β-turn fraction is comparable with that of the native state. In state V, probed at pH 11.8, the α-helix and β-turn contents were 31% and 20%, respectively. All these changes of the secondary structure are modest, which shows that the structural changes induced by the conformational transitions of ferricytochrome c involve mostly changes of the tertiary structure.

5.4 Thermodynamic Analysis[3]

Thus far, this chapter has demonstrated how the CD spectrum of the B-state transition of heme proteins could be used to compare different conformations of the protein in solution. In what follows, a simple thermodynamic model is presented by which they can be consistently analyzed. The thermal unfolding of the native state of cytochrome c can be modeled by the following reaction scheme:

$$III \underset{\longleftarrow}{\overset{K_{III,1}}{\longrightarrow}} III_h \underset{\longleftarrow}{\overset{K_{III,2}}{\longrightarrow}} III_u \qquad \text{(R2)}$$

where III_h is a thermodynamic intermediate and III_u is the thermally unfolded state.

[3] The following results have been analyzed by Dr. Reinhard Schweitzer-Stenner and published in ref 210. Schweitzer-Stenner, R., Hagarman, A., Verbaro, D., and Soffer, J. (2009) Conformational Stability of Cytochrome c Probed By Optical Spectroscopy, *Meth. Enzymol. 466*, 109-153.]
The results are relevant in this context to show how the previous temperature dependant CD spectra can be consistently analyzed with a simple thermodynamic model

Based on reaction scheme R2 the following formalism can be used to describe the temperature dependence of $\Delta\varepsilon$ at neutral pH:

$$\Delta\varepsilon = \frac{\Delta\varepsilon_0 + \Delta\varepsilon_h e^{-G^h/RT} + \Delta\varepsilon_u e^{-G^u/RT}}{Z} \tag{47}$$

where G^h and G^u are the Gibbs energies of the states III_h and III_u relative to the folded state III, R is the gas constant and T the absolute temperature. Z is the partition sum, which is written as:

$$Z = 1 + e^{-G^h/RT} + e^{-G^u/RT} \tag{48}$$

We consider the III⇔III_h transition as non-cooperative, so that

$$G^h = H^h - TS^h \tag{49}$$

with temperature independent enthalpy and entropy. In our fit to the data, the transition temperature:

$$T_h = \frac{H^h}{S^h} \tag{50}$$

is used as free parameter. However, this approach does not work for the III_h⇔III_u transition which indicates some cooperativity of the underlying unfolding/folding process in that a fit with a non-cooperative approach as utilized for III⇔III_h does not reproduce the slope of the $\Delta\varepsilon$(T) graphs above 340 K. Cooperativity, could be empirically accounted for by employing a Hill function (i.e. substituting $e^{-G^u/RT}$ by $e^{-nG^u/RT}$ in eqs. (47) and (48), n: empirical Hill coefficient). To avoid the use of physically not very well

defined parameters, however, it is more appropriate to use the approach of Uchiyama et al. (*288*) who considered a temperature dependent enthalpy, which in first order can be accounted for by:

$$H^u = H_0^u + \delta c_p (T - T_u) \tag{51}$$

where δc_p is the heat capacity difference between the two states at constant pressure and T_u is the transition temperature for the $III_h \Leftrightarrow III_u$ transition. Then, we can write the Gibbs energy as:

$$G^u = H_0^u + \delta c_p \left(T - T_u\right) + T\left(S^u \left(T_u\right) + \delta c_p \ln\left(\frac{T}{T_u}\right)\right) \tag{52}$$

The solid lines in Figure 46 and Figure 47 result from a consistent fit of the above formalism to the experimental data.

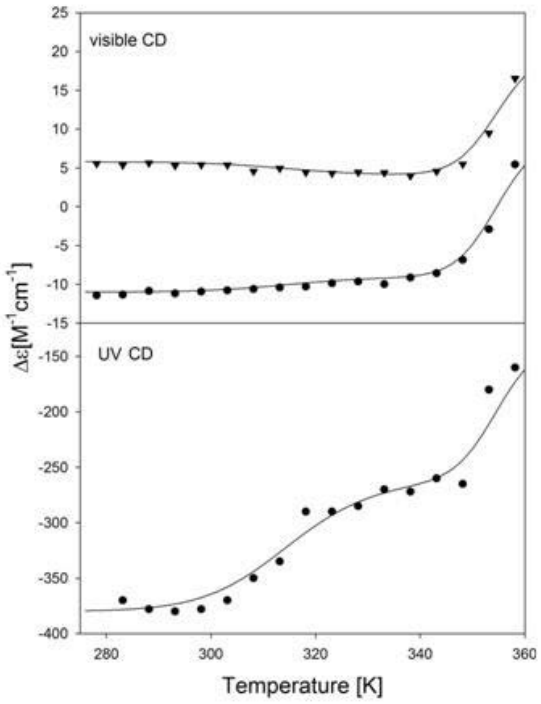


Figure 46: Taken from ref (*210*) and modified. Δε versus temperature of oxidized bovine horse heart cytochrome c between 278 and 363 K. Upper panel: Δε obtained from the CD spectra in Figure 31 at 24876 cm^{-1} (triangles) and 24010 cm^{-1} (filled circles). Lower panel: Δε obtained from the corresponding CD spectra at 44964 cm^{-1}. Solid lines results from the fits described in the text.

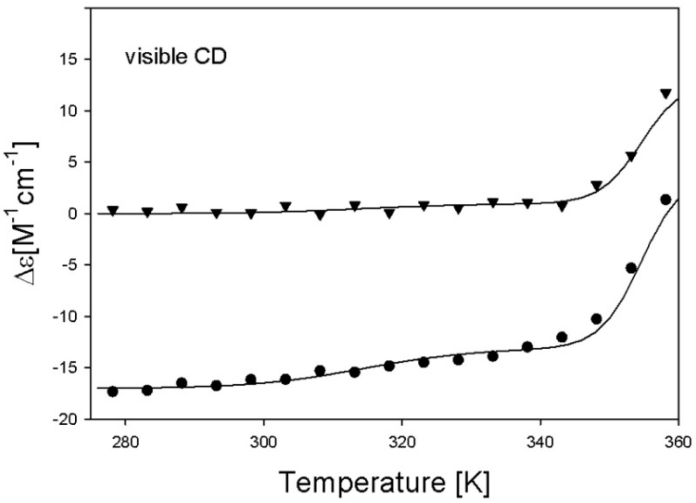


Figure 47: Taken from ref (*210*) and modified. Δε versus temperature of oxidized bovine horse heart cytochrome c between 278 and 363 K. The Δε values obtained from the CD spectra in Figure 34 at 24876 cm^{-1} (triangles) and 24010 cm^{-1} (filled circles). Solid lines results from the fits described in the text.

Consistence here means that the very same thermodynamic parameters were used to fit $\Delta\varepsilon$(T) and ε(T) data obtained with the same buffer. In order to minimize the ambiguity we took ΔH and δc_p values from the calorimetric studies of Uchiyama et al. (*288*) for the $III_h \Leftrightarrow III_u$ transition and modified them only slightly for the purpose of optimization. The thermodynamic parameters used for the fits are listed in Table 5.

The $\Delta\varepsilon$(T) graphs measured at pH 10.5 and 11.5 were also subjected to fits based on reaction scheme R2 for the respective states, which, for the pH 10.5 data, led to the respective solid line (data points in filled circles) in Figure 48. The population of an intermediate is difficult to discern from the data, though its existence becomes more obvious if one inspects the temperature dependence of the Kuhn anisotropy. The Cotton band in the unfolded state IV_u is by a factor ¾ less intense than the CD band of the III_u state populated at neutral pH and high temperatures. This difference seems to make sense in that we expect IV_u to be even more unstructured than III_u, which would lead to a less ordered heme environment.

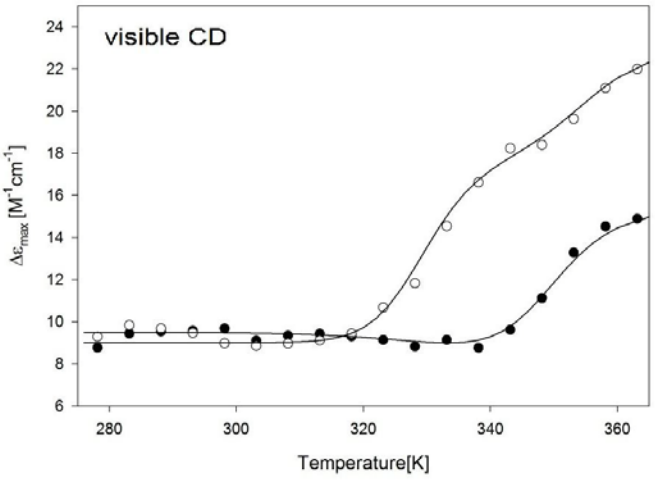


Figure 48: Taken from ref (*210*) and modified. $\Delta\varepsilon$ versus temperature of oxidized bovine horse heart cytochrome c between 278 and 363 K. The $\Delta\varepsilon$ values obtained from the CD spectra in Figure 41 (state IV, filled circles) and Figure 43 (state V, open circles) at the position of the Cotton band maximum. Solid lines results from the fits described in the text.

The attempt to fit R2 to the data acquired at pH 11.5 had only limited success in that the rather steep increase of $\Delta\varepsilon$(T) at 320 K could not be reproduced. This steepness indicates that the observed changes at this temperature are already reflecting the transition into an unfolded state and that the slight variation of $\Delta\varepsilon$(T) could be interpreted as indicating the occupation of an intermediate. The reaction scheme (R3) was therefore extended to the following based on the results presented prior:

$$V \rightleftarrows V_h \rightleftarrows V_{u1} \rightleftarrows V_{u2} \qquad \text{(R3)}$$

For the transition into V_{u1}, the same cooperative model was utilized to model the transitions $III_h \Leftrightarrow III_u$ and $IV_{u1} \Leftrightarrow IV_{u2}$. This model yielded an excellent fit to the experimental data (Figure 48). Interestingly, the $\Delta\varepsilon_{max}$ value for V_{u1} is similar to that obtained for IV_u and V_u. This similarity suggests that, e.g., III_u might actually not be the final state of the unfolding process and that the transition into III_u might be followed by

another transition into the final unfolded state at higher temperatures. Generally, this state is not accessible to the optical measurements used for the current study, owing to the onset of aggregation.

	pH 7 MOPS	pH 7 Tris HCl	pH 10.5	pH 11.5
H^h [kJ/mol]	100	100	100	200
H^u[kJ/mol]	300	250	200	200
T_h [K]	315	315	320	330
T_u[K]	355	355	350	355
c_p[kJ/mol K]	-2.8	-2.9	-2.8	-2.8

Table 5: Thermodynamic parameters obtained from fitting the $\Delta\varepsilon_{max}(T)$ and $\Delta\varepsilon_{min}(T)$ graphs in Figure 47 (bovine ferricytochrome c, MOPs buffer, pH 7), Figure 46 (Tis HCl buffer, pH 7 at room temperature) and Figure 48 (Bis/Tris buffer, pH 10.5 and 11.5).

5.5 Out-of-plane Deformations of Ferrocytochrome c Isoforms

Nonplanar deformations of heme groups have been quantified in the past by analyzing crystal structures of the respective proteins. Although this method is very useful, crystal structures for all relevant ligation, spin and redox states of a given protein are not always available. Thus, alternative methods have been used, such as NSD analyses of hemes extracted from molecular dynamics studies. Also, modifications of the heme groups by varying solution conditions (pH, ionic strength and viscosity) can generally not be explored by X-ray crystallography. Changes in pH are particularly relevant for ferricytochrome c whose heme environment is highly pH dependant. Thus, we focus on exploring and comparing oop deformations of the heme group in three ferrocytochrome c isoforms; horse heart (hhc), chicken (chc) and yeast isoenzyme 1 from *Saccromyces cerevisiae* where C102 was mutated to T to prevent aggregation (yc). We used Soret excited resonance Raman spectroscopy to obtain X and Y polarized spectra and focus on the low frequency region (<1000cm^{-1}) where the oop modes are visible.

In what follows, we first assign the low wavenumber Raman bands visible in our spectra. Secondly, we detail our normalization of the resonance Raman spectra and use the intensities and depolarization of oop modes to quantitatively describe differences between investigated heme groups. Next, we use intensities and depolarization ratios of in-plane and substituent modes to further compare the measured cytochrome c species. Finally, we compare our results to NSD results obtained for hemes extracted from X-ray crystal structures (*227, 231*) and MD simulations.

5.5.1 Band Assignments

We measured the low wavenumber-polarized resonance Raman spectra (< 1000 cm^{-1}) of horse heart, chicken, and yeast C102T cytochrome c with Soret (441.6nm) excitation. Bands of oop modes in the low wavenumber range have been previously assigned by vibrational analyses of various Ni(II) porphyrins (*289-292*), as well as various myoglobin (*293*) and cytochrome c species (*294*). Figure 49 details the band decomposition of polarized low-wavenumber resonance Raman spectra of horse heart ferrocytochrome c. The most prominent oop bands have been labeled according to Hu *et al.* (*294*).

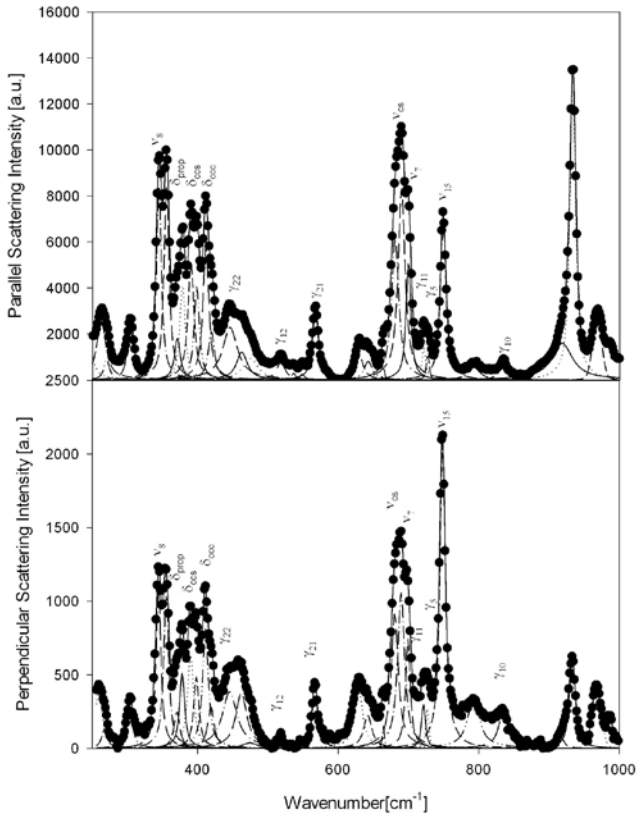


Figure 49: Taken from ref (*295*) and modified. Spectral decomposition of the polarized low wavenumber resonance Raman spectra of horse heart ferro-cytochrome c at pH8. Top- Raman intensity of parallel (x) polarization, Bottom- Raman intensity of perpendicular (y) polarization. The assignments of the most relevant bands are indicated.

The spectra in Figure 49 and corresponding spectra of chicken and yeast (spectra and decomposition not shown) were decomposed into individual bands by using our spectral decomposition program, MULTIFIT (*135*). The same wavenumbers and bandwidths were employed for corresponding bands in the x- and y-polarized spectra. Generally, we used Lorentzian profiles convoluted with a small Gaussian profile of 2.5 cm^{-1} width, which represents the spectrometer function. To analyze the out-of-plane bands, γ_5 and γ_{11}, Gaussian profiles were used convoluted with a small Lorentzian profile of 1.6cm^{-1}. Figure 50 depicts the low-wavenumber x-polarized resonance Raman spectra of all three ferrous cytochromes c, where the most prominent bands have been marked.

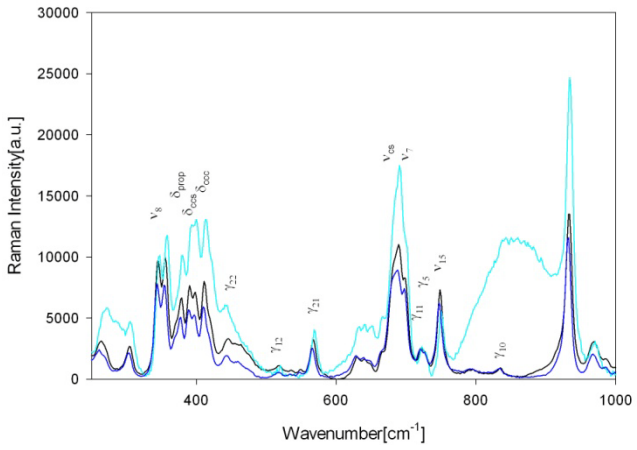


Figure 50: Taken from ref (*295*) and modified. Parallel (x) polarized resonance Raman spectra of horse heart (black), yeast (C102T) (light blue) and chicken (blue) ferrocytochrome *c*. The assignments of the most relevant bands are indicated.

Nonplanar modes become Raman active with Soret excitation predominantly via Franck-Condon coupling, i.e. by an excited state displacement along the normal coordinate of the vibration with respect to the respective equilibrium position in the deformed ground state. Examples were obtained in earlier studies of metalloporphyrins in organic solvent

(*291*). Raman bands from oop deformations such as doming (A_{2u}), ruffling (B_{1u}) and waving (E_g) have been identified in the low wavenumber region of the Raman spectra, as well as some bands of in-plane modes and of substituent deformations and stretchings. For ferrocytochrome c, the most prominent Raman active oop modes are γ_5 (doming), γ_{10} - γ_{12} (ruffling), γ_{21} and γ_{22} (waving), which correspond to the dominant SNCDs found for the heme groups of the three investigated cytochrome c molecules. Along with the identified bands of oop modes, various low-wavenumber bands assignable to in-plane bands are clearly detectable, namely the A_{1g} modes, ν_7 and ν_8, and the rather intense B_{1g} mode, ν_{15}. Finally, we identify several bands in the low-wavenumber Raman spectra from modes of peripheral substituents, i.e. deformation modes of propionic acid substituents, deformation modes of thioether bridges, the deformation mode of methyl substituents and CS stretching modes. The deformation modes of propionic acid substituents most likely gain Raman intensity from vibrational mixing with macrocycle modes as demonstrated for substituent modes of Ni-porphyrins (*291, 296, 297*). When the heme is perturbed into a lowered symmetry, substituent modes and macrocycle modes of the same symmetry vibrationally mix, especially when their vibrational frequencies are in close proximity. Some of the less considered bands are the E_u modes, ν_{50} and ν_{51}, and the B_{2g} modes, $\nu_{33(a\ and\ b)}$, at 303, 355, 462 and 472cm^{-1}, respectively.

5.5.2 Intensities and Depolarization Ratios of Out-of-plane Modes

This section first details the normalization of Raman spectra to enable a direct comparison of their Raman band intensities. To properly normalize the Raman spectra

and facilitate the comparison of coupling strengths and non-planar deformations between different isoforms the absorption spectra were first measured in the Soret region for each of the three proteins at equal concentrations depicted in Figure 51.

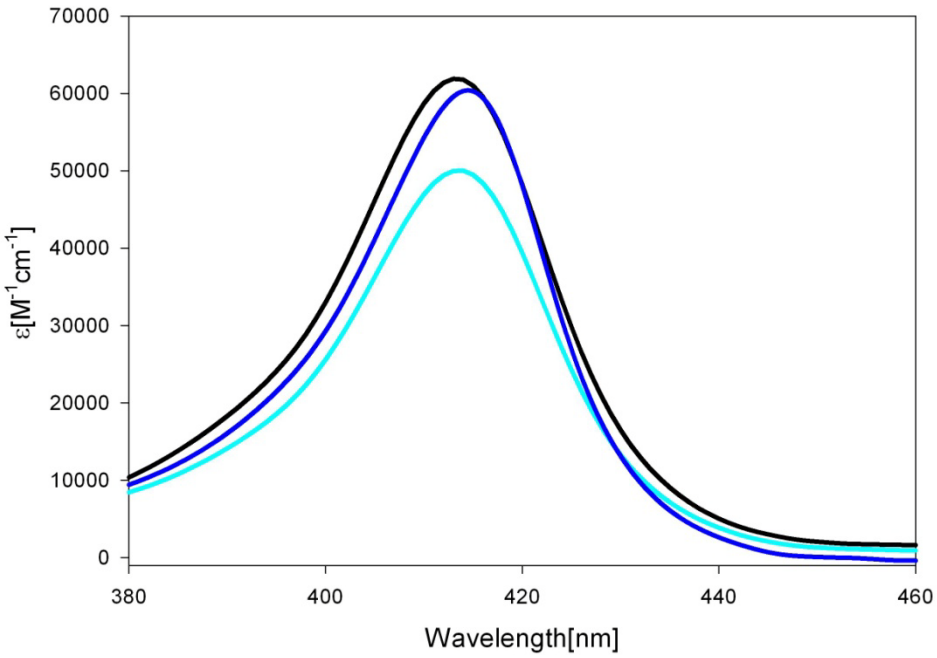


Figure 51: Taken from ref (*295*) and modified. Soret band absorption spectra of horse heart (black), yeast (light blue) and chicken (blue) ferrocytochrome c. All spectra were acquired with a protein concentration of 0.5mM and a tris buffer concentration of 0.01M. All spectra were corrected for background absorption.

The yc absorption spectrum is a factor of ~1.35 less intense than horse heart and chicken cytochromes c. To ensure that the Raman spectra of hhc and yc were affected equally with respect to self-absorption, we used a 1.35-fold higher concentration for measuring the Raman spectra of yc. The Raman spectra were then normalized on the internal standard Raman band of sodium perchlorate at 934cm^{-1} so that they could be compared with each other. To utilize the intensity ratios of corresponding bands of different cytochrome c species for a comparison of the corresponding vibronic coupling matrix elements, the polarized Raman spectra of yc were multiplied by a factor of 1.35^2

to correct for the different oscillator strength of yc on one side and hhc and chc on the other side. The necessity of this correction can be inferred from the equation:

$$R_r(l,k) = \frac{I_r(l)}{I_r(k)} = \left(\frac{\vec{\mu}_{gB}(l)}{\vec{\mu}_{gB}(k)}\right)^4 \left(\frac{c_{BB}^{\Gamma_r}(l)}{c_{BB}^{\Gamma_r}(k)}\right)^2 \qquad (53)$$

provided that the absorption band profiles are identical for the considered species. This is the case for hhc and yc; whereas those of hhc and chc are slightly different. We ignore this difference in the present study. $I_r(l)$ and $I_r(k)$ are the measured integrated intensities of the Raman active mode *r* of the two species *l* and *k*. $\vec{\mu}_{gB}(l), \vec{\mu}_{gB}(k)$ are the corresponding electronic dipole moments of the B-band transition and $c_{BB}^{\Gamma_r}(l), c_{BB}^{\Gamma_r}(k)$ are the corresponding vibronic coupling matrix elements of mode *r*. The matrix elements are linearly related to the Raman activity inducing oop deformations by the equation:

$$c_{BB}^{\Gamma_r} = \left\langle B \middle| \frac{\partial^2 \hat{H}_{el}}{\partial Q_r^{\Gamma_r} \partial Q_j^{\Gamma_j}} \middle| B \right\rangle Q_{01}^{\Gamma_r} \delta Q_j^{\Gamma_j} \qquad (54)$$

where $Q_r^{\Gamma_r}$ is the pure vibrational matrix element of vibration *r*. The other parameters in eq. 54 were defined in the theoretical background section. The factor

$$R_r'(l,k) = \sqrt{\frac{I_r(l)}{I_r(k)}} \cdot \left(\frac{\vec{\mu}_{gB}(k)}{\vec{\mu}_{gB}(l)}\right)^2 \qquad (55)$$

derived from the properly normalized spectra can be used to directly compare vibronic coupling matrix elements and thus the corresponding SNCDs. The normalized intensities, the depolarization ratios and the corresponding R_r' values (normalized on

the Raman intensities of horse heart cytochrome c) of the observed Raman bands are listed in Table 6, Table 7 and Table 8, respectively.

Raman band [cm^{-1}]	hhc	chc	yc
279(ν_9)	14.6	6.3	
303 (ν_{51})	37.7	30.4	66.9
344 (ν_8)	119.3	110.2	103.5
355 (ν_{50})	124.5	93.8	137.8
370($\delta_{prop,1}$)[a]	31.1	15.3	
378 ($\delta_{prop,2}$)[a]	61.1	53.6	120.1
390 ($\delta_{CCS,1}$)[b]	87.6	59.1	174.0
398 ($\delta_{CCS,1}$)[b]	60.8	52.7	56.3
411(δ_{CCC})[c]	99.1	90.1	406.4
420(δ_{CCC})[c]	38.9	20.5	18.9
446 (γ_{22})	103.7	44.0	258.6
462 (ν_{33b})	45.7	30.6	24.5
472 (ν_{33a})	38.8	7.3	95.0
518 (γ_{12})	17.0	10.1	
566 (γ_{21})	49.4	34.9	60.9
680 ($\nu_{cs,1}$)[d]	130.9	141.0	115.9
689 ($\nu_{cs,2}$)[d]	169.3	126.4	299.3
698 (ν_7)	71.9	63.9	89.3
721 (γ_{11})	12.2	13.5	10.6
727 (γ_5)	11.7	12.1	10.4
748 (ν_{15})	126.2	106.4	70.4
832 (γ_{10})	20.2	20.0	

Table 6: Taken from ref (*295*) and modified. Relative intensities*10^3 of low wavenumber Raman bands in the spectra of horse heart, chicken and yeast (C102T) ferrocytochrome c. The band assignment was obtained from Hu *et al*. (*294*). [a]deformation modes of propionic acid substituents; [b]deformation modes of thioether bridges; [c]deformation mode of methyl substituents, [d] CS stretching modes

An inspection of the intensities listed in Table 6 reveal that the B_{1u} modes, $\gamma_{10\text{-}12}$, and the E_g modes, γ_{21} and γ_{22}, are relatively intense, indicating ruffling and to some extent waving as the major contributions to the oop deformations of the heme group. This finding qualitatively correlates well with NSD results obtained from crystal structures and MD simulations. The intensities of the γ_5 doming mode are slightly higher for chc and hhc compared to yc. The band of the ruffling mode, γ_{11}, is slightly more intense for

chc and hhc than for yc, whereas the intensity of the ruffling modes, γ_{10} and γ_{12}, are only comparable for hhc and chc whereas chc is more intense then hhc. However, the intensities of the latter ruffling bands are very weak in all spectra. The determination is therefore uncertain, owing to its strong dependence on baseline variations. The ruffling mode, γ_{10}, was obscured in the spectrum of yc due to interference with glass peaks.

The Raman intensity of oop modes of A_{1u}, A_{2u}, B_{1u} and B_{2u} symmetry can be directly used as a measure of the corresponding Raman activity inducing non-planar deformations (by using eq. 55), if the depolarization ratios of these bands were close to 0.125 (*298*), because that would indicate that only one symmetry type of deformations contribute to the Raman tensor. The respective values listed in Table 7 clearly show that the depolarization ratios of out-of-plane bands obtained from the decomposition of the spectra seen in Figure 49 deviate, sometimes substantially, from 0.125, indicating that the Raman tensor is a mixture of the McClain tensors in eq. 39.

Raman band [cm^{-1}]	hhc	chc	yc
279(ν_9)	0.14	0.13	
303 (ν_{51})	0.12	0.16	0.24
344 (ν_8)	0.13	0.14	0.10
355 (ν_{50})	0.12	0.15	0.09
370($\delta_{prop,1}$)[a]	0.14	0.14	
378 ($\delta_{prop,2}$)[a]	0.12	0.14	0.11
390 ($\delta_{CCS,1}$)[b]	0.13	0.15	0.16
398 ($\delta_{CCS,1}$)[b]	0.12	0.16	0.19
411(δ_{CCC})[c]	0.14	0.16	0.22
420(δ_{CCC})[c]	0.13	0.21	0.57
446 (γ_{22})	0.17	0.23	0.46
462 (ν_{33b})	0.31	0.37	0.87
472 (ν_{33a})	0.04	0.33	0.71
518 (γ_{12})	0.10	0.10	
566 (γ_{21})	0.13	0.15	0.18
680 ($\nu_{cs,1}$)[d]	0.15	0.13	0.16
689 ($\nu_{cs,2}$)[d]	0.13	0.14	0.13
698 (ν_7)	0.14	0.13	0.16
721 (γ_{11})	0.20	0.09	0.18
727 (γ_5)	0.25	0.11	0.32
748 (ν_{15})	0.30	0.23	0.66
832 (γ_{10})	0.42	0.22	

Table 7: Taken from ref (*295*) and modified. Depolarization ratio of low wavenumber Raman bands in the spectra of horse heart, chicken and yeast (C102T) ferrocytochrome c. The band assignment was obtained from Hu *et al*. (*294*). [a]deformation modes of propionic acid substituents; [b]deformation modes of thioether bridges; [c]deformation mode of methyl substituents, [d] CS stretching modes.

The deviations can be rationalized in terms of an admixture of contributions from terms other than A_{1g} in the second term on the right side of eq. 43. Examples of the effects that can cause these deviations from canonical depolarization ratios of A_{1g} symmetry were detailed in the Methods section. The increased depolarization ratio of the A_{2u} doming mode γ_5 was rationalized by a ruffling (B_{1u}) or saddling (B_{2u}) deformation that would admix B_{2g} or B_{1g} vibronic coupling, respectively to the totally symmetric A_{1g} contribution to the Raman tensor. The depolarization ratios of the Raman lines of B_{1u}-type modes (γ_{10}, γ_{11}, the depolarization ratio of γ_{12} is uncertain owing to the low

intensity of the line) are also substantially larger than 0.125. The increase in the depolarization ratios suggest an admixture of vibronic coupling of $A_{2u} \otimes B_{1u} = B_{2g}$ or $A_{1u} \otimes B_{1u} = B_{1g}$ symmetry to the Raman tensor, thus indicating the presence of doming and/or propellering deformations. The depolarization ratios of E_g modes can vary between 0.125 and 0.75, so that nothing specific can be derived from the values observed for γ_{21} and γ_{22}. The depolarization ratios for the modes γ_5 and γ_{11} are all substantially larger for yc than for hhc, but smaller for chc. We invoke the formalism of Huang and Schweitzer-Stenner (*298*) (eqs. (18)-(20) in their paper) to derive the following equation for the relationship between $\delta Q_j^{\Gamma_j}(l) / \delta Q_j^{\Gamma_j}(k)$, R_r' (*l*,*k*) and the depolarization ratio of the mode *r* in species *l* and *k*:

$$\xi_A^{\Gamma_j} = \frac{\delta Q_j^{\Gamma_j}(l)}{\delta Q_j^{\Gamma_j}(k)} = \frac{c_{BB}^{A_{1g}}(l)}{c_{BB}^{A_{1g}}(k)} = \sqrt{R_r'(l,k)} \cdot \left[\frac{1+\sqrt{\frac{8\rho_r(k)-1}{3-4\rho_r(k)}}}{1+\sqrt{\frac{8\rho_r(l)-1}{3-4\rho_r(l)}}} \right] \qquad (56)$$

where Γ_j is the symmetry of the considered out-of-plane mode in D_{4h}, $c_{BB}^{A_{1g}}$ denotes the vibronic coupling matrix element for which the oop-mode and deformation have the same symmetry and $\rho_r(k)$, $\rho_r(l)$ are the depolarization ratios for the cytochrome species *k* and *l*. Normalized intensity ratios, R_r' are seen in Table 8.

Raman band	yc/hhc	chc/hhc
279(ν_9)		0.44
303 (ν_{51})	1.77	0.81
344 (ν_8)	0.87	0.92
355 (ν_{50})	1.11	0.75
370($\delta_{prop,1}$)[a]		0.56
378 ($\delta_{prop,2}$)[a]	1.97	0.88
390 ($\delta_{CCS,1}$)[b]	1.99	0.68
398 ($\delta_{CCS,1}$)[b]	0.93	0.87
411(δ_{CCC})[c]	4.10	0.91
420(δ_{CCC})[c]	0.49	0.53
446 (γ_{22})	2.49	0.42
462 (ν_{33b})	0.54	0.67
472 (ν_{33a})	2.45	0.21
518 (γ_{12})		0.60
566 (γ_{21})	1.23	0.71
680 ($\nu_{cs,1}$)[d]	0.89	1.08
689 ($\nu_{cs,2}$)[d]	1.77	0.75
698 (ν_7)	1.24	0.89
721 (γ_{11})	0.87	1.11
727 (γ_5)	0.89	1.04
748 (ν_{15})	0.56	0.84
832 (γ_{10})		0.99

Table 8: Taken from ref (*295*) and modified. Normalized Raman intensity ratios of bands assigned from Hu *et al.* (*294*). Normalized Raman intensities of yc and chc were normalized on hhc. These values were termed R_r'. [a]deformation modes of propionic acid substituents; [b]deformation modes of thioether bridges; [c]deformation mode of methyl substituents, [d] CS stretching modes.

In the following we again use hhc as standard to calculate $\xi_A^{\Gamma_j}$. By inserting the respective values for the γ_{11} mode of hhc and yc, we obtain $\xi_A^{B_{1u}}$ = 0.98, which suggests heme ruffling is nearly identical in both species. For chc, however, the $\xi_A^{B_{1u}}$ value suggests an increase of ruffling by 1.4 compared to hhc. Intensities and depolarization ratios of γ_5 can be analyzed in the same way. Compared again with hhc, this analysis yields a decrease of doming by 0.83 in yc and an increase by 1.54 for chc. For the comparison of hhc and chc, γ_{10} and γ_{12} modes were considered as well. The

corresponding intensity ratios of these modes suggest that the heme is slightly more ruffled (by a factor of 1.1) in chc than in hhc.

Next we use eq. (18) of Huang and Schweitzer-Stenner (*298*) to derive the corresponding ratios of the Jahn-Teller coupling parameters c_{BB}^{B}:

$$\xi_B^{\Gamma_i} = \frac{c_{BB}^{B}(l)}{c_{BB}^{B}(k)} = \frac{c_{BB}^{A_{1g}}(l)}{c_{BB}^{A_{1g}}(k)} \cdot \left[\frac{\sqrt{\frac{8\rho_r(l)-1}{3-4\rho_r(l)}}}{\sqrt{\frac{8\rho_r(k)-1}{3-4\rho_r(k)}}} \right] \tag{57}$$

Thus, we obtained that B-type coupling of γ_{11} is by a factor 0.83 smaller in yc than in hhc, whereas it is by a factor 0.35 smaller in chc. For this mode, this coupling type can arise from doming ($B_{1u} \otimes A_{2u} = B_{2g}$) or propellering ($B_{1u} \otimes A_{1u} = B_{1g}$). For yc, the identity of the scaling factors $\xi_A^{A_{2u}}$ for the A-type coupling of γ_5 and $\xi_B^{B_{1u}}$ the B-type factor of γ_{11} suggests that doming, rather than propellering, is responsible for the latter. For chc, however, the factors for A-type coupling of γ_5 and B-type coupling of γ_{11}, point into opposite directions, so that the latter must be assigned to propellering deformations. For γ_5, we obtained $\xi_B^{B_{1u}}$ =1.11 and 0.28 for yc and chc, respectively. Here, B-type Jahn-Teller coupling can arise from ruffling ($B_{1u} \otimes A_{2u} = B_{2g}$) or saddling ($B_{2u} \otimes A_{2u} = B_{1g}$). The ruffling option can be ruled out from the above discussed properties of γ_{11}, hence, we estimated an increase in saddling by a factor 1.11 for yc and

a decrease of this deformation by a factor 0.28 for chc. The same B-type coupling analysis was again performed on γ_{10} and γ_{12} to compare chc and hhc. For chc we thus obtained factors of 0.73 and 0.84 smaller, respectively. This result is therefore attributed to the change in propellering deformation. Therefore, the combined use of the changes of intensities and depolarization ratios affords us a comparison of the non-planar heme deformations in the investigated ferrocytochrome c species. Doubtless, the obtained numbers are subject to substantial statistical errors (~25% might be a realistic estimation) owing to the uncertainty of the intensity determination particularly for the weak y-components. We like to reemphasize the simplicity of the analysis, which assumes that the normal mode compositions and the pure electronic parts of vibronic coupling of γ_{11} and γ_5 are identical for the three cytochrome c proteins.

The observed intensity ratios of the waving modes, γ_{21} and γ_{22}, suggest the hierarchy yc > hhc >chc. γ_{22} exhibits somewhat more pronounced intensity variation than γ_{21}, but this might arise from the admixture of substituent modes to the eigenvector of the former, which would make the spectral changes difficult to interpret.

5.5.3 Intensities and Depolarization Ratios of In-plane and Peripheral Substituent Modes

Generally, Soret band excitation of heme groups yield a low wavenumber spectrum dominated by the intensities of the A_{1g}-type modes ν_7, ν_8 and ν_9. While the bands of these modes are still clearly identifiable in the spectrum of ferricytochrome c, other modes exhibit comparable intensities. This consideration particularly concerns the modes of peripheral substituents, e.g. the deformation modes of the propionic acids at

370 and 378 cm^{-1}, the CCS deformation modes of the thioether bridges at 390 and 398 cm^{-1}, and the CS stretching modes of the same entity at 680 and 689 cm^{-1}. The resonance enhancement of the Raman scattering of substituent modes certainly results from their mixing with macrocycle modes in a lower symmetry than D_{4h}, because the internal coordinates of the substituents are certainly not coupled to $\pi \rightarrow \pi^*$ transitions of the macrocycle. This mixing very much complicates the interpretation of intensities and depolarization ratios of ν_7 and ν_8. In view of the substantial mixing between these macrocycle modes and substituents modes reported for various types of Ni(II)-porphyrins (*291, 296*), it is very likely that all the modes of the relatively intense bands in the 350-400 cm^{-1} and the 650-700 cm^{-1} region exhibit some ν_8 and ν_7 character, respectively. A further complication arises from the fact that these two modes can mix with out-of-plane modes in the presence of non-planar deformations. Generally, all these mixing effects should reduce the Raman intensity of involved in-plane modes. That this result occurs has been demonstrated for ν_7 by Kitagawa et al. who showed that heme doming decreases the Raman intensity of this mode (*299*). The higher intensity of the ν_7 mode in the spectrum of yc (*R'*= 1.24) and the slightly lower one in the spectrum of chc (*R'*=0.83) correlate in fact nicely with the reduced ruffling of the heme in the former and the increased heme ruffling in the latter, as deduced from our analysis of oop modes. The intensity ratios, R_r, for the δ_{CCS} and ν_{CS} modes indicate that the orientation of the thioether linkages with respect to the heme macrocycle are different in hhc and yc, yielding different degrees of vibrational mixing with ν_8 and ν_7, respectively. It is reasonable to assume that the degree of mixing (and thus the Raman

intensity) increases with a decreasing tilt between the substituent group and the heme. The observed R_r' values for substituent modes of the thioether linkages 1 and 2 therefore suggest that the tilt of the former is larger in yc than in hhc, while the very opposite notion applies to the latter.

The depolarization ratios of ν_7 and ν_8 do not deviate much from the D_{4h} expectation value for A_{1g}-modes for hhc and chc, as it is expected for B-state scattering (*291*). For yc, the depolarization ratio of ν_7 is somewhat larger, but this is difficult to interpret, owing to these modes vibrationally mixing with substituent modes (CH_3-deformation in addition to CS (*291, 296*)). The larger depolarization ratio might result from an increase of some of the products according to the second order contribution to the expanded vibronic coupling operator, $\delta Q_j^{B_{1u}} \delta Q_{j'}^{A_{1u}}$, $\delta Q_j^{B_{1u}} \delta Q_{j'}^{A_{2u}}$, $\delta Q_j^{B_{2u}} \delta Q_{j'}^{A_{1u}}$,and $\delta Q_j^{B_{1u}} \delta Q_{j'}^{A_{2u}}$, which would admix either B_{1g} or B_{2g} to the Raman tensor. Changes of substituent orientations can vary in-plane deformations, but the depolarization ratios of the ν_4-mode (data not shown) of hhc (0.18) and yc (0.15) indicate that rhombic B_{1g}-deformations are smaller in yc than in hhc (*221*).

Another peculiarity of the low wavenumber spectra of ferrocytochrome c derivatives is the rather strong intensity of the B_{1g}-type mode, ν_{15} (*300*). This is somewhat unexpected because intrastate Jahn-Teller coupling is generally much weaker than Franck-Condon coupling, so that B_{1g}- and B_{2g}-type modes exhibit only modest intensity in heme and porphyrin spectra taken with Soret band excitation (*301*). As shown in Figure 50 and as reflected by the corresponding R_r-value in Table 6 the intensity of ν_{15} is substantially

lower for yc than for hhc and chc. Interestingly, the depolarization ratio of this mode is exceptionally low for hhc and chc (0.3 and 0.23), whereas it is quite normal for yc (0.66). These values clearly suggest that the large intensity of ν_{15} is actually produced by Franck-Condon coupling. One is tempted to interpret our results as suggesting vastly different in-plane B_{1g}-type deformations of the heme group, which would add a $B_{1g} \otimes B_{1g} = A_{1g}$ contribution to the Raman tensor. This interpretation would imply that ν_{15} is much more sensitive to changes of this deformation than ν_4 (we estimated a 'sensitivity enhancement' of 1.8 by using eq. (18) of Huang and Schweitzer-Stenner (*298*)). Alternatively, one might consider an admixture of A_{1g}-type coupling by two non-planar deformations, $\delta Q_i^{\Gamma_i}$ and $\delta Q_j^{\Gamma_j}$ via the yet unconsidered coupling term (*225, 226*):

$$c_{BB}^{\Gamma_r\,\prime} = \left\langle B \left| \frac{\partial^3 \hat{H}_{el}}{\partial Q_r^{\Gamma_r} \partial Q_i^{\Gamma_i} \partial Q_j^{\Gamma_j}} \right| B \right\rangle Q_{01}^{\Gamma_r} \delta Q_i^{\Gamma_i} \delta Q_j^{\Gamma_j} \qquad (58)$$

However, the deformations derived from X-ray structures as well as from the MD simulations discussed below and a comparison with earlier studies on the effect of non-planar deformations on depolarization ratios (*244, 298*) suggest that only the combination of the dominant deformations ruffling and saddling (B_{1u} and B_{2u}) would have an impact, but this would give rise to antisymmetric scattering, which is detectable only with Q-band excitation. Optical absorption studies have shown that the Q-band splitting is much smaller for yc than for hhc, which reflects, to a significant extent, a reduction of electronic B_{1g} perturbations for yc. However, this variation of the

electronic perturbation is unlikely to have a big affect on depolarization ratios, which mostly reflect vibronic perturbations (*211*).

The analysis reported here and similar analyses in earlier studies (*220, 225, 226, 241, 245*) have been based on one very important, though somewhat untested assumption, namely that the considered deformations of the heme macrocycle do not significantly affect the eigenvector of the considered normal mode, as long as the π-electron ring is not seriously disrupted, as is the case in hydroporphyrins (*296*). Normal mode analyses of asymmetric porphyrins are rare in the literature, in spite of their biological relevance. One of the few, however, i.e. the normal mode analysis of Ni(II)-octaethylporphyrins (Ni-OEP) with asymmetric NO_2 substitutions at *meso*-positions by Lemke *et al.*, indicate that the above assumption might be of limited validity (*297*). If one compares the eigenvector pattern of the modes ν_2, ν_3, ν_4 and ν_8 of Ni(5,15-NO_2-OEP) reported in their paper, one obtains that the pattern of ν_8 is being heavily affected by the *meso*-substitution, whereas the influence on the other A_{1g}-type modes is limited and hardly discernable from the respective normal mode patterns. This porphyrin molecule is certainly not a good model system for heme c, but the study clearly shows that asymmetric perturbations can change the eigenvectors of macrocycle modes quite dramatically. The results of Lemke *et al.* further suggest that this effect might be more likely for low wavenumber modes and that a similar perturbation could be responsible for the large depolarization ratio variations of ν_{15}. It is thinkable that the specific B_{1g} deformations which are operative in cytochrome c lead to a mixture of ν_7 and ν_{15}

coordinates, thus causing Franck-Condon activity of the latter. As a matter of fact, both modes would exhibit an A-symmetry in a C_{2v} point group, which would explain why ν_{15} is a very sensitive indicator of B_{1g}-deformations, at least for cytochrome c.

Other in-plane modes such as ν_{50} and ν_{51} (E_u) are present, indicating the presence of E_u type deformation. E_u modes are Raman inactive in D_{4h} symmetry, but become Franck-Condon active e.g. with an electric field exhibiting the same symmetry (*224*). The E_u mode, ν_{50}, is of comparable intensity in all three species with yc having the largest, though the ν_{51} mode is double the intensity in yc compared to hhc and chc. The larger intensities of the yc E_u Raman bands indicate that the electric field of the heme, which is weaker in yeast than in hhc cytochrome c (*212*), is not the only source of E_u-type deformations.

5.5.4 Comparison with NSD Results[4]

While x-ray crystallography gives accurate results for comparison to a static conformation, they might not be representative of a dynamic average of the solution phase protein. A direct comparison of results obtained from the NSD analysis of crystal structures and structures derived from MD simulations reveals notable differences.

Protein	Type	D_{oop}	B_{2u}	B_{1u}	A_{2u}	E_{gx}	E_{gy}	A_{1u}
HHC	Crystal	1.062	-0.224	1.004	-0.020	0.122	0.161	-0.022
YC	Crystal	0.892	-0.306	0.771	0.008	0.137	0.196	-0.102
HHC	MD	0.7867	-0.001	-0.765	-0.034	0.065	0.028	-0.167
CHC	MD	0.7899	-0.277	-0.670	0.074	0.229	0.170	-0.106
YC	MD	0.9775	-0.132	-0.927	0.087	0.034	0.237	-0.116

Table 9: Taken from ref (*295*) and modified. Oop results of NSD analysis from crystal structures from Shelnutt`s website (*302*) and MD simulations completed by Dr. Monique Laberge.

[4] The NSD results for the hemes extracted from MD simulations were provided by Dr. Monique Laberge in the Department of Biology, Center for Research in Molecular Modeling, Concordia University, Montreal, H3G 1M8, Canada.

NSD results obtained for crystal structures (Table 9) indicate a dominant ruffling with moderate saddling and waving deformations for hhc (1hrc)(*231*) and yeast (1ycc)(*227*) with hhc having ~0.25 Å larger ruffling deformation, 0.11 Å smaller saddling deformation, ~0.04 Å smaller doming, 0.09 Å smaller propellering, and 0.15 Å smaller total waving deformation. The respective ratios of deformation ($R^{\Gamma}=\Delta^{\Gamma}$*(yeast)/* Δ^{Γ}*(hhc)*) are 0.77 for $\Gamma=B_{1u}$ (ruffling), 1.38 for $\Gamma=B_{2u}$ (saddling), 2.36 for $\Gamma=A_{2u}$ (doming), 4.9 for $\Gamma=A_{1u}$ (propellering), and 1.3 for $\Gamma=E_g$ (waving). For ruffling we derived nearly identical heme deformations for yc and hhc, though the analysis is limited to only one oop Raman band, namely, γ_{11}. Since the difference between corresponding deformations of yc and hhc is not large, one might still consider the discrepancy between R^{B1u} =0.77 and $\xi_A^{B_{1u}}$ =0.98 as not particularly significant. For saddling, the above R^{Γ}-values are nearly in quantitative agreement with the corresponding $\xi_B^{B_{1u}}$ - values of γ_5, i.e. 1.11 for yc. The stronger waving deformation in yc is also predicted from our results. With respect to doming and propellering, the X-ray structure and our results do not agree, since crystal structures indicate that these deformation types are more pronounced in yc. However, one has to take into consideration that doming and propellering deformations are rather small compared with ruffling and saddling and therefore subject to larger errors in the NSD analysis. Apparently, the use of resonance Raman spectroscopy provides us with a very sensitive tool for probing even very small non-planar deformations.

Our NSD analysis of the hemes extracted from MD simulations (Table 9) indicate that the ruffling deformation of yeast cytochrome c is nearly 0.16 Å larger than the

corresponding hhc deformation and even 0.25 Å larger than the ruffling deformation in chc, which is not in agreement with our results. All three investigated B_{1u} bands were more intense for chc compared to hhc, thus providing a consistent picture suggesting that the heme of chc is more ruffled than hhc. With respect to doming, our results are at variance with the heme deformation of the MD structure of yc being larger than that of hhc ($\xi_A^{A_{2u}}$ = 0.83 compared with $R^{B1u}=\Delta^{B1u}(yc/\ \Delta^{B1u}(hhc)$=1.46), whereas quantitative agreement is obtained between our result (i.e. $\xi_A^{A_{2u}}$ =1.54) and the ratio $R^{A2u}=\Delta^{A2u}(chc)/\ \Delta^{A2u}(hhc)$=1.49 obtained from the corresponding MD structures of chc and hhc. These structures suggest that the heme groups of yc and chc are much more saddled than the active site of hhc; this is definitely not in agreement with our results. Concerning propellering, however, our results agree with the NSD values for chc and hhc, in that both suggest a much smaller deformation for the former ($\xi_B^{B_{1u}}$ = 0.35 compared with $R^{B1u}=\Delta^{B1u}(yc/\ \Delta^{B1u}(hhc)$=0.64). Altogether, the above analysis suggests that the ratios of non-planar heme deformations inferred from resonance Raman bands of oop modes are in reasonable agreement with the corresponding values of the predominant deformations (ruffling and saddling) obtained from X-ray-structures (*227, 231*). For smaller deformations, i.e. doming and propellering, our results are much closer to the structural properties inferred from MD simulations. The latter generally yield larger doming and propellering distortions than exhibited by the respective X-ray structures. This result, combined with the fact that these deformations have a detectable influence on the Raman cross section of vibrational modes, leads us to propose (a) that structures from MD simulations provide a more reliable picture for smaller deformations while

failing to account for the correct magnitude of the dominant distortions and (b) that Raman spectroscopy is a very suitable tool (even more than X-ray) to actually probe and explore these deformations.

Recently, Levantino *et al.* (*221*) reported that yc might be more ruffled than hhc, based on results from their analysis of vibronic Q-band splitting in horse heart and yeast ferrocytochrome c. However, these authors derived their conclusion from values which they obtained from vibronic matrix elements:

$$c_{BB}^{A_{1g}\,\prime} = \left\langle Q \middle| \frac{\partial^3 \hat{H}_{el}}{\partial Q_r^{A_{2g}} \partial Q_i^{B_{1u}} \partial Q_j^{B_{2u}}} \middle| Q \right\rangle Q_{01}^{A_{2g}} \delta Q_i^{B_{1u}} \delta Q_j^{B_{2u}} \tag{59}$$

which describes an admixture of A_{1g}-symmetry to the vibronic pattern of modes of A_{2g}-symmetry, owing to the combined effect of ruffling and saddling. This matrix element, which Levantino *et al*. found to be larger for yc than for hhc for 3 A_{2g}-type modes, is approximately proportional to the product of both deformations, which are nearly identical for both species in the X-ray structure. An additional, yet overlooked contribution also producing an A_{1g}-type contribution can result from doming and propellering, which, as argued above, might be larger than assumed. We can therefore conclude that the coupling parameters reported by Levantino *et al.* do not contradict results obtained from our Raman data and therefore do not argue against the use of X-ray structures for NSD analyses to describe predominant deformations. The question of the NSD discrepancies resulting from the use of hemes extracted from X-ray and MDS deserves further comment. The distortions of hemes extracted from X-ray structures

have an inherent uncertainty induced by the crystal packing forces (*231*), and one way to relax the models is to subject them to MDS under explicit solvation conditions. Simulations reported here were carried out for 1 ns, a timeframe sufficient to equilibrate such a small size protein. That the trajectory average heme structures subjected to NSD analysis yield results –for the smaller distortions– that are different from that of the X-ray structures is not surprising, as they reflect the dynamical heterogeneity of the protein conformational space. In this view, protein matrix dynamics would fine-tune heme properties (and function) by modulating the extent of the smaller heme deformations.

Chapter 6 Conclusions

We have shown how lower resolution techniques such as CD, absorption and Raman spectroscopy can be used to quantitatively probe tertiary and secondary structure changes as well as changes in the active site environment of cytochrome c due to these fluctuations. Overall, we were able to quantitatively estimate planar and non-planar deformations of ferrocytochrome c isoforms in a relatively time and cost effective way, where X-ray crystallography or NMR spectroscopy would have been excessive. We discovered the combined use of absorption and CD spectroscopy as a tool to explore and differentiate between electronic and vibronic perturbations of the heme and determine the electric field of the heme pocket in the heme plane. We also categorized the role of certain solution conditions in their connection to structural effects and thus, alterations of spectroscopic data, identified and characterized conformational substates of cytochrome c and discovered a thermodynamic intermediate of the alkaline transition. Also, we used some of this knowledge to structurally and thermodynamically distinguish between a thermally populated state and an alkaline state which were believed to be equal. As well, we further characterized the states adopted in the alkaline transition of ferricytochrome c. These results offer a baseline on which to build further estimates of the conformational changes that cytochrome c undergoes in relation to its biological processes. Recently, much attention has been focused on the binding of cytochrome c to artificial membranes which represents this protein's interactions with the intermembrane space of the mitochondria. The inner mitochondrial membrane has a large fraction of cardiolipin (CL), a negatively charged phospholipid. Upon binding to the inner membrane space of the mitochondria,

cytochrome c acts as a peroxidase, selectively catalyzing CL peroxidation. Cardiolipin peroxidation is a prerequisite for initiation of apoptosis. By characterizing the changes in tertiary and secondary structure upon cytochrome c binding to liposomes containing cardiolipin, a model can be created that represents this protein's conformational changes in the initial stages of the apoptotic process.

References

1. **Anfinsen, C. B., Haber, E., Sela, M., and White, F. N. (1961) Kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, *Proc. Natl. Acad. Sci. USA 47*, 1309-1314.**

2. **Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. Phys.-Chim. Biol. 65*, 44-45.**

3. **Frauenfelder, H., Sligar, S. G., and Wolynes, P. G. (1991) The Energy Landscapes and Motions of Proteins, *Science 254*, 1598-1603.**

4. **Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, Pathways, and the Energy Landscape of Protein Folding - A Synthesis, *Proteins-Structure Function and Genetics 21*, 167-195.**

5. **Radford, S. E., and Dobson, C. M. (1999) From computer simulations to human disease: Emerging themes in protein folding, *Cell 97*, 291-298.**

6. **Thomas, P. J., Qu, B. H., and Pedersen, P. L. (1995) Defective Protein-Folding as a Basis of Human Disease, *Trends Biochem. Sci. 20*, 456-459.**

7. **Dobson, C. M. (1999) Protein misfolding, evolution and disease, *Trends Biochem. Sci. 24*, 329-332.**

8. **Horwich, A. (2002) Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions, *J. Clin. Invest. 110*, 1221-1232.**

9. **Chou, P. Y., and Fasman, G. D. (1974) Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins, *Biochemistry 13*, 211-222.**

10. **DeBartolo, J., Colubri, A., Jha, A. K., Fitzgerald, J. E., Freed, K. F., and Sosnick, T. R. (2009) Mimicking the folding pathway to improve homology-free protein structure prediction, *Proc. Natl. Acad. Sci. USA 106*, 3734-3739.**

11. **DeBartolo, J., Hocky, G., Wilde, M., Xu, J. B., Freed, K. F., and Sosnick, T. R. (2010) Protein structure prediction enhanced with evolutionary diversity: SPEED, *Protein Sci. 19*, 520-534.**

12. **Liwo, A., Arlukowicz, P., Czaplewski, C., Oldziej, S., Pillardy, J., and Scheraga, H. A. (2002) A method for optimizing potential-energy functions by a hierarchical design of the potential-energy landscape: Application to the UNRES force field, *Proc. Natl. Acad. Sci. U. S. A. 99*, 1937-1942.**

13. **Taylor, J. P., Hardy, J., and Fischbeck, K. H. (2002) Biomedicine - Toxic proteins in neurodegenerative disease, *Science 296*, 1991-1995.**

14. **Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M., and W., B. D. (1991) Calorimetric determination of the enthalpy change for the alpha-helix to coil transition of an alanine peptide in water, *Proc. Natl. Acad. Sci. USA 88*, 2845.**

15. **Richardson, J. M., Lopez, M. M., and Makhatadze, G. I. (2005) Enthalpy of helix-coil transition: Missing link in rationalizing the thermodynamics of helix-forming propensities of the amino acid residues, *Proc. Natl. Acad. Sci. U. S. A. 102*, 1413-1418.**

16. **Myers, J. K., and Pace, C. N. (1996) Hydrogen bonding stabilizes globular proteins, *Biophys. J. 71*, 2033-2039.**

17. **Kauzmann, W. (1959) Some Factors in the Interpretation of Protein Denaturation, *Adv. Prot. Chem. 14*, 1-63.**

18. **Klotz, I. M., and Franzen, J. S. (1962) Hydrogen Bonds Between Model Peptide Groups in Solution, *J. Am. Chem. Soc. 84*, 3461-3466.**

19. **Susi, H., and Ard, J. S. (1969) Hydrophobic Interactions and Hydrogen Bonding of Epsilon Caprolactam in Aqueous Solution, *J. Phys. Chem. 73*, 2440-2441.**

20. **Baskakov, I., and Bolen, D. W. (1998) Forcing thermodynamically unfolded proteins to fold, *J. Biol. Chem. 273*, 4831-4834.**

21. **Henkels, C. H., and Oas, T. G. (2005) Thermodynamic characterization of the osmolyte- and ligand-folded states of Bacillus subtilis ribonuclease P protein, *Biochemistry 44*, 13014-13026.**

22. **Mello, C. C., and Barrick, D. (2003) Measuring the stability of partly folded proteins using TMAO, *Protein Sci. 12*, 1522-1529.**

23. **Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding, *Proc. Natl. Acad. Sci. U. S. A. 103*, 16623-16633.**

24. **Ramachandran, G. N., Ramachandran, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations, *J. Mol. Biol. 7*, 95-99.**

25. **Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, Wiley & Sons, New York.**

26. **Hovmöller, S., Zhou, T., and Ohlson, T. (2002) Conformations of amino acids in proteins, *Acta Cryst. D58*, 768-776.**

27. **Shi, Z., Shen, K., Liu, Z., and Kallenbach, N. R. (2006) Conformation in the Backbone in Unfolded Proteins, *Chem. Rev. 106*, 1877-1897.**

28. **Tanford, C. (1968) Protein denaturation, *Adv. Protein Chem. 23*, 121-282.**

29. **Munoz, V., and Serrano, L. (1994) Intrinsic secondary structure propensities of the amino acids, using statistical phi-psi matrixes: comparison with experimental scales, *Proteins 20*, 301-311.**

30. **Swindells, M. B., MacArthur, M. W., and Thornton, J. M. (1995) Intrinsic ϕ,ψ propensities of amino acids, derived from the coil regions of known structures, *Nat. Struct. Biol. 2*, 596-603.**

31. **Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) Analysis of Main Chain Torsion Angles in proteins: Predicition of NMR coupling Constants for native and Random Coil Conformations, *J. Mol. Biol. 255*, 494-506. .**

32. **Serrano, L. (1995) Comparison between the ϕ-distribution of the amino acids in the Protein Data Base and NMR data indicates that amino acids have various ϕ propensities in the random coil conformation, *J. Mol. Biol. 254*, 322-333.**

33. **O'Connell, T. M., Wang, L., Tropsha, A., and Hermans, J. (1999) The "random-coil" state of proteins: comparison of database statistics and molecular simulations, *Proteins 36*, 407-418.**

34. **Dyson, H. J., and Wright, P. E. (1993) Peptide Conformation and Protein Folding, *Current Opinion in Structural Biology 3*, 60-65.**

35. **Rucker, A. L., and Creamer, T. P. (2002) Polyproline II helical structure in protein unfolded states: Lysine peptides revisited, *Protein Sci. 11*, 980-985.**

36. **Syme, C. D., Blanch, E. W., Holt, C., Ross, J., Goedert, M., Hecht, L., and Barron, L. D. (2002) A Raman optical activity study of rheomorphism in caseins, synucleins and tau New insight into the structure and behaviour of natively unfolded proteins, *Eur. J. Biochem. 269*, 148-156.**

37. **Brant, D. A., and Flory, P. J. J. (1965) The Configuration of Random polypeptide Chains. II. theory, *J. Am. Chem. Soc. 87*, 2791-2800.**

38. **Tanaka, S., and Scheraga, H. A. (1976) Statistical mechanical treatment of protein conformation. II. A three-state model for specific-sequence copolymers of amino acids, *Macromolecules 9*, 150-167.**

39. **Zimm, B. H., and Bragg, J. K. (1959) Theory of the phase transition between helix and random coil, *J. Chem. Phys. 31*, 526.**

40. **Lifson, S., and Roig, A. (1961) On the Theory of Helix-Coil Transition in Polypeptides, *J. Chem. Phys. 34*, 1963.**

41. **Tiffany, M. L., and Krimm, S. (1968) New chain conformations of poly(glutamicacid) and polylysine, *Biopolymers 6*, 1767-1770.**

42. **Woody, R. W. (2009) Circular Dichroism Spectrum of Peptides in the Poly(Pro)II Conformation, *J. Am. Chem. Soc. 131*, 8234.**

43. **Tiffany, M. L., and Krimm, S. (1969) Circular Dichroism of the "Random" Polypeptide Chain, *Biopolymers 8*, 347-359.**

44. **Cowan, P. M., and McGavin, S. (1955) Structure of Poly-L-Proline, *Nature 176*, 501-503.**

45. **Dukor, R., and Keiderling, T. A. (1991) Reassessment of the Random Coil Conformation: Vibrational CD Study of Proline Oligopeptides and Related Polypeptides, *Biopolymers 31*, 1747-1761.**

46. **Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls, J. E., and Creamer, T. P. (2003) Host-guest scale of left-handed polyproline II helix formation, *Proteins: Structure, Function and Genetics 53*, 68-75.**

47. **Shi, Z., Olson, C. A., Rose, G. D., Baldwin, R. L., and Kallenbach, N. R. (2002) Polyproline II structure in a sequence of seven alanine residues, *Proc. Natl. Acad. Sci. U. S. A. 99*, 9190-9195.**

48. **Hagarman, A., Measey, T. J., Mathieu, D., Schwalbe, H., and Schweitzer-Stenner, R. (2010) Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I' band profiles and NMR scalar coupling constants, *J. Am. Chem. Soc. 132*, 540-551.**

49. **Han, W.-G., Jakanen, K. J., Elstner, M., and Suhai, S. (1998) Theoretical study of aqueous N-Acetyl-L-alanine N_-methylamide: structures and Raman, VCD, and ROA spectra, *J. Phys. Chem. B 102*, 2587-2602.**

50. **Poon, C. D., Samulsi, E. T., Weise, C. F., and Weisshaar, J. C. (2000) Do Bridging Water Molecules Dictate the Structure of a Model Dipeptide in Aqueous Solution?, *J. Am. Chem. Soc. 122*, 5612-5613.**

51. **Woutersen, S., and Hamm, P. (2000) Structure determination of trialanine in water using polarized sensitive two-dimensional vibrational spectroscopy, *J. Phys. Chem. B 104*, 11316-11320.**

52. **Woutersen, S., and Hamm, P. (2001) Isotope-edited two-dimensional vibrational spectroscopy of trialanine in aqueous solution, *J. Chem. Phys. 114*, 2727-2737.**

53. **Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins, *Proc. Natl. Acad. Sci. U. S. A. 101*, 12491-12496.**

54. **Millet, I. S., Doniach, S., and Plaxco, K. W. (2002) Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins, in *Unfolded Proteins*, pp 241-262, Academic Press Inc, San Diego.**

55. **Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques, *Biochemistry 38*, 16424-16431.**

56. **Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc. Natl. Acad. Sci. USA 101*, 12497.**

57. **Jha, A. K., Kolubri, A., Freed, K. F., and Sosnick, T. R. (2005) Statistical coil model of the unfolded state: Resolving the Reconciliation Problem, *Proc. Natl. Acad. Sci. USA 102*, 13099-13104.**

58. **Dill, K. A., and Shortle, D. (1991) Denatured State of Proteins, *Annu. Rev. Biochem. 60*, 795-825.**

59. **Kihara, D. (2005) The effect of long range interactions on the secondary structure formation of proteins, *Protein Sci. 14*, 1955-1963.**

60. **Minor, D. L., jr, and Kim, P. S. (1996) Context dependence secondary structure formation of a designed protein sequence, *Nature 380*, 730-734.**

61. **Jha, A. K., Colubri, A., Zaman, M. H., Koide, S., Sosnick, T. R., and Freed, K. F. (2005) Helix, Sheet and Polyproline II Frequencies and Strong Nearest Neighbor Effects in a Restricted Coil Library, *Biochemistry 44*, 9691-9702.**

62. **Avbelj, F. (private communication).**

63. **Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., and Dobson, C. M. (1996) Toward a description of the Conformations of Denatured States of proteins. Comparison of a Random Coil Model with NMR measurements, *J. Phys. Chem. 100*, 2661-2666.**

64. **Avbelj, F., and Baldwin, R. L. (2003) Role of backbone solvation and electrostatics in generating preferred peptide backbone conformations: Distributions of phi, *Proc. Natl. Acad. Sci. USA 100*, 5742-5747.**

65. **Pappu, R. V., Srinivasan, R., and Rose, G. D. (2000) The Flory isolated-pair hypothesis is not valid for polypeptide chains: Implications for protein folding, *Proc. Natl. Acad. Sci. USA 97*, 12565-12570.**

66. **Zaman, M. H., Shen, M.-Y., Berry, R. S., Freed, K. F., and Sosnick, T. R. (2003) Investigations into Sequence and Conformational Dependence of Backbone Entropy, Inter-basin Dynamics and the Flory Isolated-pair Hypothesis for Peptides, *J. Mol. Biol. 331*, 693-711.**

67. **Avbelj, F., and Baldwin, R. L. (2004) Origin of the neighboring residue effect on peptide backbone conformation, *Proc. Natl. Acad. Sci. USA 101*, 19067-10972.**

68. **Makowska, J., Rodziewicz-Motowidlo, S., Baginska, K., Vila, J. A., Liwo, A., Chmurzynski, L., and Scheraga, H. A. (2006) Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins, *Proc. Natl. Acad. Sci. USA 103*, 1744-1749.**

69. **Sosnick, T. R. Sampling library, pp http://godzilla.uchicago.edu/cgi-bin/rama.cgi.**

70. **Ho, B. K., and Dill, K. A. (2006) Folding Very Short Peptides Using Molecular Dynamics, *PLOS Computational Biology 2*, 0228-0237.**

71. **Eker, F., Griebenow, K., Cao, X., Nafie, L., and Schweitzer-Stenner, R. (2004) Preferred peptide backbone conformations in the unfolded state revealed by the structure analysis of alanine-based (AXA) tripeptides in aqueous solution, *Proc. Natl. Acad. Sci. USA 101*, 10054-10059.**

72. **Shi, Z., Chen, K., Liu, Z., Ng, A., Bracken, W. C., and Kallenbach, N. R. (2005) Polyproline II propensities from GGXGG peptides reveal an anticorrelation with β-sheet scales, *Proc. Natl. Acad. Sci. USA 102*, 17964 - 17968.**

73. **Tran, H. T., Wang, X., and Pappu, R. V. (2005) Reconciling Observations of Sequence-Specific Conformational Properties with Generic Polymeric Behavior of Denatured Proteins, *Biochemistry 44*, 11369-11380.**

74. **Woutersen, S., Pfister, R., Hamm, P., Mu, Y., Kosov, D. S., and Stock, G. (2002) Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations, *J. Chem. Phys. 117*, 6833-6840.**

75. **Ding, L., Chen, K., Santini, P. A., Shi, Z., and Kallenbach, N. R. (2003) The Pentapeptide GGAGG Has PII Conformation, *J. Am. Chem. Soc. 125*, 8092-8093.**

76. **Schweitzer-Stenner, R., Eker, F., Griebenow, K., Cao, X., and Nafie, L. (2004) The conformation of tetraalanine in water determined by polarized Raman, FTIR and VCD spectroscopy, *J. Am. Chem. Soc. 126*, 2768-2776.**

77. **Eker, F., Cao, X., Nafie, L., and Schweitzer-Stenner, R. (2002) Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR and VCD spectroscopy study, *J. Am. Chem. Soc. 124*, 14330-14341.**

78. **Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2003) Stable conformations of tripeptides in aqueous solution studied by UV Circular Dichroism Spectroscopy, *J. Am. Chem. Soc. 125*, 8178-8185.**

79. **Avbelj, F., Gradolnik, S. G., Grdadolnik, J., and Baldwin, R. L. (2006) Intrinsic backbone preferences are fully present in blocked amino acids, *Proc. Natl. Acad. Sci. USA 103*, 1272-1277.**

80. **McColl, I. H., Blanch, E. W., Hecht, L., Kallenbach, N. R., and Barron, L. D. (2004) Vibrational Raman optical activity characterization of poly(L-proline II) helix in alanine oligopeptides, *J. Am. Chem. Soc. 126*, 5076-5077.**

81. **Asher, S. A., Mikhonin, A. V., and Bykov, S. (2004) UV Raman Demonstrates that α-Helical Polyalanine Peptides Melt to Polyproline II Conformations, *J. Am. Chem. Soc. 126*, 8433-8440.**

82. **Makowska, J., Rodziewicz, S., Baginska, K., Makowski, M., Vila, J. A., Liwo, A., Chmurzyński, L., and Scheraga, H. A. (2007) Further Evidence for the Absence of Polyrpoline II Stretch in the XAO Peptide, *Biophys. J. 92*, 2904-2917.**

83. **Zagrovic, B., Lipfert, J., Sorin, E. J., Millett, I. S., van Gunsteren, W. F., Doniach, S., and Pande, V. S. (2005) Unusual compactness of a polyproline II structure, *Proc. Natl. Acad. Sci. USA 102*, 11698-11703.**

84. **Schweitzer-Stenner, R., and Measey, T. (2007) The Alanine-Rich XAO Peptide Adopts a Heterogeneous Population, Including Turn-Like and PPII Conformations, *Proc. Natl. Acad. Sci. USA 104*, 6649–6654.**

85. **Graf, J., Nguyen, P. H., Stock, G., and Schwalbe, H. (2007) Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study, *J. Am. Chem. Soc. 129*, 1179-1189.**

86. **Schweitzer-Stenner, R. (2009) Distribution of Conformations Sampled by the Central Amino Acid Residue in Tripeptides Inferred From Amide I Band Profiles and NMR Scalar Coupling Constants, *J. Phys. Chem. B 113*, 2922-2932.**

87. **Mu, Y., and Stock, G. (2002) Conformational dynamics of trialanine in water: A water dynamics study, *J. Phys. Chem. B 106*, 5294-5301.**

88. **Mu, Y., Kosov, D. S., and Stock, G. (2003) Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments, *J. Phys. Chem. B 107*, 5064.**

89. **Kentsis, A., Mezei, M., Gindin, T., and Osman, R. (2004) Unfolded State of Polyalanine Is a Segmented Polyproline II Helix, *Proteins: Structure, Function and Genetics 55*, 493-501.**

90. **Mezei, M., Fleming, P. J., Srinivasan, R., and Rose, G. D. (2004) Polyproline II Helix Is the Preferred Conformation for Unfolded Polyalanine in Water, *Proteins: Structure, Function and Genetics 55*, 502-507.**

91. **Beck, D. A. C., Alonso, D. O., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occuring amino acids and reflection of these propensities in proteins, *Proc. Natl. Acad. Sci. U. S. A. 105*, 12259-12264.**

92. **Best, R. B., and Hummer, G. (2009) Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides, *J. Phys. Chem. B 113*, 9004-9015.**

93. **Best, R. B., Buchete, N.-V., and Hummer, G. (2008) Are Current Molecular Dynamics Force Fields too Helical?, *Biophys. J. 94*, L07-L09.**

94. **Garcia, A. E. (2004) Characterization of non-alpha conformations in Ala peptides, *Polymer 120*, 885-890.**

95. **Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A Point-Charge Force Field for Molecular Mechanics Simulations of Protein Based on Condensed-Phase Quantum Mechanical Calculations, *J. Comp. Chem. 24*, 1999-2012.**

96. Eker, F., Cao, X., Nafie, L., Griebenow, K., and Schweitzer-Stenner, R. (2003) The Structure of Alanine Based Tripeptides in Water and Dimethyl Sulfoxide Probed by Vibrational Spectroscopy, *J. Phys. Chem. B 107*, 358-365.

97. Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A., and Creamer, T. P. (2001) Host-guest study of left handed polyproline II helix formation, *Biochemistry 2001*, 14376-14383.

98. Pappu, R. V., and Rose, G. D. (2002) A simple model for polyproline II structure in unfolded states of alanine-based peptides, *Protein Sci. 11*, 2437-2455.

99. Shi, Z., Woody, R. W., and Kallenbach, N. R. (2002) Is polyproline II a major backbone conformation in unfolded proteins?, *Adv. Protein Chem. 62*, 163-240.

100. Stapley, B. J., and Creamer, T. P. (1999) Lysine Peptides Revisited, *Protein Sci. 8*, 587-595.

101. Chellgren, B. W., and Creamer, T. P. (2004) Effects of H2O and D2O polyproline lane II helical structure, *J. Am. Chem. Soc. 126*, 14734-14735.

102. Liu, Z. G., Chen, K., Ng, A., Shi, Z. S., Woody, R. W., and Kallenbach, N. R. (2004) Solvent dependence of PII conformation in model alanine peptides, *J. Am. Chem. Soc. 126*, 15141-15150.

103. Drozdov, A. N., Grossfield, A., and Pappu, R. V. (2004) Role of Solvent in determining Conformational Preferences of Alanine Dipeptide in Water, *J. Am. Chem. Soc. 126*, 2574-2581.

104. Law, P. B., and Daggett, V. (2010) The relationship between water bridges and the polyproline II conformation: a large-scale analysis of molecular dynamics simulations and crystal structures, *Protein Engineering, Design & Selection 23*, 27-33.

105. Fleming, P. J., Fitzkee, N. C., Mezei, M., Srinivasan, R., and Rose, G. D. (2005) A Novel Method Reveals that Solvent Water Favors Polyproline II over β-strand Conformation in Peptides and Unfolded Proteins: Conditional Hydrophobic Accessible Surface Area (CHASA), *Protein Sci. 14*, 111-118.

106. Hagarman, A., Measey, T., Doddasamayajula, R. S., Dragomir, I., Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2006) Conformational Analysis of XA and AX Dipeptides in Water by Electronic Circular Spectroscopy, *J. Phys. Chem. B 110*, 6979-6986.

107. Gnanakaran, S., and Garcia, A. E. (2003) Validation of an All-Atom Protein Force Field: From Dipeptides to larger Peptides., *J. Phys. Chem. B 107*, 12555-12557.

108. Garcia, A. E. (2004) Characterization of non-alpha conformations in Ala peptides. , *Polymer 120*, 885-890.

109. **Wirmer, J., and Schwalbe, H. (2002) Angular dependence of (1)J(N-i,C-alpha i) and (2)J(N-i,C alpha(i-1)) coupling constants measured in J-modulated HSQCs, *J. Biomol. NMR 23*, 47-55.**

110. **Löhr, F., and Rüterjans, H. (1995) (H)NCAHA and (H)CANNH Experiments for the Dtetermination of Vicinal Coupling Constants Related to the Phi Torsional Angle, *J. Biomol. NMR 5*, 25-36.**

111. **Löhr, F., and Rüterjans, H. (1999) Alternative E.COSY techniques for the measurement of (3)J(C '(i-1),C-i(beta)) and (3)J(H-i(N),C-i(beta)) coupling constants in proteins, *J. Biomol. NMR 13*, 263-274.**

112. **Hennig, M., Bermel, W., Schwalbe, H., and Griesinger, C. (2000) Determination of psi torsion angle restraints from (3)J(C-alpha,C-alpha) and 3J(C-alpha,H-N) coupling constants in proteins, *J. Am. Chem. Soc. 122*, 6268-6277.**

113. **Weisemann, R., Ruterjans, H., Schwalbe, H., Schleucher, J., Bermel, W., and Griesinger, C. (1994) Determination of H(N),H-ALPHA and H(N), C' Coupling Constants in C-13,N-15-Labeled Proteins, *J. Biomol. NMR 4*, 231-240.**

114. **Surewicz, W. K., and Mantsch, H. H. (1988) New insight into protein secondary structure from resolution-enhanced infrared spectra, *Biochim. Biophys. Acta 952*, 115-130.**

115. **Bandekar, J. (1992) Amide modes and protein conformation, *Biochim. Biophys. Acta 1120*, 123-143.**

116. **Chen, X. G., Schweitzer-Stenner, R., Asher, S. A., Mirkin, N. G., and Krimm, S. (1995) Vibrational Assignments of trans-N-Methylacetamide and Some of Its Deuterated Isotopomers from Band Decomposition of IR, Visible, and Resonance Raman Spectra, *J. Phys. Chem. 99*, 3074-3083.**

117. **Mirkin, N. G., and Krimm, S. (1991) *Ab initio* vibrational analysis of hydrogen-bonded trans- and cis-N-methylacetamide, *J. Am. Chem. Soc. 113*, 9742-9747.**

118. **Schweitzer-Stenner, R. (2001) Visible and UV-resonance Raman spectroscopy on model peptides, *J. Raman Spectrosc. 32*, 711.**

119. **Sieler, G., and Schweitzer-Stenner, R. (1997) The Amide I Mode of Peptides in Aqueous Solution Involves Vibrational Coupling Between The Peptide Group and Water Molecules of The Hydration Shell, *J. Am. Chem. Soc. 119*, 1720-1726.**

120. **Ham, S., Kim, J. H., Lee, H., and Cho, M. H. (2003) Correlation between electronic and molecular structure distortions and vibrational properties. II. Amide I modes of NMA-nD(2)O complexes, *J. Chem. Phys. 118*, 3491-3498.**

121. **Krimm, S., and Bandekar, J. (1986) Vibrational Spectroscopy of Peptides and Proteins, *Adv. Protein Chem. 38*, 181.**

122. **Torii, H., and Tasumi, M. (1998) Ab Initio Molecular Orbital Study of the Amide I Vibrational Interactions Between the Peptide Groups in Di- and Tripeptides and Considerations on the Conformation of the Extended Helix, *J. Raman Spectrosc. 29*, 81-86.**

123. **Choi, J.-H., Ham, S., and Cho, M. (2003) Local amide I mode frequencies and coupling constants in polypeptides, *J. Phys. Chem. B 107*, 9132-9138.**

124. **Measey, T., Hagarman, A., Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2005) Side Chain Dependence of Intensity and Wavenumber Position of Amide I' in IR and Visible Raman Spectra of XA and AX Dipeptides, *J. Phys. Chem. B 109*, 8195-8205.**

125. **Schweitzer-Stenner, R. (2004) Secondary structure analysis of polypeptides based on an excitonic coupling model to describe the band profile of amide I of IR, Raman and vibrational circular dichroism spectra, *J. Phys. Chem. B 108*, 16965-16975.**

126. **Schweitzer-Stenner, R. (2006) Advances in vibrational spectroscopy as a sensitive probe of peptide and protein structure. A critical review, *Vibr. Spectrosc. 42*, 98-117.**

127. **Schweitzer-Stenner, R., Measey, T., Hagarman, A., and Dragomir, I. (2010) The Structure of unfolded peptides and proteins explored by Raman and IR spectroscopies, in *Vibrational spectroscopy on peptides and proteins.* (Longhi, S., & Uversky, V.N., Ed.), p in press, Wiley & sons, Chichester.**

128. **Choi, J.-H., Ham, S., and Cho, M. (2002) Inter-peptide interaction and delocalization of amdie I vibrational excitons in myoglobin and flavodoxin, *J. Chem. Phys. 117*, 6821-6832.**

129. **Hamm, P., Lim, M., DeGrado, W. F., and Hochstrasser, R. (1999) Stimulated Photon Echos from Amide I Vibrations, *J. Phys. Chem. B 103*, 10049-10053.**

130. **Moran, A., and Mukamel, S. (2004) The origin of vibrational mode couplings in various secondary structural motifs of polypeptides, *Proc. Natl. Acad. Sci. USA 101*, 506-510.**

131. **Schweitzer-Stenner, R., Measey, T., Hagarman, A., Eker, F., and Griebenow, K. (2006) Salmon Calcitonin an Amyloid β: Two Peptides with Amyloidogenic Capacity Adopt Different Conformational Manifolds in Their Unfolded States, *Biochemistry 45*, 2810-2819.**

132. **Glasoe, P. K., and Long, F. A. (1960) Use of glass electrodes to measure acidities in Deuterium Oxide, *J. Phys. Chem. 64*, 188-190.**

133. **Schweitzer-Stenner, R., Measey, T., Kakalis, L., Jordan, F., Pizzanelli, S., Forte, C., and Griebenow, K. (2007) Conformations of Alanine-Based Peptides in Water Probed by FTIR, Raman, Vibrational Circular Dichroism, Electronic Circular Dichroism, and NMR Spectroscopy, *Biochemistry 46*, 1587-1596.**

134. **Shaka, A. J., Lee, C. J., and Pines, A. (1988) Iterative Schemes for Bilinear Operators - Application to Spin Decoupling, *J. Magn. Reson. 77*, 274-293.**

135. **Jentzen, W., Unger, E., Karvounis, G., Shelnutt, J. A., Dreybrodt, W., and Schweitzer-Stenner, R. (1995) Conformational Properties of Nickel(II) Octaethylporphyrin in Solution. 1. Resonance Excitation Profiles and Temperature Dependence of Structure-Sensitive Raman Lines, *J. Phys. Chem. 100*, 14184-14191.**

136. **Schweitzer-Stenner, R. (2008) Conformational Analysis of Unfolded Peptides by Vibrational Spectroscopy, in *Unfolded Proteins. From Denatured States to Intrinsically Disordered.* (Creamer, T. A., Ed.), pp 101-142, Novalis Press, New York.**

137. **Motta, A., Reches, M., Pappalardo, L., Andreotti, G., and Gazit, E. (2005) The Preferred Conformation of the Tripeptide Ala-Phe-Ala in Water Is an Inverse γ-turn: Implications for Protein Folding and Drug Design, *Biochemistry 144*, 14170-14178.**

138. **Gong, H., and Rose, G. D. (2008) Assessing the solvent-dependent surface area of unfolded proteins using an ensemble model, *Proc. Natl. Acad. Sci. USA 105*, 3321-3326.**

139. **Hutchinson, E. G., and Thornton, J. M. (1994) A Revised Set of Potentials for β-turn Formation in Proteins, *Protein Sci. 3*, 2207-2216.**

140. **Measey, T., and Schweitzer-Stenner, R. (2005) Simulation of amide I' band profiles of trans polyproline based on an excitonic coupling model, *Chem. Phys. Lett. 408*, 123-127.**

141. **Chen, K., Liu, Z., and Kallenbach, N. R. (2004) The polyproline II conformation in short alanine peptides is non cooperative, *Proc. Natl. Acad. Sci. USA 101*, 15352-15357.**

142. **Xu, C., Wang, J., and Liu, H. (2008) Hamiltonian replica exchange approach and its application to the study of side-chain type and neighbor effects on peptide backbone conformations, *J. Chem. Theory Comput. 4*, 1348-1359.**

143. **Tiffany, M. L., and Krimm, S. (1968) New chain conformations of poly(glutamicacid) and polylysine., *Biopolymers 6*, 1767-1770.**

144. **Bemporad, F., Taddei, N., Stefani, M., and Chiti, F. (2006) Assessing the role of aromatic residues in the amyloid aggregation of human muscle acylphosphatase, *Protein Sci. 15*, 862-870.**

145. **Tjernberg, L., Hosia, W., Bark, N., Thyberg, J., and Johansson, J. (2002) Charge Attraction and β Propensity Are Necessary for Amyloid Fibril Formation from Tetrapeptides, *J. Biol. Chem. 277*, 43243-43246.**

146. **Tjernberg, L. O., Naslund, J., Lindqvist, F., Johansson, J., Karlstrom, A. R., Thyberg, J., Terenius, L., and Nordstedt, C. (1996) Arrest of beta-amyloid fibril formation by a pentapeptide ligand, *J. Biol. Chem. 271*, 8545-8548.**

147. **Kim, W., and Hecht, M. H. (2006) Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide, *Proc. Natl. Acad. Sci. USA 103*, 15825-15830.**

148. **Xiong, H., Buckwalter, B. L., Shieh, H.-M., and Hecht, M. H. (1995) Periodicity of polar and nonpolar amino acids is the major determinant of secondary structure in self-assembling oligomeric peptides, *Proc. Natl. Acad. Sci. USA 92*, 6349-6353.**

149. **Richardson, J. S. (1981) Protein Anatomy, *Adv. Prot. Chem. 34*, 167-339.**

150. **Duddy, W. J., Nissink, J. W. M., Allen, F. H., and Milner-White, E. J. (2004) Mimicry by Asx- and ST-turns of the Four Main Types of β–turn in Protein, *Protein Sci. 13*, 3051-3055.**

151. **Wan, W. Y., and Milner-White, E. J. (1999) A recurring two-hydrogen-bond motif incorporating a serine or threonine residue is found both at α–helical termini and in other situations, *J. Mol. Biol. 286*, 1650-1666.**

152. **Avbelj, F., and Baldwin, R. L. (2003) Role of backbone solvation and electrostatics in generating preferred peptide backbone conformations: Distributions of phi. , *Proc. Natl. Acad. Sci. USA 100*, 5742-5747.**

153. **Avbelj, F., and Baldwin, R. L. (2004) Origin of the neighboring residue effect on peptide backbone conformation. , *Proc. Natl. Acad. Sci. USA 101*, 19067-10972.**

154. **Jha, A. K., Kolubri, A., Freed, K. F., and Sosnick, T. R. (2005) Statistical coil model of the unfolded state: Resolving the Reconciliation Problem. , *Proc. Natl. Acad. Sci. USA 102*, 13099-13104.**

155. **Bernado, P., Blanchard, L., Timminis, P., Marion, D., Rugrok, R. W. H., and Blackledge, M. (2005) A structural model for unfolded proteins from dipolar couplings and small-angle x-ray scattering, *Proc. Natl. Acad. Sci. USA 102*, 17002-17007.**

156. **Bernado., P., Bertoncini, C. W., Griesinger, C., Zweckstetter, M., and Blackledge, M. (2005) Defining Long-Range Order and Local Disorder in Native α–Synuclein Using Residual Dipolar Couplings, *J. Am. Chem. Soc. 127*, 17968-17969.**

157. **Vitalis, A., and Pappu, R. V. (2008) ABSINTH: A New Continuum Solvation Model for Simulation of Polypeptides in aqueous Solutions, *J. Comp. Chem. 30*, 673-699.**

158. **Shi, Z., Chen, K., Liu, Z., Ng, A., Bracken, W. C., and Kallenbach, N. R. (2005) Polyproline II propensities from GGXGG peptides reveal an anticorrelation with β-sheet scales. , *Proc. Natl. Acad. Sci. USA 102*, 17964 - 17968.**

159. **Grdadolnik, J., Grdadolnik, S. G., and Avbelj, F. (2008) Determination of Conformational Preferences of Dipeptides Using Vibrational Spectroscopy, *J. Phys. Chem. B 112*, 2712-2718.**

160. **Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A., and Creamer, T. P. (2001) Host-guest study of left handed polyproline II helix formation. , *Biochemistry 2001*, 14376-14383.**

161. **Plaxco, K. W., Morton, C. J., Grimshaw, S. B., Jones, J. A., Pitkeathly, M., Campbell, I. D., and Dobson, C. M. (1997) The effects of guanidine hydrochloride on the 'random coil' conformations and NMR chemical shifts of the peptide series GGXGG, *J. Biomol. NMR 10*, 221-230.**

162. **Eker, F., Cao, X., Nafie, L., Griebenow, K., and Schweitzer-Stenner, R. (2003) The Structure of Alanine Based Tripeptides in Water and Dimethyl Sulfoxide Probed by Vibrational Spectroscopy., *J. Phys. Chem. B 107*, 358-365.**

163. **Eker, F., Cao, X., Nafie, L., and Schweitzer-Stenner, R. (2002) Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR and VCD spectroscopy study. , *J. Am. Chem. Soc. 124*, 14330-14341.**

164. **Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I. M., Stewart, J. M., and Baldwin, R. L. (1985) Nature of the charged-group effect on the stability of the C-peptide helix, *Proc. Natl. Acad. Sci. USA 82*, 2349-2353.**

165. **Zagrovic, B., Lipfert, J., Sorin, E. J., Millett, I. S., van Gunsteren, W. F., Doniach, S., and Pande, V. S. (2005) Unusual compactness of a polyproline II structure. , *Proc. Natl. Acad. Sci. USA 102*, 11698-11703.**

166. **Gnanakaran, S., and Garcia, A. E. (2005) Helix-coil transition of alanine peptides in water: Force field dependence on the folded and unfolded structures, *Proteins 59*, 773-782.**

167. **Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A 2nd generation force field for the simulation of proteins; nucleic acids; and organic molecules. , *J. Am. Chem. Soc. USA 117*, 5179-5197.**

168. **Garcia, A. E., and Sanbonmatsu, K. Y. (2002) α-Helical stabilization by side chain shielding of backbone hydrogen bonds., *Proc. Natl. Acad. Sci. U. S. A. 99*, 2782-2787.**

169. **Serrano, L. (1995) Comparison between the ϕ-distribution of the amino acids in the Protein Data Base and NMR data indicates that amino acids have various ϕ propensities in the random coil conformation., *J. Mol. Biol. 254*, 322-333.**

170. **Edman, E. T. (1979) A comparison of the structures of electron transfer proteins, *Biochim. Biophys. Acta 549*, 107-144.**

171. **Belikova, N. A., Vladimirov, Y. A., Osipov, A. N., Kapralov, A. A., Tyurin, V. A., Potapovich, M. V., Basova, L. V., Peterson, J., Kurnikov, I. V., and Kagan, V. E. (2006) Peroxidase activity and structural transitions of cytochrome c bound to cardiolipin-containing membranes, *Biochemistry 45*, 4998-5009.**

172. **Garber, E. A. E., and Margoliash, E. (1994) Circular dichroism studies of the binding of mammalian and non-mammalian cytochromes c to cytochrome c oxidase, cytochrome c peroxidase, and polyanions, *Biochim. Biophys. Acta 1187*, 289-295.**

173. **Green, D. R., and Kroemer, G. (2004) The pathology of mitochondrial cell death, *Science 305*, 626-629.**

174. **Jiang, X., and Wang, X. (2004) Cytochrome c-mediated apoptosis, *Ann. Rev. Biochem. 73*, 87-106.**

175. **Liu, X., Kim, C. N., Yang, J., Jemmerson, R., and Wang, X. (1996) Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c, *Cell 86*, 147-157.**

176. **Purring-Koch, C., and McLendon, G. (2000) Cytochrome c binding to Apaf-1: the effects of dAtp and ionic strength, *Proc. Natl. Acad. Sci. USA 97*, 11928-11931.**

177. **Jemmerson, R., Liu, J., Hausauer, D., Lam, K.-P., Mondino, A., and Nelson, R. D. (1999) A conformational change in cytochrome c of apoptotic and necrotic cells is detected by monoclonal antibody binding and mimicked by association of the native antigen with synthetic phospholipid vesicles, *Biochemistry 38*, 3599-3609.**

178. **Döpner, S., Hildebrandt, P., Rosell, F. I., Mauk, A. G., von Walter, M., Soulimane, T., and Buse, G. (1999) The structural and functional role of lysine residues in the binding domain of cytochrome c for the redox process with cytochrome c oxidase, *Eur. J. Biochem. 261*, 379-391.**

179. **Weber, C., Michel, B., and Bosshard, H. R. (1987) Spectroscopic Analysis of the Cytochrome c Oxidase-Cytochrome c Complexes: Circular Dichroism and Magnetic Circular Dichroism Measurements Reveal Changes of Cytochrome c Heme Geometry Imposed by Complex Formation, *Proc. Natl. Acad. Sci. USA 84*, 6687-6691.**

180. **Louie, G. V., and Brayer, G. D. (1990) High-resolution Refinement of Yeast Iso-1-Cytochrome c and Comparisons With Other Eukaryotic Cytochromes c, *J. Mol. Biol. 214*, 527-555.**

181. **Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) High-resolution Three-dimensional Structure of Horse Heart Cytochrome c, *J. Mol. Biol. 214*, 585-595.**

182. **Moore, G. W., and Pettigrew, G. W. (1990) *Cytochrome c –evolutionary, structural and physicochemical aspects*, Springer, Berlin, Heidelberg, New York.**

183. **Berghuis, A. M., and Brayer, G. D. (1992) Oxidation State-Dependent Conformational Changes in Cytochrome c, *J. Mol. Biol. 223*, 959-976.**

184. **Filosa, A., Yang, Y., Ismail, A. A., and English, A. M. (2001) Two-dimensional infrared correlation spectroscopy as a probe of sequential events in the thermal unfolding of cytochromes c, *Biochemistry 40*, 8256-8263.**

185. **Theorell, H., and Åkesson, Å. (1941) Studies on Cytochrome c, *J. Am. Chem. Soc. 63*, 1804-1820.**

186. **Taler, G., Schejter, A., Navon, G., Vig, I., and Margoliash, E. (1995) The nature of the thermal equilibrium affecting the iron coordination of ferric cytochrome c, *Biochemistry 34*, 14209-14212.**

187. **Dyson, H. J., and Beattie, J. K. (1982) Spin state and unfolding equilibria of ferri-cytochrome c in acidic solution, *J. Biol. Chem. 257*, 2267-2273.**

188. **Indiani, C., de Sanctis, G., Neri, F., Santos, H., Smulevich, G., and Coletta, M. (2000) Effect of pH on Axial Ligand Coordination of Cytochrome c" from *Methylophilus methylotrophus* and Horse Heart Cytochrome c., *Biochemistry 39*, 8234-8242.**

189. **Cohen, D. S., and Pielak, G. J. (1994) Stability of yeast iso-1-ferricytochrome c as a function of pH and temperature, *Protein Sci. 3*, 1253-1280.**

190. **Barker, P. D., and Mauk, A. G. (1992) pH-Linked conformational regulation of a metalloprotein oxidation-reduction equilibrium: electrochemical analysis of the alkaline form of cytochrome c, *J. Am. Chem. Soc. 114*, 3619-3624.**

191. **Battistuzzi, G., Borsari, M., Loschi, L., Martinelli, A., and Sola, M. (1999) Thermodynamics of the alkaline transition of cytochrome c, *Biochemistry 38*, 7900-7907.**

192. **Döpner, S., Hildebrandt, P., Rosell, F. I., and Mauk, A. G. (1998) The alkaline conformational transitions of ferricytochrome c studied by resonance Raman spectroscopy, *J. Am. Chem. Soc. 120*, 11246-11255.**

193. **Rossel, F. I., Ferrer, J. C., and Mauk, A. G. (1998) Proton-linked protein conformational switching: definition of the alkaline conformational transition of yeast iso-l-ferricytochrome c, *J. Am. Chem. Soc. 120*, 11234-11245.**

194. **Blouin, C., Guillemette, J. G., and Wallace, C. J. A. (2001) Resolving the individual components of a pH-induced conformational change, *Biophys. J. 81*, 2331-2338.**

195. **Assfalg, M., Bertini, I., Dolfi, A., Turano, P., Mauk, A. G., Rossel, F. I., and Gray, H. B. (2003) Structural Model for an alkaline Form of Ferricytochrome c, *J. Am. Chem. Soc. 125*, 2913-2922.**

196. **Schejter, A., and George, P. (1964) The 695 mμ band of ferricytochrome c and its relationship to protein conformation, *Biochemistry 3*, 1045-1049.**

197. **Filosa, A., and English, A. M. (2000) Probing local thermal stabilities of bovine, horse, and tuna ferricytochromes c at pH 7, *JBIC 4*, 448-454.**

198. **Theorell, H., and Åkesson, Å. (1939) Absorption Spectrum of Further Purified Cytochrome c, *Science 90*, 67.**

199. **Pinheiro, T. J. T., Elöve, G., Watts, A., and Roder, H. (1997) Structural and Kinetic Description of Cytochrome c Unfolding Induced by the Interaction with Lipid Vesicles, *Biochemistry 36*, 13122-13132.**

200. **Eaton, W. A., and Hochstrasser, R. M. (1968) Singe Crystal Spectra of Ferrimyoglobin in Polarized Light, *J. Chem. Phys. 49*, 985-995.**

201. **McKnight, J., Cheesman, M. R., Thomson, A. J., Miles, J. S., and Munro, A. W. (1993) Identificationof charge transfer transitions in the optical spectrum of low-spin ferric cytochrome -P450 *Bacillus Megaterium.*, *Eur. J. Biochem. 213*, 683-687.**

202. **Battistuzzi, G., Loschi, L., Borsari, M., and Sola, M. (1999) Effects of nonspecific ion-protein interactions on the redox chemistry of cytochrome c, *JBIC 4*, 601-607.**

203. **Battistuzzi, G., Borsari, M., Dallari, D., Lancelloti, I., and Sola, M. (1996) Anion Binding to Mitochondrial Cytochromes c Studied Through Electrochemistry - Effects of the Neutralization of Surface Charges on the Redox Potential, *Eur. J. Biochem. 241*, 208-214.**

204. **Feng, Y., and Englander, S. W. (1990) Salt dependent Structural Change and Ion Binding in Cytochrome c Studied by Two-Dimensional Proton NMR, *Biochemistry 29*, 3505-3509.**

205. **Moench, S. J., Shi, T.-M., and Satterlee, J. S. (1991) Proton-NMR studies of the effects of ionic strength and pH on the hyperfine-shifted resonances and phenylalanine-82 environment of three species of mitochondrial ferricytochrome c, *Eur. J. Biochem. 197*, 631-641.**

206. **Trewhella, J., Carlson, V. A. P., Curtis, E. H., and Heidorn, D. B. (1988) Differences in the solution structures of oxidized and reduced cytochrome c measured by small-angle X-ray scattering, *Biochemistry 27*, 1121-1125.**

207. **Liu, G., Grygon, C. A., and Spiro, T. G. (1989) Ionic Strength Dependence of CYtochrome c Structure and Trp-59 H/D Exchange from Ultravoilet Resonance Raman Spectroscopy, *Biochemistry 28*, 5046-5050.**

208. **Barkigia, K. M., Chantranupong, L., Smith, K. M., and Fajer, J. (1988) Structural and Theoretical-Models of Photosynthetic Chromophores - Implications for Redox, Light-Absorption Properties and Vectoral Electron Flow, *J. Am. Chem. Soc. 110*, 7566-7567.**

209. **Schweitzer-Stenner, R., Levantino, M., Cupane, A., Wallace, C. J. A., Laberge, M., and Huang, Q. (2006) Functional Relevant Electric-Field Induced Perturbations of the Prosthetic Group of Yeast Ferrocytochrome c Mutants Obtained from a Vibronic Analysis of Low Temperature Absorption Spectra, *J. Phys. Chem. B 110*, 12155-12161.**

210. **Schweitzer-Stenner, R., Hagarman, A., Verbaro, D., and Soffer, J. (2009) Conformational Stability of Cytochrome c Probed By Optical Spectroscopy, *Meth. Enzymol. 466*, 109-153.**

211. **Schweitzer-Stenner, R., Gorden, J. P., and Hagarman, A. (2007) The asymmetric band profile of the soret band of deoxymyoglobin is caused by electronic and vibronic perturbations of the heme group rather than by a doming deformation, *J. Chem. Phys. 127*, 135103.**

212. **Schweitzer-Stenner, R. (2008) The Internal Electric Field in Cytochrome C Explored by Visible Electronic Circular Dichroism Spectroscopy, *J. Phys. Chem. B 112*, 10358-10366.**

213. **Woody, R. W. (1996) Theory of CIrcular Dichroism of Proteins, in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.), pp 25-67, Plenum Press, New York.**

214. **Cupane, A., Leone, M., Vitrano, E., and Cordone, L. (1995) Low Temperature optical absorption spectroscopy: an approach to the study of stereodynamic properties of hemeproteins, *Eur. Biophys. J. 23*, 385-398.**

215. **Stallard, B. R., Champion, P. M., Callis, P. R., and Albrecht, A. C. (1983) Advances in calculating Raman excitation profiles by means of the transform theory, *J. Chem. Phys. 78*, 712-722.**

216. **Schweitzer-Stenner, R. (2001) Polarized Resonance Raman Dispersion Spectroscopy on Metalporphyrins. , *J. Porphyrins Phthalocyanines 5*, 198.**

217. **Simpson, W. T. (1949) On the theory of the π-electron system in porphines, *J. Chem. Phys. 18*, 1218.**

218. **Longuet-Higgins, H. C., Rector, C. W., and Platt, J. R. (1950) Molecular orbital calculations on porphine and tetrahydroporphine, *J. Chem. Phys. 18*, 1174-1181.**

219. **Gouterman, M. (1959) Study of the effect of substitution on the absorption spectra of porphyrins, *J. Chem. Phys. 30*, 1139.**

220. **Schweitzer-Stenner, R. (1989) Allosteric linkage-induced distortions of the prosthetic group in haem proteins as derived by the theoretical interpretation of the depolarization ratio in resonance Raman scattering, *Q. Rev. Biophys. 22*, 381.**

221. **Levantino, M., Huang, Q., Cupane, A., Laberge, M., Hagarman, A., and Schweitzer-Stenner, R. (2005) The importance of vibronic perturbations in ferrocytochrome c spectra: A reevaluation of spectral properties based on low-temperature optical absorption, resonance Raman, and molecular-dynamics simulations, *J. Chem. Phys. 123*, 054508.**

222. **Dragomir, I., Hagarman, A., Wallace, C., and Schweitzer-Stenner, R. (2007) Optical band splitting and electronic perturbations of the heme chromophore in cytochrome c at room temperature probed by visible electronic circular dichroism spectroscopy, *Biophys. J. 92*, 989-998.**

223. **Manas, E. S., Vaderkooi, J. M., and Sharp, K. A. (1999) The effects of protein environment on the low temperature electronic spectroscopy of cytochrome c and microperoxidase-11, *J. Phys. Chem. B 103*, 6334-6348.**

224. **Schweitzer-Stenner, R., and Bigman, D. (2001) Electronic and vibronic contributions to the band splitting in optical spectra of heme proteins, *J. Phys. Chem. B 105*, 7064-7073.**

225. **Schweitzer-Stenner, R., Stichternath, A., Dreybrodt, W., Jentzen, W., Song, Z. P., Shelnutt, J. A., Faurskov Nielson, O., Medforth, C. J., and Smith, K. M. (1997) Raman dispersion spectroscopy on the highly saddled nickel(II)-octaethyltetraphenylporphyrin reveals the symmetry of nonplanar distortions and the vibronic strength of normal modes, *J. Chem. Phys. 107*, 1794-1814.**

226. **Lemke, C., Dreybrodt, W., Shelnutt, J. A., Quirke, J. M. E., and Schweitzer-Stenner, R. (1998) Polarized Raman dispersion spectroscopy probes planar and non-planar distortions of Ni(II) porphyrins with different peripheral substituents, *J. Raman Spectrosc. 29*, 945-953.**

227. **Jentzen, W., Song, X.-Z., and Shelnutt, J. A. (1997) Structural Characterization of Synthetic and Protein-Bound Porphyrins in Terms of the Lowest-Frequency Normal Coordinates of the Macrocycle, *J. Phys. Chem. B 101*, 1684.**

228. **Warshel, A. (1977) Energy-Structure Correlation in Metalloporphyrins and the Control of Oxygen Binding by Hemoglobin, *Proc. Natl. Acad. Sci. U. S. A. 74*, 1789.**

229. **Laberge, M., Köhler, M., Vanderkooi, J. M., and Freidrich, J. (1999) Sampling field heterogeneity at the heme of c-type cytochromes by spectral hole burning spectroscopy and electrostatic calculations, *Biophys. J. 77*, 3293-3304.**

230. **Song, Y. J., Haddad, R. E., Jia, S. L., Hok, S., Olmstead, M. M., Nurco, D. J., Schore, N. E., Zhang, J., Ma, J. G., Smith, K. M., Gazeau, S., Pecaut, J., Marchon, J. C., Medforth, C. J., and Shelnutt, J. A. (2005) Energetics and structural consequences of axial ligand coordination in nonplanar nickel porphyrins, *J. Am. Chem. Soc. 127*, 1179-1192.**

231. **Jentzen, W., Ma, M.-G., and Shelnutt, J. A. (1998) Conservation of the Conformation of the Porphyrin Macrocycle in Hemoproteins., *Biophys. J. 74*, 753-763.**

232. **Hobbs, J. D., and Shelnutt, J. A. (1995) Conserved Nonplanar Heme Distortions in Cytochromes c, *J. Protein Chem. 14*, 19-25.**

233. **Shelnutt, J. A., Song, X. Z., Ma, J. G., Jia, S. L., Jentzen, W., and Medforth, C. J. (1998) Nonplanar porphyrins and their significance in proteins, *Chem. Soc. Rev. 27*, 31-41.**

234. **Jentzen, W., Simpson, M. C., Hobbs, J. D., Song, X., Ema, T., Nelson, N. Y., Medforth, C. J., Smith, K. M., Veyrat, M., Mazzanti, M., Ramasseul, R., Marchon, J. C., Takeuchi, T., Goddard, W. A., and Shelnutt, J. A. (1995) Ruffling in a Series of NIckel (II) Meso-Tetrasubstituted Porphyrins as a Model for the Conserved Ruffling of the Heme of Cytochromes c, *J. Am. Chem. Soc. 117*, 11085-11097.**

235. **Shelnutt, J. A. (2000) *The Porphyrin Handbook*, Vol. 7, Academic Press, San Diego.**

236. **Laberge, M., Huang, Q., Schweitzer-Stenner, R., and Fidy, J. (2003) The endogenous calcium ions of horseradish peroxidase C are required to maintain the functional nonplanarity of the heme, *Biophys. J. 84*, 2542-2552.**

237. **Kohler, M., Friedrich, J., Laberge, M., and Vanderkooi, J. M. (1997) Influence of the pH on the pocket field of cytochrome c type proteins, *Chem. Phys. 224*, 327-335.**

238. **Dopner, S., Hildebrandt, P., Heibel, G. E., Vanhecke, F., and Mauk, A. G. (1995) The Effect of pH and Hydrogen-Deuterium Exchange on the Heme Pocket Structure of Cytochrome c Probed by Resonance Raman Spectroscopy, *J. Mol. Struct. 349*, 125-128.**

239. **Sinibaldi, F., Mei, G., Polticelli, F., Piro, M. C., Howes, B. D., Smulevich, G., Santucci, R., Ascoli, F., and Fiorucci, L. (2005) ATP specifically drives refolding of non-native conformations of cytochrome c, *Protein Sci. 14*, 1049-1058.**

240. **Li, D., Stuehr, D. J., Yeh, S.-R., and Rousseau, D. L. (2004) Heme Distortion modulated by Ligand-Protein Interactions in Inducible Nitirc-oxide Synthase., *J. Biol. Chem. 279*, 26489-26499.**

241. **Schweitzer-Stenner, R., Cupane, A., Leone, M., Lemke, C., Schott, J., and Dreybrodt, W. (2000) Anharmonic Protein Motions and Heme Deformations in Myoglobin Cyanide Probed by Absorption and Resonance Raman Spectroscopy, *J. Phys. Chem. B 104*, 4754.**

242. **Huang, Q., Szigeti, K., Fidy, J., and Schweitzer-Stenner, R. (2003) Structural disorder of native horseradish peroxidase C Probed by Resonance Raman and Low-Temperature Optical Absorption Spectroscopy, *J. Phys. Chem. B 107*, 2822-2830.**

243. **Huang, Q., Al-Azzam, W., Griebenow, K., and Schweitzer-Stenner, R. (2003) Heme structural perturbation of PEG-modified horseradish peroxidase c in aromatic organic solvents probed by optical absorption and resonance Raman dispersion spectroscopy, *Biophys. J. 84*, 3285-3298.**

244. **Huang, Q., Medforth, C. J., and Schweitzer-Stenner, R. (2005) Nonplanar Heme Deformations and Excited State Displacements in Nickel Porphyrins Detected by Raman Spectroscopy at Soret Excitation, *J. Phys. Chem. A*, 10493.**

245. **Schweitzer-Stenner, R., Huang, Q., Hagarman, A., Laberge, M., and Wallace, C. J. A. (2007) Static Normal Coordinate Deformations of the Heme Group in Mutants of Ferrocytochrome c from *saccharomyces cervisiae* Probed by Resonance Raman Spectroscopy., *J. Phys. Chem. B 111*, 6527.**

246. **Ulmer, D. D. (1965) Optical rotary dispersion of oxidized and reduced cytochrome c, *Biochemistry 4*, 902-907.**

247. **Urry, D. W., and Doty, P. (1965) On the conformation of horse heart ferri- and ferrocytochrome-c, *J. Am. Chem. Soc. 87*, 2756-2758.**

248. **Myer, Y. P. (1968) Conformation of cytochromes. III. Effect of urea, temperature, extrinsic ligands and pH variation on the conformation of horse heart ferricytochrome-c, *Biochemistry 7*, 765-776.**

249. **Drucker, H., Campbell, L. L., and Woody, R. W. (1970) Optical rotary properties of the cytochrome c_8 species of *Desulfovibrio*, *Biochemistry 9*, 1519-1527.**

250. **Hsu, M. C., and Woody, R. W. (1971) The origin of the heme cotton effects in myoglobin and hemoglobin, *J. Am. Chem. Soc. 93*, 3515-3525.**

251. **Blauer, G., Sreerama, N., and Woody, R. W. (1993) Optical activity of hemoproteins in the Soret region. circular dichroism of the heme undecapeptide of cytochrome c in aqueous solution, *Biochemistry 32*, 6674-6679.**

252. **Kiefl, C., Sreerama, N., Haddad, R., Sun, L., Jentzen, W., Lu, Y., Qiu, Y., Shelnutt, J. A., and Woody, R. W. (2002) Heme distortions in sperm-whale carbonmonoxy myoglobin: correlations between rotational strengths and heme distortions in MD-generated structures, *J. Am. Chem. Soc. 124*, 3385-3394.**

253. **Woody. (2004) The circular dichroism of protein folding intermediates *Meth. Enzymol. 380*, 242-351.**

254. **Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal. Biochem. 209*, 32-44.**

255. **Tinoco, I. (1962) Theoretical aspects of optical activity, *Adv. Chem. Phys. 4*, 113-160.**

256. **Woody, R. W. (2009) Circular Dichroism Spectrum of Peptides in the Poly(Pro)II Conformation, *J. Am. Chem. Soc. 131*, 8234-8245.**

257. **Wallace, B. A. (2000) Synchrotron radiation circular-dichroism spectroscopy as a tool for investigating protein structures, *J. Synchrotron Rad 7*, 289-295.**

258. **Greenfiled, N. J. (1996) Methods to Estimate the Conformtation of Proteins and Polypeptides from Circular Dichroism Data, *Analytical Biochemistry 235*, 1-10.**

259. **Böhm, G., Muhr, R., and Jaenicke, R. (1992) Quantitative-Analysis of Protein Far UV Circular-Dichroism Spectra by Neural Networks, *Protein Eng. 5*, 191-195.**

260. **Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) Evaluation of Secondary Structure of Proteins from UV Circular-Dichroism Spectra Using an Unsupervised Learning Neural-Network, *Protein Eng. 6*, 383-390.**

261. **Merelo, J. J., Andrade, M. A., Prieto, A., and Morán, F. (1994) Proteinotopic Feature Maps, *Neurocomputing 6*, 443-454.**

262. **Provencher, S. W., and Glöckner, J. (1981) Estimation of Globular Protein Secondary Structure from Circular Dichroism, *Biochemistry 20*, 33-37.**

263. **Manavalan, P., and Johnson, W. C., Jr. (1987) Variable Selection Method Improves the Prediction of Protein Secondary Structure from Circular Dichroism Spectra, *Anal. Biochem. 167*, 76-85.**

264. **Sreerama, N., and Woody, R. W. (1994) Poly(pro)II Helices in Globular Proteins - Identification and Circular Dichroic Analysis, *Biochemistry 33*, 10022-10025.**

265. **Sreerama, N., and Woody, R. W. (1994) Protein Secondary Structure fro Circular Dichroism Spectroscopy - Combining Variable Selection Principle and Cluster Analysis with Neural Network, Rigid Regression and Self-Consistent Methods, *J. Mol. Biol. 242*, 497.**

266. **Black, K. M., and Wallace, C. J. A. (2007) Probing the role of the conserved beta-II turn Pro-76/Gly-77 of mitochondrial cytochrome c, *Biochem. Cell Biol. 85*, 366-374.**

267. **Sutherland, J. C., Desmond, E. J., and Takacs, P. Z. (1980) Versatile Spectrometer for Experiments Using Synchrotron Radiation at Wavelenghts Greater than 100 nm, *Nucl. Instr. Meth. 172*, 195-199.**

268. **Snyder, P. A., and Rowe, E. M. (1980) The 1st use of Synchrotron Radiation for Vacuum Ultraviolet Circular Dichroism Measurements, *Nucl. Instr. Meth. 172*, 345-349.**

269. **Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Research 32*, W668-W673.**

270. **Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases, *Biopolymers 89*, 392-400.**

271. **Parrish, J. C., Guillemette, J. G., and Wallace, C. J. A. (2001) A tale of two charges: Distinct roles for an acidic and a basic amino acid in the structure and function of cytochrome c, *Biochem. Cell. Biol. 79*, 83-91.**

272. **Wallace, C. J. A., and Clark-Lewis, I. (2000) Site-specific independent double labeling of proteins with reporter atoms, *Biochem. Cell. Biol. 78*, 79-86.**

273. **Addison, A. W., and Stephanos, J. J. (1986) Nitrosyl (III) hemoglobin: autorerduction and spectroscopy, *Biochemistry 25*, 4104-4113.**

274. **Spilotros, A., Levantino, M., and Cupane, A. (2010) Conformational substates of ferricytochrome c revealed by combined optical absorption and electronic circular dichroism spectroscopy at cryogenic temperature, *Biophys. Chem. 147*, 8-12.**

275. **Schweitzer-Stenner, R., Shah, R., Hagarman, A., and Dragomir, I. (2007) Conformational substates of horse heart cytochrome c exhibit different thermal unfolding of the heme cavity, *J. Phys. Chem. B 111*, 9603-9607.**

276. **Shah, R., and Schweitzer-Stenner, R. (2008) Structural changes of horse heart ferricytochrome c induced by changes of ionic strength and anion binding, *Biochemistry 47*, 5250-5257.**

277. **Davis, L. A., Schejter, A., and Hess, G. P. (1973) Alkaline Isomerization of oxidized cytochrome c, *J. Biol. Chem. 249*, 2624-2632.**

278. **Verbaro, D., Hagarman, A., Soffer, J., and Schweitzer-Stenner, R. (2009) The pH Dependence of the 695 nm Charge Transfer Band Reveals the Population of an Intermediate State of the Alkaline Transition of Ferricytochrome c at Low Ion Concentrations, *Biochemistry 48*, 2990-2996.**

279. **Weinkam, P., Zimmermann, J., Sagle, L. B., Matsuda, S., Dawson, P. E., Wolynes, P. G., and Romesberg, F. E. (2008) Cahracterization of Alkaline Transitions in Ferricytochrome c Using Carbon- Deuterium Infrared Probes, *Biochemistry 47*, 13470-13480.**

280. **Hagarman, A., Duitch, L., and Schweitzer-Stenner, R. (2008) The Conformational Manifold of Ferricytochromec Explored by Visible and Far-UV Electronic Circular Dichroism Spectroscopy, *Biochemistry 47*, 9667-9677.**

281. **Filosa, A., Ismail, A. A., and English, A. M. (1999) FTIR-monitored thermal titration reveals different mechanisms for the alkaline isomerization of tuna compared to horse and bovine cytochromes c, *JBIC 4*, 717-726.**

282. **Krishna, M. M. G., Maity, H., Rumbley, J. N., Lin, Y., and Englander, S. W. (2006) Order of steps in the cytochrome c folding pathway: Evidence for a sequential stabilization mechanism, *J. Mol. Biol. 359*, 1410-1419.**

283. **Lees, J. G., Smith, B., Wien, F., Miles, A., and Wallace, B. A. (2004) *CDtool*—An integrated software package for circular dichroism spectroscopic data processing, analysis and archiving, *Anal. Biochem. 332*, 285-289.**

284. **Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics 18*, 211-212.**

285. **Manas, E. S., Wright, W. W., Sharp, K. A., Friedrich, J., and Vanderkooi, J. M. (2000) The influence of protein environment on the low temperature electronic spectroscopy of Zn-substituted cytochrome c, *J. Phys. Chem. B 104*, 6932-6941.**

286. **Prabhu, N. V., Dalosto, S. D., Sharp, K. A., Wright, W. W., and Vanderkooi, J. M. (2002) Optical Spectra of Fe(II) Cytochrome c Interpreted Using Molecular Dynamics Simulations and Quantum mechanical Calculations., *J. Phys. Chem. B 106*, 5561-5571.**

287. **Banci, L., Bertini, I., Redding, T., and Turano, P. (1998) Monitoring the conformational flexibility of cytochrome c at low ionic strength by H-NMR spectroscopy, *Eur. J. Biochem. 1998*, 271-278.**

288. **Uchiyama, S., Hasegawa, J., Tanimoto, Y., Moriguchi, H., Mizutani, M., Igarashi, Y., Sambongi, Y., and Kobayashi, Y. (2002) Thermodynamic characterization of variants of mesophilic cytochrome c and its thermophilic counterpart, *Protein Eng. 15*, 445.**

289. **Rush, T. S., Kozlowski, P., Piffat, C. A., Kumble, R., Zgierski, M. Z., and Spiro, T. G. (2000) Computational modeling of metalloporphyrin structure and vibrational spectra: Porphyrin ruffling in NiTPP, *J. Phys. Chem. B 104*, 5020-5034.**

290. **Li, X. Y., Czernuszewicz, R. S., Kincaid, J. R., and Spiro, T. G. (1989) Consistent Porphyrin Force Field .3. Out-of-plane Modes in the Resonance Raman Spectra of Planar and Ruffled Nickel Octaethylporphyrin, *J. Am. Chem. Soc. 111*, 7012-7023.**

291. **Unger, E., Beck, M., Lipski, R. J., Dreybrodt, W., Medforth, C. J., Smith, K. M., and Schweitzer-Stenner, R. (1999) A new method for evaluating the conformations and normal modes of macromolecule vibrations with a reduced force field. 2. Application to nonplanar distorted metal porphyrins, *J. Phys. Chem. B 103*, 10022-10031.**

292. **Tsai, H.-H., and Simpson, M. C. (2004) Isolated impact of ruffling on the vibrational spectrum of Ni porphyrins. Diagnosing out-of-plane distortions, *J. Phys. Chem. A 108*, 1224-1232.**

293. **Hu, S. Z., Smith, K. M., and Spiro, T. G. (1996) Assignment of protoheme Resonance Raman spectrum by heme labeling in myoglobin, *J. Am. Chem. Soc. 118*, 12638-12646.**

294. **Hu, S. Z., Morris, I. K., Singh, J. P., Smith, K. M., and Spiro, T. G. (1993) Complete Assignment of CYtochrome c Resonance Raman Spectra via Enzymatic Reconstitution with Isotopically Labeled Hemes, *J. Am. Chem. Soc. 115*, 12446-12458.**

295. **Hagarman, A., Wallace, C. J. A., Laberge, M., and Schweitzer-Stenner, R. (2008) Out-of-plane Deformations of the Heme Group in Different Ferrocytochrome c Proteins Probed by Resonance Raman Spectroscopy, *J. Raman Spectrosc. 39*, 1848-1858.**

296. **Lipski, R., Unger, E., Dreybrodt, W., Militello, V., Leone, M., and Schweitzer-Stenner, R. (2001) Vibrational analysis of Ni(II)- and Cu(II)-octamethylchlorin by polarized resonance Raman and Fourier transformed infrared spectroscopy., *J. Raman Spectrosc. 32*, 521-541.**

297. **Lemke, C., Schweitzer-Stenner, R., Shelnutt, J. A., Quirke, J. M. E., and Dreybrodt, W. (2001) Vibrational Analysis of Metalloporphyrins with Electron-Withdrawing NO2 Substituents at Different Meso Positions, *J. Phys. Chem. A 105*, 6668-6679.**

298. **Huang, Q., and Schweitzer-Stenner, R. (2005) Non-planar heme deformations and excited state displacements in horseradish peroxidase detected by Raman spectroscopy at Soret excitation, *J. Raman Spectrosc. 36*, 363-375.**

299. **Kitagawa., T., Abe., M., Kyogoku., Y., Oposhi., H., Watanabe, E., and Yoshida, Z. (1976) Resonance Raman spectra of metallooctaethylporphyrins. Low frequency vibrations of porphyrin and iron-axial ligand stretching modes, *J. Phys. Chem. 80*, 1181.**

300. **Cartling, B. (1988) in Biological Application on Raman Spectroscopy, 217.**

301. **Unger, E., Bobinger, U., Dreybrodt, W., and Schweitzer-Stenner, R. (1993) Vibronic coupling in Ni(II) Porphine Derived from Resonant Raman Excitation Profiles, *J. Phys. Chem. B 97*, 9956-9968.**

302. **Shelnutt, J. A., p http://jasheln.unm.edu.**

Appendix A: List of Abbreviations

α_l – left handed alpha helix

α_r – right handed alpha helix

ρ – depolarization ratio

aβ – antiparallel beta strand

aβt – transition region between polyproline II and beta strand

Asx – asparagine or aspartic acid

APAF-1 – apoptotic protease activating factor 1

ATP – adenosine triphosphate

bhc – bovine heart cytochrome c

Bis/Tris – (Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane)

CCD – charge-coupled device

CCL_4 – carbon tetrachloride

CD – circular dichroism

CHASA – conditional hydrophobic accessible surface area

chc – chicken cytochrome c

Complex III – cytochrome bc_1 complex/ cytochrome reductase

Complex IV – cytochrome oxidase

CS_2 – carbon disulfide

CT – charge transfer

D_2O – deuterium oxide

DCl – deuterium chloride

DFT – density functional theory

DMSO – dimethyl sulfoxide

DPR – depolarization ratio

E.COSY – enhanced correlation spectroscopy

EPR – electron paramagnetic resonance

FIPH – Flory isolated pair hypothesis

FTIR – Fourier transform infrared

HCl – hydrochloric acid

hhc – horse heart cytochrome c

HOMO – highest occupied molecular orbital

HPLC – high performance liquid chromatography

HSQC – heteronuclear spin quantum correlation

HT – Herzberg-Teller

I_x – Raman intensity scattered with a polarization parallel to the incoming excitation frequency

I_y - Raman intensity scattered with a polarization perpendicular to the incoming excitation frequency

IDP – intrinsically disordered protein

ip – in-plane

kDa – kilodalton

LCAO-MO – linear combination of atomic orbitals – molecular orbitals

LUMO – lowest unoccupied molecular orbital

MD – molecular dynamics

MOPS – 3-(N-morpholino)propanesulfonic acid

N – native

NADH – H-nicotinamide adenine dinucleotide

NaOH – sodium hydroxide

NMR – nuclear magnetic resonance

NSD – normal coordinate structural decomposition

oop – out-of-plane

ORD – optical rotary dispersion

pβ – parallel beta strand

PDB – protein data bank

PPII – polyproline II

ROA – Raman optical activity

SAXS – small angle

SNCD – static normal coordinate deformation

SR – synchrotron radiation

ST – serine or threonine

TFA – trifluoroacetic acid

TFE – trifluoroethanol

TOCSY – total correlation spectroscopy

Tris – tris(hydroxylmethyl)aminomethane

TSP – 3-(trimethylsilyl) propionic-2,2,3,3-d_4 acid

U – unfolded

UV – ultra violet

VCD – vibrational circular dichroism

VUV – vacuum ultraviolet

Wat – water

XAO –$X_2A_7O_2$

yc (C102T) – yeast cytochrome c with a mutated cysteine to a threonine at position 102

Vita

Andrew Michael Hagarman was born in Syracuse, NY, USA on October 6, 1982 to Dr. James A. and Patricia V. Hagarman. Shortly thereafter in 1984, they moved to Audubon, PA, where they have resided since. He attended Drexel University and received a Bachelor of Science in Chemistry. Upon completion of his undergraduate degree, he stayed at Drexel University under the guidance of Dr. Reinhard Schweitzer-Stenner in the Chemistry department and along the way to completing his PhD, received his Masters of Science in Chemistry with a focus in biophysical spectroscopy. He has been a contributing author on sixteen peer reviewed journal articles, with more in preparation, and two book articles throughout his research career. Of those peer reviewed journal articles, he was a first author on four. He has presented his work orally and in poster format at many local, national and international conferences. During his time in the graduate program in the Chemistry department at Drexel he taught many physical and general chemistry labs. Upon starting his research career he was able to travel to the University of Puerto Rico, Rio Piedras in Puerto Rico to learn vibrational spectroscopic techniques. Also, he was able to travel abroad and work in the laboratory of Dr. Harald Schwalbe in Frankfurt, Germany at Johan Wolfgang Goethe University learning advanced NMR techniques. He traveled to Biotools Inc. in Jupiter, Florida to study Raman Optical Activity as well. While in his final year in the PhD program at Drexel he got engaged to Maria Alessi while in Paris, France and the two of them gave birth to their first child, Matthew Augustus Hagarman on January 8, 2010. Upon completion of the PhD program at Drexel, he will be starting a Post-Doctoral research position at Duke University under the guidance of Dr. Terrence Oas.

Publications Contributing to this Thesis

1. Hagarman, A., Measey, T. J., Mathieu, D., Schwalbe, H., and Schweitzer-Stenner, R. (2010) Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I' band profiles and NMR scalar coupling constants, *J. Am. Chem. Soc. 132*, 540-551.

2. Dragomir, I., Hagarman, A., Wallace, C., and Schweitzer-Stenner, R. (2007) Optical band splitting and electronic perturbations of the heme chromophore in cytochrome c at room temperature probed by visible electronic circular dichroism spectroscopy, *Biophys. J. 92*, 989-998.

3. Schweitzer-Stenner, R., Shah, R., Hagarman, A., and Dragomir, I. (2007) Conformational substates of horse heart cytochrome c exhibit different thermal unfolding of the heme cavity, *J. Phys. Chem. B 111*, 9603-9607.

4. Schweitzer-Stenner, R., Gorden, J. P., and Hagarman, A. (2007) The asymmetric band profile of the soret band of deoxymyoglobin is caused by electronic and vibronic perturbations of the heme group rather than by a doming deformation, *J. Chem. Phys. 127*, 135103.

5. Verbaro, D., Hagarman, A., Soffer, J., and Schweitzer-Stenner, R. (2009) The pH Dependence of the 695 nm Charge Transfer Band Reveals the Population of an Intermediate State of the Alkaline Transition of Ferricytochrome c at Low Ion Concentrations, *Biochemistry 48*, 2990-2996.

6. Hagarman, A., Duitch, L., and Schweitzer-Stenner, R. (2008) The Conformational Manifold of Ferricytochromec Explored by Visible and Far-UV Electronic Circular Dichroism Spectroscopy, *Biochemistry 47*, 9667-9677.

7. Hagarman, A., Wallace, C. J. A., Laberge, M., and Schweitzer-Stenner, R. (2008) Out-of-plane Deformations of the Heme Group in Different Ferrocytochrome c Proteins Probed by Resonance Raman Spectroscopy, *J. Raman Spectrosc. 39*, 1848-1858.

8. Schweitzer-Stenner, R., Hagarman, A., Verbaro, D., and Soffer, J. (2009) Conformational Stability of Cytochrome c Probed By Optical Spectroscopy, *Meth. Enzymol. 466*, 109-153.