Title

Mapping molecular flexibility of spin labeled proteins on the nanosecond and longer time scales via CW lineshape analysis and osmolyte-perturbation EPR

Author

López, Carlos Javier

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Mapping molecular flexibility of spin labeled proteins on the nanosecond and longer time scales *via* CW lineshape analysis and osmolyte-perturbation EPR

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Biochemistry and Molecular Biology

by

Carlos Javier López

2012

ABSTRACT OF THE DISSERTATION

Mapping molecular flexibility of spin labeled proteins on the nanosecond and longer time scales *via* CW lineshape analysis and osmolyte-perturbation EPR

by

Carlos Javier López

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2012

Professor Wayne L. Hubbell, Chair

Proteins in solution exhibit structural fluctuations on a wide range of characteristic time scales, from fast backbone motions in the ps-ns range to slow conformational exchange in the μs-ms time domain. These motions can play important roles in function, thus elucidation of molecular mechanisms underlying function requires experimental techniques capable of measuring motions on these time scales. Site directed spin labeling and EPR spectroscopy (SDSL-EPR) has been established as an important tool in protein science for studying structure, identifying dynamically disordered sequences, and monitoring conformational switching triggered by chemical or physical signals in soluble and membrane-bound proteins. However, it remains to be investigated whether fast backbone motions contribute to the spectra of the spin label side chain R1 in well-ordered sequences of regular secondary structure, and how slow

conformational exchange between substates can be detected at equilibrium; these are goals of this dissertation.

The results from this work support a model wherein variations in the EPR spectral lineshape of R1-labeled proteins measures the amplitude of fast but constrained backbone fluctuations on the ns time scale, and show that these motions are correlated with the local packing density. In addition, it is shown that osmolyte perturbation SDSL (OP-SDSL) can be used to identify conformational equilibria between substates that have characteristic lifetimes > 100 ns. The newly-developed OP-SDSL strategy was employed along with continuous wave (CW) lineshape analysis to map molecular flexibility of sperm whale myoglobin in folded and partially folded states and of core-repacking mutants of T4 lysozyme. For myoglobin, the results reproduce and extend earlier NMR studies, thus validating the OP-SDSL method and highlighting advantages of the inherent EPR time scale. For T4 lysozyme, it is found that core repacking mutations that generate internal cavities can give rise to new conformational substates in solution, despite the fact that earlier studies show that the corresponding crystal structures of WT and mutants were essentially identical. The results also show that ligand binding to some of the engineered cavities dramatically shifts the populations of substates towards the native-like state.

Collectively, the results of this research advance the SDSL-EPR methodology for facile mapping of protein flexibility over the important ps-ms time domain and demonstrate the utility of the technology for exploring the molecular basis of protein function.

The dissertation of Carlos Javier López is approved.

James U. Bowie

Arnold J. Berk

Dean Bok

Robert T. Clubb

Wayne L. Hubbell, Committee Chair

University of California, Los Angeles

2012

To my family and love ones for their continued support throughout my professional career.

TABLE OF CONTENTS

Chapter 6 Study of the effects of hydrophobic core substitutions on the conformational flexibility of T4 lysozyme by site-directed spin labeling

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Wayne L. Hubbell for his guidance and support and for sharing his scientific wisdom with me throughout my graduate career, all of which made possible the completion of this dissertation. I would like to thank Cherie Hubbell for her support, patience, encouragement, and for helpful comments on this manuscript. I am grateful to the members of my thesis committee for their guidance and insightful questions. I am grateful to past and current members of the Hubbell laboratory for their contribution to my scientific development during the last 6 years. In particular, I would like to thank Drs. Mark R. Fleissner and Michael D. Bridges for sharing their technical expertise and for the invaluable in-depth scientific discussions and career advices. Dr. Mark R Fleissner kindly provided some of the data shown in Figures 6.27 and 6.28. Thanks to Dr. Christian Altenbach for helpful comments on the manuscript of chapter 2.

Finally, I would like to acknowledge Dr. Joseph Horwitz for technical assistance in the use of the CD spectrometer and Evan Brooks for providing technical assistance in generating the α-helix I mutants in rIFABP and for his support in the expression and purification of the T4 lysozyme K→R variants and mutants bearing the p-AcPhe unnatural amino acids. The project presented in chapter 3 is a modified version of a published article in the Protein Science journal (2009), vol.18; 1637-1652 with permission to reproduce the copyright material limited to this dissertation granted by John Wiley and Sons (License #: 2816670563492). The project presented in chapter 4-5 was done in collaboration with Shirley Oga, while the work presented in chapter 7 was done in a collaborative effort with Dr. Mark. R. Fleissner and Liming Tan. This work was supported by NIH grant 5R01 EY005216 and the Jules Stein Professor Endowment.

VITA

1999-2000	Dean's List Award University of Puerto Rico, Mayaguez
2002-2004	Summer Undergraduate Research Program trainee sponsored by Alfred P. Sloan Foundation
2004	Pharmaceutical Industry internship as process support scientist Abbott Laboratories Barceloneta, Puerto Rico
2004	B.Sc., Industrial Biotechnology Magna Cum Laude University of Puerto Rico, Mayaguez Campus. Mayaguez, Puerto Rico
2005	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles
2007	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

López, C.J., Esposito, E., and Martin, C.T. 2002. Test of a model for translocation-driven promoter release by T7 RNA polymerase. 3rd Annual Biomedical Research conference for minority students (ABRCMS). New Orleans, LA. (poster)

Van Eps, N., **López, C.J.**, and Hubbell, W.L. 2008. Identifying conformational exchange in proteins with site-directed spin labeling. 52nd Annual Biophysical Society Meeting. Long Beach, CA. (poster).

López, C.J., Fleissner, M.R., Guo, Z., Kusnetzow, A.K., and Hubbell, W.L. 2009. Osmolyte-perturbation reveals conformational equilibria in spin-labeled proteins. *Protein Sci*, 18:1637-1652.

López, C.J. 2009.Osmolyte-perturbation reveals conformational equilibria in spin-labeled proteins. 15th Annual Vision Science Conference. Lake Arrowhead, California (talk).

López, C.J., Oga, S., and Hubbell, W.L. 2010. Exploring conformational flexibility of partially folded states of apomyoglobin with site-directed spin labeling. The 24th Annual Symposium of the Protein Society. San Diego, California (poster).

López, C.J., Oga, S., and Hubbell, W.L. 2010. Exploring conformational flexibility of partially folded states of apomyoglobin with site-directed spin labeling. Joint Conference of the 14th In Vivo ESR/EPR Spectroscopy & Imaging and the 11th International EPR Spin Trapping/Spin Labeling. San Juan, Puerto Rico (poster).

Armstrong, B.D., Choi, J., **López, C**., Wesener, D.A., Hubbell, W., Cavagnero, S., and Han, S. 2011. Site-specific hydration dynamics in the nonpolar core of a molten globule by dynamic nuclear polarization of water. *J. Am. Chem. Soc*. 113:5987-5995.

López, C.J. and Hubbell, W.L. 2012. Structural and dynamic response to core repacking substitution in T4 lysozyme detected by site-directed spin labeling. 56th Annual Biophysical Society Meeting. San Diego, California (poster).

López, C.J., Oga, S., and Hubbell, W.L. 2012. Mapping molecular flexibility of partially folded states of proteins with site-directed spin labeling: A case study of myoglobin. (paper in preparation).

Chapter 1: Introduction to protein dynamics

1.1 Historical perspective: From lock-and-key to dynamic energy landscapes

By the late 1950's it was clear that proteins were not static systems and must move in order to perform their function [1]. In fact Daniel Koshland realized that the lock-and-key model proposed by in Emeril Fisher in the 19th century [2] was not appropriate to explain the physical nature of enzyme specificity, thus he proposed a new mechanism of molecular recognition known as the induced-fit model, which posits that substrate binding to an enzyme leads to a change in conformation of the protein, whereby specific functional groups of the enzyme are brought into their proper position to catalyze the reaction [3]. However, both the lock-and-key and induced-fit models consider the protein as being in a single conformation prior to the substrate-binding event. Within two years after the introduction of the induced-fit paradigm, Kendrew [4] and Perutz [5] reported the first x-ray structures of myoglobin and hemoglobin, respectively, which provided a rigid and static view of these molecules. On the other hand, experimental evidence from techniques such as fluorescence [6], nuclear magnetic resonance [7], and hydrogen exchange experiments [8] depicted a different view where proteins in solution behave more as "kicking and screaming stochastic molecules" [9], whereby protein motions allow even interior groups to become accessible to the solvent. In 1976, Copper indicated these transient fluctuations were an intrinsic property of proteins at thermodynamic equilibrium due to their small size and limited compressibility [10]. Despite the significant progress in the study of protein structure, dynamics, and function during these early years, a comprehensive theory that could describe the relation between *equilibrium* fluctuations of proteins in solution with functional properties was missing until elegant flash photolysis experiments done by Austin and

coworkers on myoglobin shed light on the relation between protein flexibility and the molecular mechanism of protein function [12].

1.1.1 Flash photolysis experiments in myoglobin leading to the energy landscape theory. To elucidate the mechanism of oxygen binding to myoglobin, Antonini and Brunori [11] performed flash photolysis experiments to measure the kinetics of light-activated dissociation and rebinding of oxygen to myoglobin at physiological temperatures. These studies suggested that the reaction of oxygen with myoglobin $Mb + O_2 \Leftrightarrow MbO_2$ was a simple one-step process. However, flash-photolysis experiments of MbCO and MbO_2 done at cryogenic temperatures ($T < 200$ K) in order to freeze-out any motion, revealed that the rebinding of the diatomic ligands to the heme active center was non-exponential in time [12]. These findings were interpreted as suggesting that myoglobin does not adopt a single conformation, but it rather exists as a large ensemble of different structures in solution, each of which can exhibit a different rebinding rate. The kinetic experiments suggested a path of CO migration inside the myoglobin molecule that could be approximated using the following reaction model [12]:

$$A \leftarrow B \Leftrightarrow C \Leftrightarrow D \Leftrightarrow S, \tag{1.1}$$

where S represents the CO in the solvent, B, C, D represent CO in internal pockets in the myoglobin molecule, and A represents CO covalently bound to the heme iron. The kinetic data recorded at physiological temperatures suggested that states B, C, D, and S were in *pre-equilibrium*. Close examination of the high-resolution structure of myoglobin provided no discernible pathway for the migration of the diatomic ligand inside the protein. For many years the locations of the B, C, and D remained unknown, however advances in cryogenic x-ray crystallography [13] time-resolved x-ray diffractions experiments [14,15], and molecular dynamic simulations [16] have revealed the locations of these “states” and the migration

pathway of the ligand inside the protein. The results from these studies are summarized in Figure 1.1.

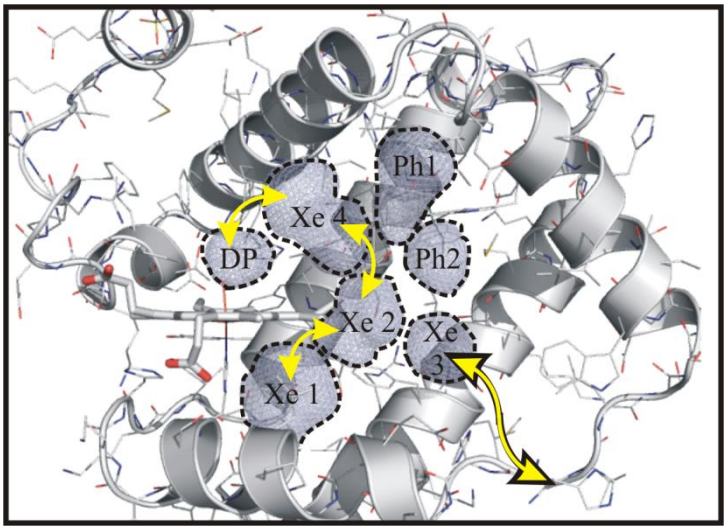


Figure 1.1: Pathway of ligand migration inside of myoglobin. Ribbon diagram of the myoglobin structure showing the internal cavities and pathway of migration of diatomic ligands (*i.e.*,O_2, CO) as inferred from studies by Tomita et al. [15]. The molecular surfaces of the internal cavities are shown as mesh representation and are enclosed by dashed lines. The primary docking site for diatomic molecules after dissociation is known as distal heme pocket (DP). Four of the internal cavities important for the internal migration are also xenon-binding sites as shown by Telton et al. [17] (pdb: 1J52), thus they are denoted Xe 1 – Xe 4. Additional cavities known as phantom 1 and 2 are shown. The ligand migration pathway is indicated by the yellow arrows. State A proposed by Austin et al. [12] corresponds to CO covalently bound to the heme iron, states B, C, and D correspond to the ligand in the DP, Xe 4 cavity

These studies on myoglobin lead to a new paradigm to describe the molecular mechanism of protein function, in which protein motions can be described by an energy landscape, whereby the primary sequence can fold into a large number of different structures known as substates [12,18]. The energy landscape is a construct of 3N dimensions, where N is the # of atoms of in the protein and the hydration shell [19], and is organized in hierarchical tiers, with valleys within valleys within valleys as depicted in Figure 1.2A. At the top of the hierarchy, in tier 0, are the taxonomic substates, also known as conformational substates, which are generally small in number and live sufficiently long, so that their properties can be measured and described in molecular detail. Transition between conformational substates (conformational exchange) involves high-amplitude, low frequency fluctuations of the peptide backbone between states separated by energy barriers of several *k*T (Figure 1.2B). Within each conformational substate, there are a large number of tier 1 substates that are separated by energy barriers of less than 1 *k*T. Tier 1 substates are numerous and short-lived, so their properties cannot be described in molecular detail, thus a taxonomic approach is impractical since these states can only be described statistically. These states are known as statistical substates and transitions among them involve small-amplitude nanosecond-picoseconds motions of the backbone and side-chain about their average positions. Each statistical substate contains a

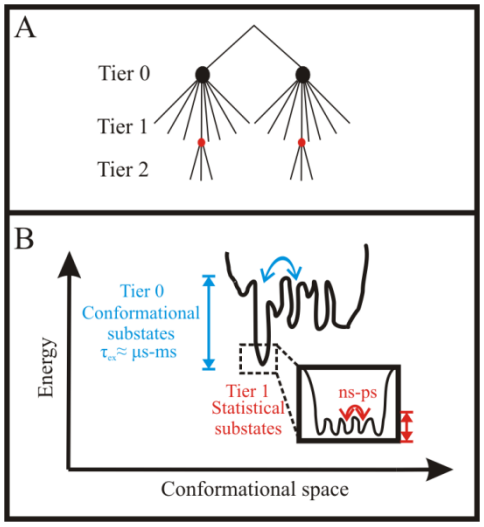


Figure 1.2: Diagrammatic representation of the free energy landscape. A) Tree diagram of the taxonomic (tier 1), statistical (tier), and few-level substates (tier 2). B) Cross section of the free energy landscape. Each "valley" in the tier 0 region corresponds to a conformational substate in equilibrium. Within each tier 0 valley, a large numbers of tier 1 valleys exist in dynamic equilibrium. The relative difference in free energy between tier 0 and tier 1 valleys are highlighted. The curved arrows represent transitions between substates.

small number of substates with lower energy barriers known as "few-level substates" that interconvert even at a mK temperature range [19].

1.1.2 *Dynamic energy landscapes and protein function*. The distribution of the populations of conformational substates in the energy landscape is determined by their relative free energy and the energy barriers separating them. The energy landscape, however, is dynamic [20,21] and coupled to a particular set of temperature, pressure, and solvent conditions, thus changes in the surrounding, by either a physical (*i.e.*, temperature pressure, pH) or functional signal (*i.e.*, ligand binding, protein-protein interaction), modify the shape of the energy landscape with a concomitant re-distribution of the relative populations of substates (Figure 1.3A). In regard to protein function, each conformational substate can have different functional properties. For example, myoglobin has three taxonomic states in dynamic equilibrium, namely A_0, A_1, and A_3, each of which exhibit different ligand rebinding rates to the heme iron [20]. In recent years it has been suggested that the conformational states A_1 and A_0 might also have different functions [22], such that A_1, which is more populated at high pH, stores oxygen and A_0, which is more populated at low pH, catalyzed the conversion of the signaling molecule NO to NO^-_3 (Figure 1.3B).

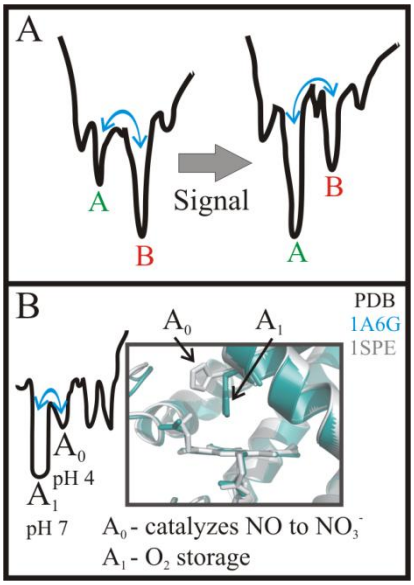


Figure 1.3: Dynamic energy landscape. A) Redistribution of the population in equilibrium due to a chemical or functional signal. Hypothetical conformational substates A and B are indicated. B) Two of the three taxonomic substates of myoglobin in dynamic equilibrium are shown. Substate A_0 (pdb 1SPE), is dominant at low pH and substate A_1 (pdb 1A6G) is dominant at high pH. The structural features and functional properties of each substates are indicated.

Transitions between statistical substates are also functionally important. For example, thermal fluctuations on the nanosecond to picosecond time scale, due to transition between

statistical substates, play a role in opening the gates that connect the internal cavities inside myoglobin allowing the ligand to migrate from the solvent (S) to the binding pocket (DP) as shown in Figure 1.1 [23].

This new paradigm of pre-existing equilibrium and dynamic energy landscape proposed by Frauenfelder and others [18,20,21,24] to describe the molecular mechanism of protein function is juxtaposed to the induced-fit model often encountered in textbooks and is still a matter of intense debate [24,25]. As an example, a series of studies published in the last 10 years attributed observed conformational changes in many proteins associated with protein-protein interaction and catalysis as either involving induced-fit or population shift (Table 1.1).

Table 1.1: Classification of conformational changes associated with protein-protein interaction (from Goh et al., 2004) and catalysis[†]

Protein	Binding partner	State 1 PDB	State 2 PDB	Max. structural changes (Å)	Classification	References
Myosin	Actin			9	Induced fit	Holmes et al., 2003
Spe7	TrxShear3			7	Population Shift	James et al., 2003
Adk[†]	-	2RH5	2RGX	19	Population Shift	Henzler-Wildman, 2007
Gαi subunit	RGS14	1BOF	1KJY	10	Induced fit	Kimple et al., 2002
FliS	FliC	1ORJ	1ORY	29	Induced fit	Evdokimov et al. 2003
PKA[†]	-	1CMK	1ATP	7	Population Shift	Masterson et al., 2011
NtrC	-	1DC7	1DC8	12	Population Shift	Volkman et al., 2001

The apparent contradiction in the mechanistic explanation of molecular function may arise because in certain cases pre-existing equilibrium is difficult to detect experimentally due to the low population (< 5%) of functionally important excited states (*see* example in section 1.2.1). However, based on the extensive experimental evidence supporting each mechanism, a more complex interplay between intrinsic dynamics and protein function has been proposed in the last few years in which both induced fit and conformational selection can be responsible for catalysis [33,34] and protein interactions [35] in a given system.

1.2 Studying protein dynamics

It is now clear that proteins in solution exhibit conformational flexibility over a wide range of time and length scales Figure 1.4. Fast motions occurring in the picosecond to nanosecond time scale generally arise from motions of individual bonds or groups of atoms which are faster than the rotational tumbling of the protein and include atomic vibrations, side chain rotation, and local backbone fluctuations. Motions from the microsecond to second time scale generally involve low frequency-high amplitude conformational rearrangements of several residues and include protein folding, ligand binding, catalysis, substrate binding, and allosteric regulation.

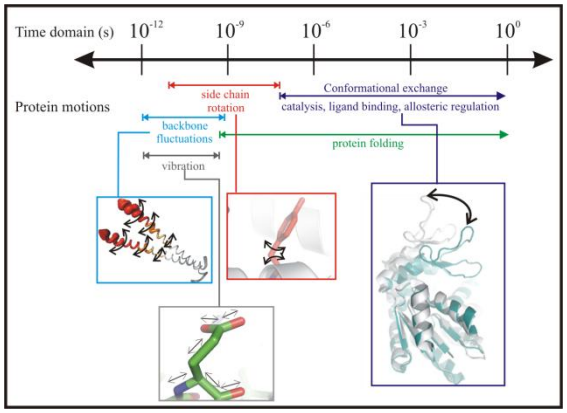


Figure 1.4: Relevant protein motions spanning the picoseconds to second time scale.

In regard to protein function, motions on the nanosecond to picosecond time scale play a role in ligand binding and protein-protein-interactions [36-39], and in providing flexible hinge regions to facilitate large scale slower motions that are important for catalysis [28]. Transitions between conformational substates taking place in the microsecond to millisecond time scale are functionally important in ligand binding [40], catalysis [28,31,41], and their existence gives rise to promiscuity in protein-ligand and protein-protein interaction in signal transduction [42,43]. Thus to understand the molecular mechanism of protein function it is essential to develop

experimental strategies that are capable of studying protein flexibility over a wide-range of time (picoseconds to seconds) and length scales (~ 0.1 to > 10Å).

Various experimental approaches have been developed to study many aspects of protein dynamics. For example, the B-factor values in x-ray crystallography, also known as the Debye-Waller factor, have been used as an indirect measurement of protein flexibility as it gives atomic-level specific information about mean square displacement arising from thermal motions [44,45]. In addition, conformational flexibility in proteins have been deduced by crystallizing the same protein in multiple states (*ie.*, +/- ligand, +/- binding partner, different space group); many examples of which can be found at the database of macromolecular movement established by Gerstein and Krebs [46] (http://bioinfo.mbb.yale.edu/MolMovDB). However, these x-ray based observations do not give information on the time scale of the motions and complication arises due to crystal-lattice contacts and static disorder as both phenomena can affect the experimental B-factor values and select for specific conformers. Solution techniques used to study protein dynamics include fluorescence [47], hydrogen exchange with mass spectrometry [48], raman spectroscopy [49], 2-dimensional infrared spectroscopy [50], NMR spectroscopy [51] and EPR spectroscopy (*see* next chapter).

Solution NMR spectroscopy has been particularly effective in providing high-resolution structural and dynamic information over a wide range of time scales (picosecond to second) of small proteins in solution [35,51]. In NMR, motions taking place in the nanosecond to picosecond time scale can be detected by measuring the nuclear spin relaxation rates[1], namely R_1, R_2, and steady-state heteronuclear NOE [51]. To monitor motions of the backbone on this time regime, most NMR studies focus on studying the reorientation of bond vectors connecting a

[1] A more in depth description of the relaxation processes taking place in magnetic resonance is described in detail in chapter 2 of this dissertation.

pair of NMR active spins, such as the ^{1}H-^{15}N and ^{13}CO-^{13}Cα. Molecular motions occurring in the picosecond to nanosecond time scale cause reorientation of the bond vector with respect to the applied magnetic field, and this reorientation leads to local magnetic field fluctuations as depicted in Figure 1.5, which in turn affects the relaxation rates of the nuclei. The three relaxation rates are related to a spectral density function (*J(ω)*), which is proportional to the amplitude of the fluctuating magnetic field at the frequency ω as is defined by the "model-free" formalism described by Lipari and Szabo, [53,54]:

$$J(\omega) = \frac{S^2\tau_m}{1+\omega^2\tau_m^2} + \frac{(1-S^2)\tau}{1+\omega^2\tau^2}, \tag{1.2}$$

where τ_m is the isotropic rotational diffusion of the protein, S^2 is the order parameter describing the amplitude of motion, and τ is defined by the following relation:

$$\frac{1}{\tau} = \frac{1}{\tau_m} + \frac{1}{\tau_e}, \tag{1.3}$$

where τ_e is the time scale of motion of the bond vector (*ie.*, ^{1}H-^{15}N). In this formalism, S^2 relates to diffusion in a cone of angle θ,

$$S^2_{cone} = \frac{1}{4}cos^2\theta(1 + cos\theta)^2 \tag{1.4}$$

It follows that $S^2 = 1$ indicates complete restriction of the amplitude of motion (θ = 0°) on the nanosecond to picosecond time scale and $S^2 = 0$ indicates unrestricted isotropic internal motions (θ = 90°). In addition to backbone motions, motions of the side chains can also be studied with relaxation techniques [35,51]. To measure protein motions on the microsecond to millisecond time scale, the Lipari-Szabo "model-free" approach is inadequate to describe R_2, thus a new term (R_{ex}) is introduced to account for these slow fluctuations. R_{ex} is the relaxation rate due to conformational exchange occurring in the microsecond to millisecond time scale. However, the "model-free" approach can only provide an estimate of the magnitude of R_{ex}, thus for a more

quantitative characterization of protein fluctuations on the microsecond to millisecond time domain, NMR spectroscopist use a method known as Carr-Pusell-Meiboom-Gill (CPMG) relaxation dispersion [55], which has the capability of identifying functionally important conformational substates that are populated even at low levels (< 5%). The relaxation dispersion method relies on the fact that the spin relaxation rate R_2 can be sensitive to the presence of minor conformers as long as the interconversion rate (k_{ex}) is in the millisecond time scale and there are large chemical shift differences between the states. In the CPMG experiment, the R_{ex} term, arising from chemical shift difference between exchanging states, is suppressed by applying a series of refocusing π pulses in a spin echo experiment[2], such that the effective decay rate of the magnetization (R_2^{eff}) decreases as a function of the frequency of the pulses (ν_{CPMG}[3]) as depicted in Figure 1.6 provided that there is an exchange process taking place. Flat dispersion profiles (Figure 1.6B) for a particular nucleus indicates absence of local conformational exchange, whereas non-flat dispersion profiles are analyzed to obtain thermodynamic (relative populations), kinetic (rate of interconversion: k_{ex}), and structural information (chemical shift differences) of the states. This approach has been useful to study transitions with exchange rates in the ~0.1 MHz to KHz regime [51]. However, it has been reported that in certain cases conformational exchange in the microsecond to millisecond time domain leads to complete loss of NMR spectral line, thus no direct structural or kinetic information can be obtained for such cases [56].

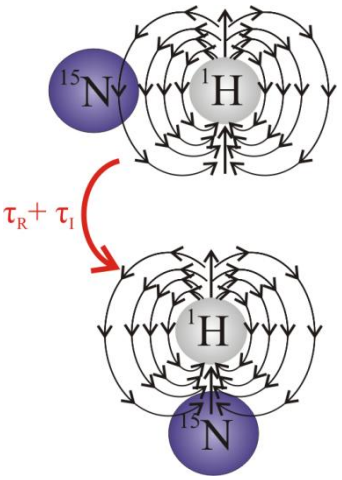


Figure 1.5: Orientation-dependent magnetic field inhomogeneity felt by the NMR-active nucleus ^{15}N covalently bound to a proton ^{1}H. Both protein tumbling (τ_R) and internal motions (τ_I) cause reorientation of the ^{15}N-^{1}H bond vector.

[2] The spin echo experiment is described in detail in chapter 2 of this dissertation.

[3] $\nu_{CPMG} = 1/(4\tau)$, where 2τ is the time interval between the center of the consecutive π pulses [55].

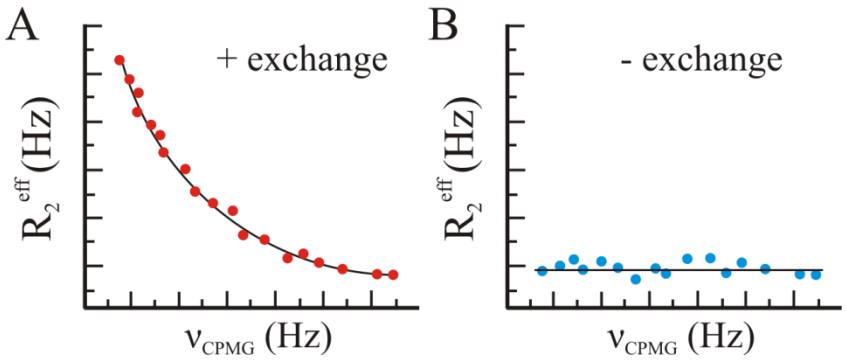


Figure 1.6: Idealized relaxation dispersion profiles of two spins in a protein. Fit of the dispersion profile of a spin sensing exchange is shown in the left panel. A flat dispersion profile is indicative of no exchange.

1.2.1 *Identifying low-lying functionally important excited states by NMR spectroscopy: NMR studies in cavity mutant L99A of T4 lysozyme*. The enzyme T4 lysozyme (T4L) has long been used as a paradigm to understand the relationship between structure, stability, and folding in proteins (*see* review by Baase et al. [57]). Over 700 mutants have been engineered and studied in detail, and ~350 of which have been crystallized in different space groups [57]. The mutant L99A of T4 lysozyme proved to be an excellent candidate for detailed characterization by NMR spectroscopy as the large-to-small substitution at position 99 created a larger nonpolar cavity [58], shown in Figure 1.7A, capable of binding different hydrophobic ligands such as benzene, ethylbenzene, and p-xylene [59-60]. The crystal structure of the mutant form (pdb: 1l90) revealed that the cavity was inaccessible to the solvent and the overall structure was similar that of the wild-type form (pdb: 3lzm), which does not bind any of these ligand. Thus it was hypothesized that a new conformation of the protein not observed in the x-ray structure must exist to expose the cavity to the solvent, thereby allowing ligand entry.

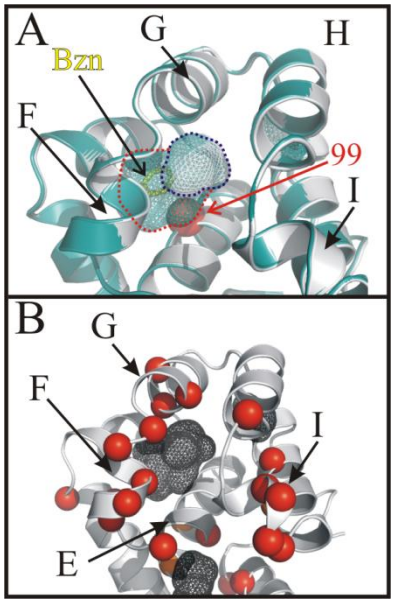


Figure 1.7: Structure and dynamics of the T4L mutant L99A. A) The high resolution structures of WT (gray: 3lzm) and L99A mutant (teal: 1l90) of T4L are superimposed. The benzene molecule bound to the L99A mutant is shown (yellow: 3hh4). Molecular surfaces of native and mutant cavities are shown as wireframe representation. The L99A mutation increases the volume of the native cavity from 34Å^3 to 150Å^3 (*see* blue to red dashed lines). B) Ribbon diagram of T4L. Red spheres at the C_α indicate the residues with large R_{ex} terms [55] and non-flat dispersion curves in the L99A mutant [61].

To obtain insights into the conformational properties and dynamics of the L99A mutant of T4L, Mulder and coworkers [55,61] measured the backbone ^{15}N and ^{1}H and side chain ^{2}H and ^{13}C relaxation rates to probe fast (ps to ns) and intermediate time scale (μs to ms) motion of the mutant and wild-type proteins. Model-free analysis showed essentially no change in the ns to ps motions between the wild-type protein and the L99A mutant, indicating that these motions were not responsible for the ligand binding event. Conversely, many residues in the L99A mutant required new R_{ex} term not observed in the wild-type protein. As discussed above, R_{ex} term serves as an indicator of conformational fluctuations occurring in the microsecond to millisecond time scale. The regions with new or increased R_{ex} terms corresponded to residues within helices E, F, G, and I (colored red in Figure 1.7B). To obtain structural and kinetic information of the substates involved, CPMG relaxation dispersion experiments were performed [61]. In these experiments individual values of k_{ex} for 30 side chain ^{13}C and 50 backbone ^{15}N were fit globally to a two-site exchange process. Global fit of the dispersion profiles resulted in k_{ex} values of 1.03 ± 0.5 kHz and 1.37 ± 0.4 kHz for backbone ^{15}N and side chain ^{13}C spins, respectively. Temperature-dependence studies of rates of interconversion revealed that the exited state (denoted P_e) was 2 kcal/mol higher in free energy than the ground state (P_g) with unfavorable enthalpic contribution of 7 kcal/mol (due to decrease in stabilizing interactions) and a favorable entropic contribution of 5 kcal/mol (more disorder).

This difference in free energy between the excited and ground states corresponds to a relative population for the P_e state of 3%. Interestingly, the exchange rates observed in relaxation dispersion experiments are in the same order of magnitude as the off-rate constants of ligands (0.3 to 0.8 kHz; [62]), thus the authors initially suggested that the conformational exchange process detected by relaxation dispersion and ligand binding events were coupled [61]. However, more recent relaxation dispersion studies in combination with CS-Rosetta modeling provided insights into the structure of the excited state and, most importantly, revealed that the excited state was not responsible for ligand binding and that ligand binding likely occur *via* the ground state by some unknown mechanism (Bouvignies et al. [63]). These well-designed NMR experiments allowed for a complete structural, kinetic, and thermodynamic characterization of an excited state with distinct functional properties, which could not be detected by other methods.

1.3 Populating low-lying excited states with pressure-perturbation methods: High-pressure NMR studies on apomyoglobin

As mentioned above, experiments done in the mid 1970's demonstrated that proteins in solution exist as an ensemble of conformational substates with different structures and distinct functional properties. In pioneer experiments done in the late 1970's, Frauenfelder and co-workers [64][4] studied the rebinding of oxygen and CO to myoglobin at high-pressure. These experiments showed that pressure exerts both *elastic* and *conformational* shifts on proteins (Figure 1.8). The elastic shift refers to a general compression of a given substate without changing its conformation, while the conformational shift refers to a shift in the relative

[4] According to Frauenfelder et al. [64], the results of the studies done in the 1970's were not published until 1990 due to lack of understanding of the results.

populations of the conformational substates in equilibrium. The change in the population distribution at equilibrium was an indication that the different substates have distinct partial molar volumes because pressure favors smaller volumes of the system.

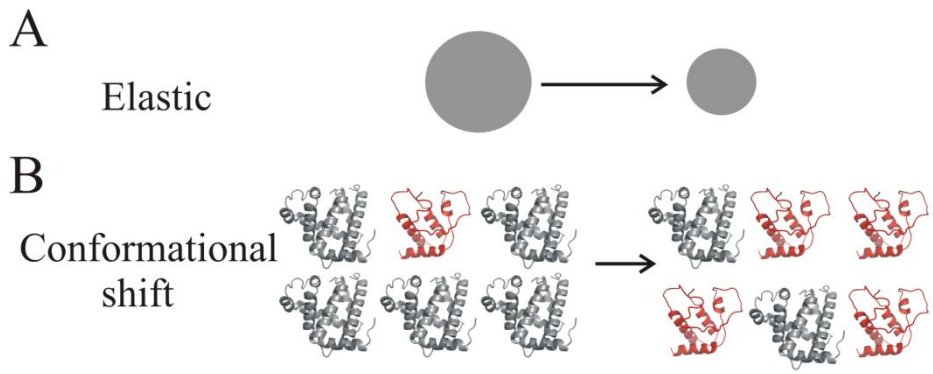


Figure 1.8: Pressure-induced effects on proteins. A) General compression of a conformer. B) Shift in the conformational equilibrium between two states with distinct molar volumes. The red ribbon model represents the substate with lower molar volume.

Since these studies done in the 1970's, many high-pressure studies in combination with x-ray crystallography [65] and various spectroscopic techniques have been done to study the conformational properties of proteins (*see* reviews by Li and Akasaka, [66]). Akasaka realized that pressure perturbation methods could be used as a means to populate functionally important low-lying (~ 1%) excited substates of low partial molar volume for direct spectroscopic detection [67]. To test this hypothesis, high-pressure NMR studies were done in sperm whale apomyoglobin [68]. Apomyoglobin was selected as a model system due to the wealth of information on the structure and dynamics of the protein in multiple folded states. Various conformers of apomyoglobin, namely, native (N), molten globule (MG), and unfolded (U) states have been populated in equilibrium at various pH [70] or populated kinetically [71,72], each of which have been extensively characterized in detail by various methods, including NMR spectroscopy [56,73-75].

To study the effect of pressure perturbation on the conformational properties of apomyoglobin, Kitahara and coworkers did ^{1}H one-dimensional and $^{15}N/1H$ two-dimensional spectroscopy in the pressure range of 30 to 3000 bar [69]. The one- and two-dimensional NMR spectra of the protein at 30 bar exhibited well-dispersed resonances characteristic of a folded structure and only showed disorder for residues 80-101 (highlighted in Figure 1.9), which were previously identified as being conformationally flexible in the μs-ms time scale in NMR studies done at atmospheric pressure [73]. Between 500 to 1500 bar the 2 dimensional spectral intensity of some of the resonances corresponding to residues within helices C, G, part of the helix H, and the CD loop broadened beyond detection, which is a hallmark of structural heterogeneity in those regions. The state populated in this pressure range was denoted I and is shown in Figure 1.9. At 2000 bar, the one- and two-dimensional spectra exhibited no dispersion and loss of signal, respectively, which indicated that the entire sequence lost its native 3 dimensional fold and became highly heterogeneous. These observations suggested that at the pressure range of ~ 2000 bar, apomyoglobin adopted a molten globule state. At 3000bar, the ^{1}H and ^{15}N resonances reappeared at full intensities in the spectral positions corresponding to unfolded polypeptides with a radius of gyration similar to that of the previously characterized acid-unfolded state. These findings suggested that at 3000 bar, a highly flexible unfolded state is populated. Interestingly, Kitahara et al. [69] found that the pressure-induced conformational changes were completely reversible at pressures below 2000 bar, which was interpreted as an indication that these states should exist in an equilibrium mixture at 1 bar.

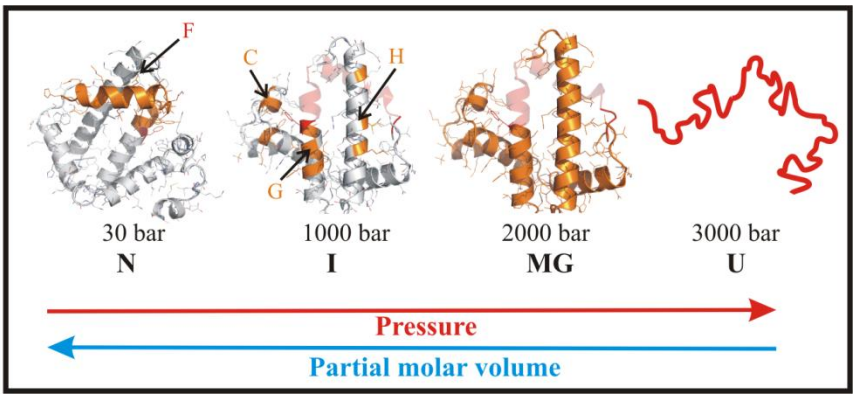


Figure 1.9: Diagrammatic representation of the various conformational states of apomyoglobin populated at the different pressures. The regions that are conformationaly disordered, as judged by NMR line broadening, are highlighted in orange in the ribbon diagram. The completely unfolded polypeptide is shown in red. The four dominant states populated at different pressures are: native (N), intermediate (I), molten globule (MG), and unfolded (U). The red and blue arrows indicate the directionality of increasing values of pressure and partial molar volumes, respectively.

Based on NMR spectral intensities for pressures below 1200 bar, an equilibrium constant was calculated for each pressure. From the fit of the experimental data, the authors concluded that there was a loss in partial molar volume of 75 ml/mol from the N-MG transition and a difference in free energy of 2.5 kcal/mol between the these states at 1 bar, corresponding to a relative population of ~ 1% for the MG state. On the basis of these observations in addition to similar findings in many other proteins [67], Kitahara and co-workers postulated a "*volume theorem*" depicted in Figure 1.9, which posits that the partial molar volume of a protein decreases with disorder. Importantly, this volume theorem predicts that functionally important "invisible" excited states with lower molar volumes can be effectively populated with variable pressure to levels detectable by spectroscopic methods. Thus, pressure-perturbation provides a means to identify and characterize those states.

1.4 Future prospect: Consideration of protein dynamics in protein design strategies.

The field of protein design is an exciting area of research that has emerged in the last few years with the aim of engineering artificial enzymes that are capable of performing important synthetic reactions not catalyzed by naturally occurring enzymes. The design of new biocatalyst is of great practical interest due to its potential applications in the area of therapeutics, bioremediation, chemical synthesis, and biotechnology among others. During the last several years there has been significant progress in the field of protein engineering through rational computational design methods [76], *in vitro* directed evolution methods [78], and *in vivo* production of catalytic antibodies [78].

Notwithstanding the significant progress and success in the design of novel biocatalysts, most artificially designed enzymes and catalytic antibodies exhibit catalytic efficiencies several orders of magnitude lower than natural enzymes. For instance, the biomolecular rate constant (k_{cat}/K_m) of natural enzymes is in the order of 10^6-10^8 $M^{-1}s^{-1}$ with rates of acceleration over uncatalyzed reaction (k_{cat} / k_{uncat}) in the order of 10^6-10^{12}, while the most efficient catalytic antibodies exhibit rate constants of 10^2-10^4 $M^{-1}s^{-1}$ and rates of acceleration of 10^3-10^5 over background [78].

Currently most enzyme design efforts focus (1) on the design of idealized active sites with optimal geometry and electrostatics for the transition state and (2) in the selection of scaffold that can fit the transition state model of the catalytic cycle [79], however the role of dynamics in the actual catalytic transformation, in substrate binding, and product release are often parameters that are overlooked or not selected for. Although it has been generally accepted that substrate binding and product release are important processes that are mediated by conformational fluctuations,

until recently, there was little evidence for the importance of protein flexibility on the actual catalytic step. In a recent study, Bhabha and coworkers [41] demonstrated that conformational fluctuations on the ms time scale in *E. coli* dihydrofolate reductase are essential for the actual chemical step in the reduction of dihydrofolate to tetrahydrofolate, thereby linking dynamics with the chemical transformation event.

Given the connection between protein flexibility and function, the role of dynamics should be a critical parameter added to the protein design strategy such that relevant dynamic modes become part of the selection process along with the optimization of the geometry and electrostatics of the active sites into the biocatalyst of interest. Such an endeavor will require detailed characterization of the conformational properties of newly designed biocatalysts with solution methods such as NMR and EPR spectroscopy. In this way, the observed properties of the designed variants (*i.e.*, catalytic efficiency, substrate binding and release) are explored from the functional dynamics perspective in order to identify "designable dynamic properties" that can be applied in future design strategies; some of these efforts are already underway [80-82]. Furthermore, a combination of computational design methods with directed evolution strategies are currently being used synergistically in protein design experiments, with the former being used for optimizing the geometry and electrostatics of the transition state and the latter being used as a fine-tuning strategy to add mutations that may account for functionally important dynamic modes not included in the original selection strategy [76,79]. These efforts undoubtedly will be essential in the design of more efficient biocatalysts that can match the catalytic efficiencies of natural enzymes.

1.5 Conclusions

As shown above protein dynamics is essential for biological function, thus to understand the molecular mechanism of protein function it is necessary to develop experimental strategies that are capable of studying protein flexibility over a wide-range of time and length scales. Solution NMR spectroscopy has been particularly effective in providing high-resolution structural and dynamic information over a wide range of time scales (picosecond to second) of small proteins in solution, but it is challenged in characterizing large proteins, membrane-bound proteins, and transient complexes. Site-directed spin labeling (SDSL) in combination with electron paramagnetic resonance (EPR) spectroscopy has the potential to provide information on motions occurring in the picosecond to millisecond time scale without limitations to the size or complexity of the system and in native-like conditions. In SDSL, a paramagnetic nitroxide side chain is introduced in a site-specific manner by mutating a selected residue to cysteine followed by modification of the reactive –SH group with a methanethiosulfonate reagent to generate a nitroxide side chain. The EPR spectrum of a spin-labeled protein encodes information on the overall motion of the nitroxide in the picosecond to nanosecond time scale, which in turn is modulated by the local environment around the nitroxide side chain. Motion of the nitroxide side chain on the picosecond to nanosecond time scale determines T_2 relaxation[5] and contributes directly to the EPR spectral lineshape. Hence, SDSL can be used to map sequence-dependent fast backbone motions of soluble and membrane-bound proteins. Although fast backbone fluctuations occurring on the nanosecond to picosecond time scale contribute directly to the EPR spectrum, exchange between conformational substates with lifetimes longer than ~100 ns is too

[5] A more detailed description of the T_2 relaxation process and its effect on the EPR spectral lineshape are described in detail in chapter 2 of this dissertation.

slow to produce magnetic relaxation, thus detection of such motions by SDSL requires different experimental strategies, one of which was developed as part of this dissertation (*see* Chapter 3).

Mapping fast and slow backbone fluctuations of ground-state conformations and low-lying excited states populated either at low pH or *via* mutagenesis is an area of particular interest in SDSL and the main motivation for this dissertation. In the next chapter (Chapter 2), the reader will be introduced to the principles of magnetic resonance and site-directed spin labeling. In Chapter 3, a new experimental strategy developed for detecting conformational substates in slow-exchange will be introduced. In the subsequent chapters 4-6, the newly developed experimental strategy is employed along with T_2-based lineshape analysis to (1) map the conformational properties of sperm whale myoglobin in 4 folded states in equilibrium conditions and (2) explore the relationship between stability and flexibility by monitoring changes in the backbone flexibility of multiple variants of T4 lysozyme bearing core re-packing destabilizing substitutions similar to the L99A mutant discussed in section 1.2.1. In chapter 7, a new methodology developed for site-specific immobilization of spin-labeled proteins will be introduced, which eliminates the undesired effect of protein tumbling on the EPR spectral lineshape of small proteins at X-band frequency, while at the same time enabling a flow system whereby the structural and dynamic effects of other molecules (*i.e.*, ligands, osmolytes, binding partners) can be assayed sequentially on the same spin-labeled protein sample.

1.6 References

1. Linderstrøm-Lang, K.U. and Shellman, J.A. 1959 *Enzyme* 1:443-510.

2. Koshland, D.E. 1958. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci.* 48:98-123.
3. Fisher, E. 1894 "Einfluss der Configuration auf die Wirkung der Enzyme". *Ber. Dt. Chem. Ges.* 27: 2985–93
4. Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181:662–6.
5. Perutz, M.F., Rossman, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C.T. 1960. Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5A resolution, obtained by x-ray analysis. *Nature.* 185:416–422.
6. Lakowicz, J.R. and Weber, G. 1973. Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins in the nanosecond time scale. *Biochemistry* 12:4171-4179.
7. Allerhand, A., Doddrell, D., Glushko, V., Cochrand, D., Wenkert, E., Lawson, P., and Gurd, F. 1971. Conformational and segmental motion of native and denatured ribonuclease A in solution. Application of naturally abundant carbon-13 partially relaxation Fourier transform nuclear magnetic resonance. *J. Am. Chem. Soc.* 93:544-546.
8. Hvdit, A. and Nelson, O. 1966. Hydrogen exchange in proteins. *Adv. Prot. Chem*. 21:287-386.
9. Weber, G. 1975. Energetics of ligand binding to proteins. *Adv. Prot. Chem*. 29:1-83.
10. Cooper, A. 1976. Thermodynamic fluctuations in protein molecules. *Proc. Natl. Acad. Sci.* 73:2740-2741.

11. Antonini, E. and Brunori, M. 1971. Hemoglobin and myoglobin and their rection with ligands. Amsterdam, North-Holland publishing company.
12. Austin, R.H., Beeson, K.W., Einstein, L., Frauenfelder, H., and Gunsalus, I.C. 1975. Dynamics of ligand binding to myoglobin. *Biochemistry* 14:5355-5373.
13. Schlitchting, I., Berendzen, J., Phillips, G.N., Sweet, R.M. 1994. Crystal structure of photolyzed myoglobin. *Nature* 371:808-812.
14. Srajer, V. et al., 1996. Photolysis of the carbon monoxide complex of myoglobin:nanosecond time-resolved crystallography. *Science* 274:1726-1729.
15. Tomita, A., Sato, T., Ichiyanagi, K., Nozawa, H., Chollet, A., Kawai, F., Park, S,Y., Tsuduki, T., Yamato, T., Koshihara, S., and Adacji, S. 2009. Visualizing breathing motion of internal cavities in concert with ligand migration in myoglobin. *Proc. Natl. Acad. Sci.* 106:2612-2616.
16. Elbert, R and Karplus, M. 1987 Multiple conformational states of proteins:a molecular dynamics analysis of myoglobin. *Science* 235:318-321.
17. Tilton, R.F., Kuntz, I.D., and Petsko, G.A. 1984. Cavities in proteins:Sutructure of a metmyoglobin xenon complex solved to 1.9Å. *Biochemistry* 23:2849-2857.
18. Frauenfeder, H. Petsko, G.A., and Tsernoglou, D. 1979 Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* 280:558-563.
19. Frauenfelder, H. 2002. Proteins:paradigms of complexity. *Proc. Natl. Acad. Sci.* 99:2479-2480.
20. Frauenfelder, H., McMahon, B.J., Austin, R.H., Chu, K., and Groves, J.T. 2001. The role of structure, energy landscape, dynamics, and allostery in the enzyme function of myoglobin. *Proc. Natl. Acad. Sci.* 98:2370-2374.

21. Kumar, S., Ma, B., Tsai, C.J., Sinha, N., and Nussinov, R. 2000. Folding and binding cascades: Dynamic landscapes and population shifts. *Prot. Sci.* 9:10-19.

22. Frauenfelder, H., McMahon, B.H., and Fenimore, P.W. 2003. Myoglobin:The hydrogen atom of biology and a paradigm of complexity. *Proc. Natl. Acad. Sci.* 100:8615-8617.

23. Tomita, A., Kreutzer U., Adachi, S., Koshihara, S., and Jue, T. 2010. "It's hollow":the function of pores within myoglobin. *J. Exp. Biol*. 213:2748-2754.

24. Tsai, C.J., Kumar, S., Ma, B., and Nussinov, R. 1999. Folding funnels, binding funnels, and protein function. *Prot. Sci.* 8:1181-1190.

25. Goh, C., Milburn, D., Gerstein, M. 2004. Conformational changes associated with protein-protein interactions. *Curr. Opin. Struc. Biol*. 14:1-6.

26. Holmes, K.C., Angert, I., Kull, F.J., Jahn, W., Schroder. R.R. 2003. Electroncryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature* 425:423-427.

27. James, L.C., Roversi, P., and Tawfic, D.S. 2003. Antibody multispecificity mediated by conformational diversity. *Science* 299:1362-1367.

28. Hendzer-Wildman, K.A., Ming, L., Thai, V., Kerns, S.J., Karplus, M., and Kern D. 2007. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450:913-918.

29. Kimple, R.J., Kimple, M.E., Betts, L., Sondek, J., Siderovski, D.P. 2002. Structural determinants of Go-Loco induced inhibition of nucleotide release by Galpha subunits. *Nature* 416:878-881.

30. Edvokimov, A.G., Phan, J., Tropea, G.E., Waugh, D.S. 2003 Similar modes of polypeptide recognition by export chaperones in flagellar biosysthesis and type III secretion. *Nat. Struc. Biol*. 10:789-793.

31. Masterson, L.R., Shi, L., Metcalfe, E., Gao, J., Taylor, S.S., and Veglia, G. 2011. Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. *Proc. Natl. Acad. Sci.* 108:6969-6974.

32. Volkman, B.F., Lipsom, D., Wemmer, D.E., Kern, D. 2001. Two-state allosteric behavior of a single domain signaling protein. *Science* 291:2429-2433.

33. Bakan, A. and Bahar, I. 2009. The intrinsic dynamics of enzymes play a dominant role in determining the structural changes induced upon inhibitor binding. *Proc. Natl. Acad. Sci.* 106:14349-14354.

34. Hammes, G.G., Chang, Y., and Oas T.G. 2009. Conformational selection or induced-fit:A flux description of reaction mechanism. *Proc. Natl. Acad. Sci.* 106 13737-13741.

35. Boehr, D.D. and Wright, P.E. 2008. How do proteins interact? *Science* 320:1429-1430.

36. Stivers, J.T., Abeygunawardana, C., Mildvan, A.S., and Whitman, C.P. 1996. ^{15}N relaxation studies of free and inhibitor-bound 4 oxalocrotonate tautamerase:backbone dynamics and entropy changes of an enzyme upon inhibitor binding. *Biochemistry* 35:16036-16047.

37. Hodsdon, M.E. and Cistola, D.P. 1997. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N relaxation and ^{1}H exchange. *Biochemistry* 36:2278-2290.

38. Freedberg, D.I., Ishima, R., Jacob, J., Wang, Y., Kustanovich, I., Louis, J.M., Schnell, J., and Torchita, D.A. 2002 Rapid structural fluctuations of the HIV protease flaps in solution. Relationship to crystal structures and comparisons with predictions of dynamics calculations. *Prot. Sci.* 11:221-232.
39. Adams, P.D., Loh, A.P., and Oswald, R.E. 2004. Backbone dynamics of an oncogenic mutant of cdc42Hs shows increased flexibility at the nucleotide binding site. *Biochemistry* 43:9968-9977.
40. Hodsdon, M.E., and Cistola, D.P. 1997 Discrete backbone disorder in the nuclear magnetic resonance structure of the apo intestinal fatty acid-binding protein:implications for the mechanism of ligand entry. *Biochemistry* 36:1450-1460.
41. Bhaba, G., Lee, J., Ekiert, D.C., Gam., J., Dyson, H.J., Benkovic, S.J., and Wright, P.E. 2011. A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science* 332:234-237.
42. Woodside, D.G. 2002. Dancing with multiple partners. Sci STKE 2002:PE14
43. Ma, B.,Shatsky, M., Wolfson, H.J., and Nussinov, R. 2002. Multiple diverse ligand binding at a single protein site: a matter of pre-existing populations. *Prot. Sci.* 11:184-197.
44. Yuan, Z., Zhao, J., and Wang, X. 2003. Flexibility analysis of enzyme active sites by crystallographic temperature factors. *Protein Eng*. 16:109-114.
45. Radivojac, P., Obradovic, Z., Smith, D.K., Zhu, G., Vucetic, S., Brown, C.J., Lawson, J.D., and Dunker, A.K. 2004. Protein flexibility and intrinsic disorder. *Prot. Sci.* 13:71-80.

46. Gerstein, M. and Krebs, W. 1998. A database of macromolecular motions. *Nucl. Acid. Res*. 26:4280-4290.

47. Weiss, S. 2000. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat. Struc. Mol. Biol*. 7:724-729.

48. Wales, T.E., and Engen, J.R. 2006. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spec. Rev*. 25:158-170

49. Thomas, G.J. 1999. Raman spectroscopy of protein and nucleic acid assemblies. *Annu. Rev. Biophys. Biomol. Struc*. 28:1-27

50. Zanni, M.T. and Hochstrasser, R.M. 2001. Two-dimensional-infrared spectroscopy (2D-IR): a promising new method for the time resolution of structure. *Curr. Opin. Struc. Biol.* 11:516-522.

51. Mittermaier, A, and Kay. L.E. 2006. New tools provide new insights in NMR studies of protein dynamics. *Science* 312:224-228.

52. Boehr, D.D., Dyson, J.H., and Wright, P.E. 2006. AN NMR perspective on enzyme dynamics. *Chem Rev*. 106:3055-3079.

53. Lipari, G. and Szabo, A. 1982(1) Model-free approach to the interpretation of nuclear magnetic resonance in relaxation of macromolecules. Theory and range of validity. *J. Am. Chem. Soc.* 104:4546-4559.

54. Lipari, G. and Szabo, A. 1982(2) Model-free approach to the interpretation of nuclear magnetic resonance in relaxation of macromolecules. Analysis of experimental results. *J. Am. Chem. Soc.* 104:4559-4570.

55. Mulder, F.A.A., Mittermaier, A., Hon, B., Dahlquist F.W., and Kay, L.E. 2001. Studying excited states of proteins by NMR spectroscopy. *Nat. Struc. Biol*. 8:932-935.

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

57. Baase, W.A., Liu, L., Tronrud, D.E., and Matthews, B.W. 2010. Lessons from the lysozyme of phage T4. *Prot. Sci.* 19:631-641.

58. Eriksson, A.E., Baase, W.A., Wozniak, J.A., and Matthews, B.W. 1992. A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* 355:371-373.

59. Morton, A., Baase, W.A., and Matthews, B.W. 1995(1). Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry* 34:8564-8575.

60. Morton, A. and Matthews, B.W. 1995(2). Specificity of ligand binding in a buried nonpolar cavity in T4 lysozyme. *Biochemistry* 34:8576-8588.

61. Mulder, F.A.A., Hon, B., Muhandiram, D.R., Dahlquist, F.W., and Kay, L.E. 2000. Flexibility and ligand exchange in a buried cavity mutant of T4 lysozyme studied by multinuclear NMR. *Biochemistry* 39:12614-12622.

62. Feher, V.A., Baldwin, E.P., and Dahlquist, F.W. 1996. Access of ligands to cavities within the core of a protein is rapid. *Nat. Struc. Biol*. 3:516-521.

63. Bouvignies, G., Vallurupalli, P., Hansen, D.F., Correia, B.E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D., and Kay, L.E. 2011. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature*, 477:111-114.

64. Frauenfelder, H., Alberding, N.A., Ansari, A., Braunstein, E., Cowen, B.R., et al. 1990. Proteins and pressure. *J. Phys. Chem* 94:1024-1037.

65. Collins, M,D., Hummer, G., Quillin, M.L., Matthews, B.W., and Gruner, S.M. 2005. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. *Proc. Natl. Acad. Sci.* 102:16668-16671.

66. Li, H. and Akasaka, K. 2006. Conformational fluctuations of proteins revealed by variable pressure NMR. *Bioch. Biophy. Acta* 1764:331-345.

67. Akasaka, K. 2006. Probing conformational fluctuation of proteins by pressure perturbation. *Chem Rev.* 106:1814–1835.

68. Akasaka, K. and Yamada, H. 2001. On-line cell high pressure nuclear magnetic resonance technique. Applications to proteins studies. *Meth. Enzymol.* 338:134-158.

69. Kitahara, R., Yamada, H., Akasaka, K., and Wright, P.E. 2002. High pressure NMR reveals that apomyoglobin is in an equilibrium mixture from the native to the unfolded. *J. Mol. Biol*. 320:311-319.

70. Eliezer, D., Jennings, P.A., Dyson. H.J., and Wright. P.E. 1997. Populating the equilibrium molten globule state of apomyoglobin under conditions suitable for characterization by NMR. *FEBS Letters* 417:92-96.

71. Jennings, P.A. and Wright, P.E. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* 262:892-896.

72. Eliezer, D., Jennings, P.A., and Wright, P.E. 1995. The radius of gyration of an apomyoglobin folding intermediate. *Science* 270:487-488.

73. Eliezer, D., Yao, J., Dyson, J., and Wright, P.E. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struc. Biol*. 5:148-155.

74. Eliezer, D., Chung, H., Dyson, H.J., and Wright, P.E. 2000. Native and non-native secondary structure and dynamics of the pH 4 intermediate of apomyoglobin. *Biochemistry* 39:2894-2901.

75. Yao, J., Chung, J., Eliezer, D., Wright, P.E., and Dyson, H.J. 2001. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry* 40:3561-3571.

76. Bolon, D.N., Voigt, C.A., and Mayo, S.L. 2002. *De novo* design of biocatalyst. *Curr. Opin. Chem. Biol*. 6:125-129.

77. Brustad, E.M. and Arnold, F.H. 2011. Optimizing non-natural protein function with directed evolution. *Curr. Opin. Chem. Biol.* 15:201-210.

78. Hilvert, D. 2000. Critical analysis of antibody catalysis. *Annu. Rev. Biochem*. 69:751-793.

79. Röthlisberger, D., Khersonsky, O., Wollacott, A.M., Jiang, L., DeChancie, E., Betker, J., Gallaher, J.L., et al., Tawfik, D.S., and Baker, D. 2008. Kemp elimination catalysis by computational enzyme design. *Nature* 453:190-197.

80. Walsh, S. T. R., Lee, A. L., DeGrado, W. F., and Wand J.A. 2001. Dynamics of a De Novo Designed Three-Helix Bundle Protein Studied by ^{15}N, ^{13}C, and ^{2}H NMR Relaxation Methods. *Biochemistry 40*:9560–9569.

81. Dobson, N., Dantas, G,. Baker, D., and Varani, G. 2006. High-resolution structural validation of the computational redesigned human U1A protein. *Structure* 14:847-856.

82. Crowhurst, K.A. and Mayo, S.L. 2008. NMR-detected conformational exchange observed in a computationally design variant of Gβ1. *Prot. Eng. Des. Sel.* 21:577-587.

Chapter 2: Introduction to Electron Paramagnetic Resonance (EPR) Spectroscopy and Site-Directed Spin Labeling (SDSL)

2.1 Synopsis

EPR spectroscopy is a technique that can be employed to study species containing unpaired electrons. Most ground-state molecules have all the electrons paired; however there are two main types of molecules that contain unpaired electrons which are of significance in biological systems, namely, transition metal ions (*i.e.,* Mn, Fe, Cu, Co) and free radicals. Since the discovery of the EPR phenomenon by Yevgeny K. Zavoisky in 1944, EPR spectroscopy has become an important tool in chemistry. The application of EPR spectroscopy to biological molecules can be traced back to the 1950's [1], but it wasn't until 1965 that extrinsic synthetic probes, known as spin labels, were used as paramagnetic sensors in otherwise "EPR-silent" proteins [2]. The development of techniques in molecular biology, such as site-directed mutagenesis in the 1980's, paved the way for the application of EPR spectroscopy to recombinant proteins *via* introduction of spin labels as probes in a site-specific manner [3], thereby expanding the experimental tools that can be employed to study the molecular structure and dynamics of proteins.

In this chapter some of the basic fundamentals of EPR spectroscopy will be introduced with a particular emphasis on its application to spin-labeled proteins. Other applications of EPR spectroscopy (*i.e.*, study of metalloproteins, spin-trapping, EPR oximetry) will not be discussed, but readers interested in such applications are encouraged to review the following references [4-6]. The aim of this chapter is to provide a general audience sufficient background on EPR spectroscopy to understand the data and interpretations presented in the subsequent chapters,

thus mathematical treatment of the EPR phenomenon is kept to a minimum. A more quantitative description of the basic principles of magnetic resonance and EPR spectroscopy can be found in the following references: Carrington and McLachlan [7], Marsh [8], and Poole [9].

2.2 Introduction to Spectroscopy

Generally speaking, in spectroscopy one studies the interaction between matter with electromagnetic radiation by measuring the absorption, emission, and scattering of light by matter. The interactions between light and matter can cause transitions between quantum-mechanically allowed energy levels. For example, when a molecule in state α (let α be the ground state) interacts with electromagnetic radiation (*see* left panel in Figure 2.1), it may absorb a photon of frequency ν and transition to higher energy state β if the frequency satisfies $E_\beta - E_\alpha = h\nu$, where h is the Planck constant (6.626×10^{-34} J·s). In the absence of radiation, a molecule can decay from a higher energy state to a lower energy state, thereby emitting a photon with a frequency that satisfies $E_\beta - E_\alpha = h\nu$ (spontaneous emission). Exposure to electromagnetic radiation can also cause transition from β to α (right panel on Figure 2.1), with resulting emission of frequency ν and energy of hν (stimulated emission).

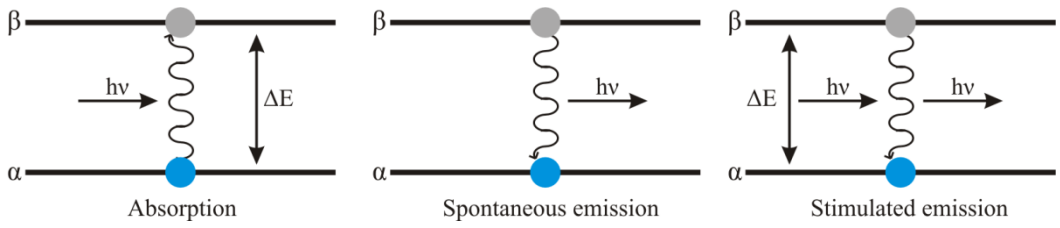


Figure 2.1: Transitions associated with absorption and emission of electromagnetic radiation

The field of spectroscopy has been subdivided into several branches depending on the energy involved in the quantum transitions. For example, molecular vibrations require energy of absorption in the infrared regions of the electromagnetic spectrum ($\nu \sim 10$ to 3×10^{14} Hz), thus

the experimental technique that is used to study such phenomena is known as Infrared (IR) Spectroscopy. In EPR spectroscopy, one observes the transitions between quantum mechanical energy levels of an unpaired electron when an external magnetic field (**H**) is applied; such transitions require frequencies in the microwave region (GHz) of the electromagnetic spectrum.

2.3 Magnetic Resonance

Electrons have an intrinsic angular momentum (known as spin) that can be explained in a classical way as arising due to the electron "spinning' about its own axis[1]. This intrinsic angular momentum is in addition to the orbital angular momentum that arises due to the motions of the elementary particle through space. Because an electron has a magnetic moment (denoted μ_e), it acts like a bar magnet when placed in an external magnetic field (**H**). Experimental evidence has shown that electrons have a spin (**S**) of ½, therefore $m_s = \pm ½$ (m_s is the magnetic component that can take values of -s, s+1,…,s). Like a bar magnet, when an electron is subjected to an external magnetic field, the magnetic dipole orients with respect to the direction of the applied magnetic field (Zeeman effect). A single unpaired electron has two allowed energy states ($m_s = \pm ½$) as depicted in Figure 2.2, with a minimum energy orientation when μ_e is aligned with the magnetic field (designated – ½) and a maximum of energy orientation when μ_e is aligned opposite to the magnetic field (+ ½). Because of the negative charge of the electron, the energies of the ± ½ states are opposite from those of the protons; that is, for a proton

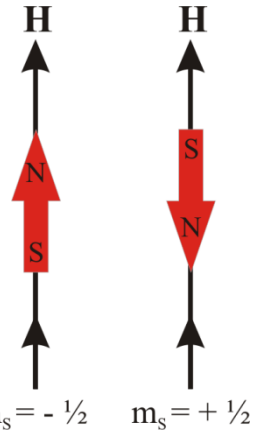


Figure 2.2: Allowed energy states of the magnetic moment of the electron with respect to the applied magnetic field (**H**).

[1] Care should be taken in explaining the spin in classical way, since spin is a purely quantum mechanical phenomenon.

the parallel orientation corresponds to + ½. The difference in energy between the two states can be given by the following sets of equation:

$$E = -\mu_e \cdot H \tag{2.1}$$

$$\mu_e = - g\beta m_s, \tag{2.2}$$

where β is the Bohr magneton = 9.274×10^{-24} J/Tesla, **H** is the strength of the applied magnetic field in Tesla (1T = 10^4 Gauss), $m_s = \pm ½$, and g is known as the g-factor, which for a free electron is 2.0023. Combining equations 2.1 and 2.2 yields one of the most fundamental equations in EPR spectroscopy, which describes the basic *resonance condition*:

$$\Delta E = h\nu = + ½\, g\beta\mathbf{H} - (- ½\, g\beta\mathbf{H}) = g\beta\mathbf{H}, \tag{2.3}$$

where h is the Plank constant = 6.626×10^{-34} Js .From equation 2.3 it should be noted that in the absence of an applied magnetic field (H ≅ 0 Tesla[2]), all the orientations of the spin have the same energy (Δ E = 0) and the energy of the 2 spin states increases linearly with the magnetic field strength as depicted in Figure 2.3. The resonance condition can be obtained either by varying the magnetic field with a fixed frequency or by applying a constant field while scanning the frequency. The frequency associated with the resonant absorption is often expressed in terms of angular frequency (ω) and is written in the general form

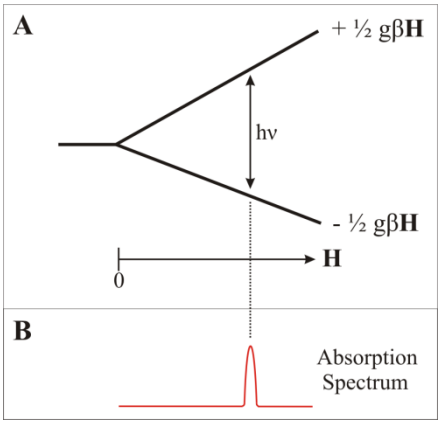


Figure 2.3 A) Divergence of the energies of the two spin states as a function of H. B) Absorption spectrum of the transition depicted in panel A.

$$\omega = \gamma H, \tag{2.4}$$

[2] The strength of the Earth's magnetic field varies between 20-60μT, depending on the location, and has declined about 10% over the last 100 years [10].

where γ is the electron gyromagnetic ratio. Due to technical limitations, in EPR spectroscopy the frequency is held constant and the magnetic field is scanned. The *field of resonance* is achieved when the energy of the radiation applied matches the energy difference between the $E_{+½}$ and $E_{-½}$ states (in some textbook these states are denoted $E_{\downarrow}$ and $E_{\uparrow}$, respectively). Table 2.1 lists the field of resonance for different microwave frequencies most commonly used in EPR spectroscopy.

Table 2.1 Field of resonance at different frequencies calculated with equation 2.3.

Frequency (GHz)	Microwave band	Field of resonance (G)
1.1	L	392
4.0	S	1439
9.75	X	3507
24.0	K	8633
34.0	Q	12230
94.0	W	33811

2.4 Boltzmann distribution and relaxation effects

In EPR spectroscopy one observes a net absorption of microwave frequencies by unpaired electrons because the two possible energy levels of the electron (↑ and ↓) are not equally populated; that is there are a greater number of electrons aligned parallel to the applied magnetic field than there are opposed to it. The extent of absorption, which defines the signal intensity, can be described in accord to the Boltzmann distribution law:

$$\left(\frac{N_{\downarrow}}{N_{\uparrow}}\right) = e^{-\left(\Delta E/kT\right)} = e^{-\left(g\beta \boldsymbol{H}/kT\right)}, \tag{2.5}$$

where T is the temperature (in Kelvin) and *k* is the Boltzmann's constant. The difference in populations between the two energy states is actually very small. For example, at X-band frequency (~ 9.75 GHz), which corresponds to a resonant field of ~ 0.35 Tesla (*see* table 2.1; ~3500 G) and at room temperature (298 K), $\frac{N_{\downarrow}}{N_{\uparrow}} = 0.998$, thus for a sample with 10,000 spins, an average of 5005 spins will be oriented parallel to the magnetic field and 4995 will be oriented antiparallel to the magnetic field. That represents a net difference of 10 spins (0.1% of the total number of spins). The population difference can, in principle, be increased by increasing the

frequency of radiation and hence the field or by decreasing the temperature, thereby enhancing the sensitivity, however other complications may arise (*i.e.*, line-broadening, saturation) when the sample is cooled to low temperatures.

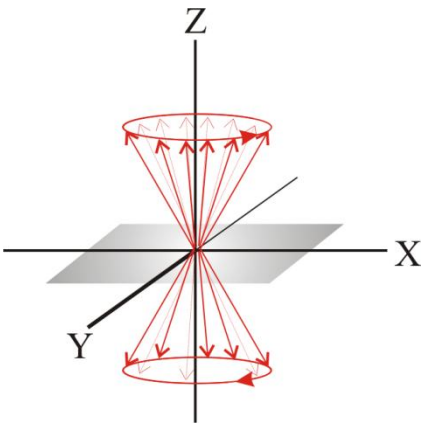


Figure 2.4 Orientations of the magnetic moment of the electron with respect to the applied magnetic field (along the z-axis) at equilibrium. The shaded region represents the *xy* plane.

The applied magnetic field (**H**) is designated to be parallel to the z axis in the cartesian coordinates system as shown in Figure 2.4. Since there is a greater number, albeit small, of spins oriented parallel to **H**, there is a net **magnetization** (**M**) along the *z*-axis under equilibrium conditions (magnetization is defined as the net sum of all the electronic magnetic moments in the sample). As shown in Figure 2.4, the electrons precessing about the *z*-axis at the *larmor frequency* (*see* arrows in the ovals) have random components in the *xy* plane, so that the individual magnetic components in the *x* and *y* directions (*transverse components*) cancel each other out, thus M_x and $M_y = 0$ at equilibrium.

When a matching electromagnetic radiation is applied, the system is no longer in thermal equilibrium and the populations in the upper and lower energy levels can become equalized, thus $M_z = 0$ (this phenomenon is called *saturation*) unless there is an efficient mechanism that restores the thermodynamic equilibrium. The mechanism that restores the system to thermodynamic equilibrium is known as *relaxation*. The rate at which M_z moves back to its thermodynamic equilibrium is proportional to $1/T_1$, where T_1 is known as the *spin-lattice relaxation time*. This type of relaxation occurs because the coupling between the electron spins and their surroundings (also known as bath or lattice because it was first studied in crystalline solids) effectively transfers the energy from the spins in the form of heat to the lattice, thereby

"heating up" the system. T_1 has been defined as the time it takes M_z for to return to ~ 63 % of its initial value and depends on how strong the spins are "coupled"[3] to the environment. For example, experiments have shown that paramagnetic ions (*i.e.,* Fe^{+3}) are strongly coupled to the lattice, therefore T_1 is short at room temperature, whereas for free radicals, including spin labels, the coupling between the spin and the lattice is weak , thus T_1 is long (μs time scale). This has important implications in the detection of the EPR signal as will be discussed in section 2.5.1.

As mentioned above, at equilibrium conditions the transverse magnetization ($\mathbf{M}_x$ and $\mathbf{M}_y$) is zero, however when microwave radiation at the resonance frequency is applied, the magnetization in the *xy* plane is no longer zero (*see* Figure 2.5). The net magnetization in the *xy* plane will precess about the *z*-axis at the larmor frequency, however random fluctuations in the local magnetic field around each spin cause dephasing along the *xy* plane (that is, some spins will precess faster than others) until the net $\mathbf{M}_{xy}$ transverse magnetization returns to its equilibrium value of zero. The rate at which $\mathbf{M}_{xy}$ returns to zero is proportional to $1/T_2$, where T_2 is defined as the *spin-spin relaxation time*. In contrast to spin-lattice relaxation, spin-spin relaxation is a non dissipative process (no energy is transferred to the lattice) and it does not change the total number of spin in the $N_\uparrow$ and $N_\downarrow$ states, thereby it cannot alleviate saturation.

[3]Coupled to the lattice means that there must be enough lattice thermal energy (kT) in the form of vibrations of the same magnitude as $g\beta\mathbf{H}$. **Thus T_1 is temperature-dependent.**

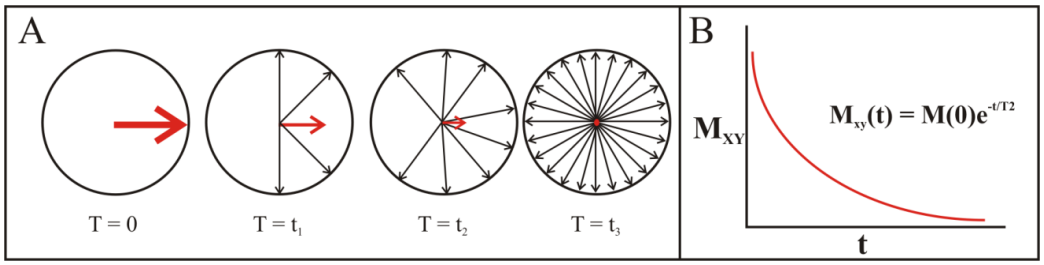


Figure 2.5 Spin-spin relaxation. A) Change in the magnetization along the *xy* plane after an external oscillating magnetic field satisfying the resonance condition is applied (the viewer is looking down the z-axis). The red arrow represents the net magnetization and the black arrows represent individual spins. B) The change in the M_{xy} magnetization is described by an exponential decay curve.

2.4.1 Bloch equations. It is clear from the previous section that magnetic resonance does not focus on individual spins, but rather with the ensemble of spins that give rise to the bulk magnetization (**M**). Felix Bloch realized that the changes in the net magnetization under the influence of electromagnetic radiation can be described *classically* with a set of equations that takes into account the effects of the applied magnetic field and relaxation [11]. The first equation describes the torque "experienced" by **M** when a magnetic field is applied:

$$\frac{dM}{dt} = \gamma M \times \boldsymbol{H}, \tag{2.6}$$

where $\gamma = -g\beta/\hbar$ and $M = \Sigma\mu_i$ (the total magnetic moment of all the individual spins). When the contributions from T_1 and T_2 relaxations are added, the time change of M on each of the axis (*x*,*y*, and *z*) takes the following form:

$$\frac{dM_z}{dt} = -\frac{(M_z - M_0)}{T_1} \tag{2.7}$$

$$\frac{dM_x}{dt} = -\frac{(M_x - M_0)}{T_2} \tag{2.8}$$

$$\frac{dM_y}{dt} = -\frac{(M_y - M_0)}{T_2} \tag{2.9}$$

since at equilibrium conditions M*x* and M*y* = 0, equations 2.8 and 2.9 take the following form:

$$\frac{dM_x}{dt} = -\frac{M_x}{T_2} \text{ and } \frac{dM_y}{dt} = -\frac{M_y}{T_2}$$

Combining equations 2.6 – 2.9 to account for the torque exerted by the field and relaxation effects on the net magnetization yields the **Bloch Equation**:

$$\frac{dM}{dt} = \gamma M \times \boldsymbol{H} - \frac{k(M_z - M_0)}{T_1} - \frac{iM_x}{T_2} - \frac{jM_y}{T_2}, \tag{2.10}$$

where *k*, *i*, and *j* are unit vectors in the *z*, *x*, and *y* directions, respectively. Equation 2.10 describes the "evolution" of the magnetization in the *laboratory* frame of reference, which is a stationary coordinate system, however the Bloch equation is usually solved using a *rotating* frame. The transformations and derivations of the Bloch equations in the rotating frame are not discussed here, but interested readers are referred to Humpries and McConnell [12] and Cavanagh et al. [13].

2.5 Parameters that determine the EPR spectrum

2.5.1 T_1 and T_2 relaxation. The line width of the absorption spectrum is modulated by the relaxation mechanism, which is often defined in terms of spin-spin relaxation time

$$\frac{1}{T_2} = \frac{1}{T_2'} + \frac{1}{2T_1}, \tag{2.11}$$

where $(1/T_2')$ arises from fluctuations of the internal field along the z-axis.

The contribution of $1/(2T_1)$ can be rationalized in terms of the Heisenberg uncertainty principle. This principle implies that if a particular system exists in a certain energy level for a short period of time, then the energy of the state is not well-defined as depicted in Figure 2.6.

Under this condition, the energy range of the state is related to its lifetime (τ) and can be obtained with the following relation:

$$\Delta E * \Delta\tau \approx h, \tag{2.12}$$

since $\Delta E = h\nu$, then equation 2.12 can be rearranged to get

$$\Delta\nu \approx 1/\Delta\tau \tag{2.13}$$

In EPR spectroscopy, the frequency is kept fixed and the magnetic field is swept, therefore the line width of the resonance is calculated by combining equations 2.3 with equation 2.12 to get

$$g\beta\Delta H \approx h\Delta\nu, \tag{2.14}$$

solving for ΔH (line width of the resonance in Tesla) gives the following relation:

$$\Delta H = \left(\frac{h}{g\beta}\right) * \left(\frac{1}{\Delta\tau}\right) \tag{2.15}$$

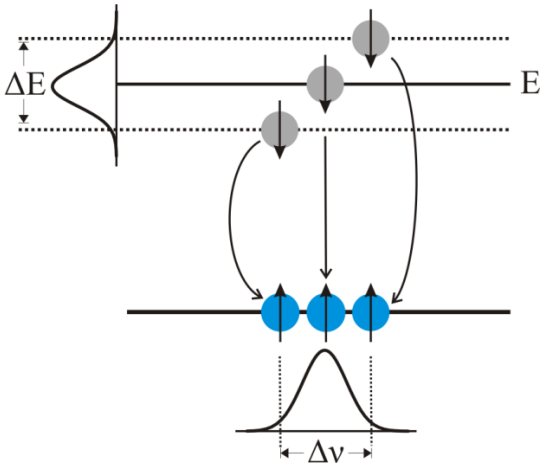


Figure 2.6 Origin of lifetime broadening. The energy of the exited state is not well-defined due to its short lifetime, therefore the resonant frequency and hence the field of resonance are broadened.

The broadening due to T_1 relaxation is known as *lifetime broadening* and applies to all the spins in the system (homogeneous). For paramagnetic ions (*i.e.*, Fe^{3+}, Co^{2+}), the strong coupling of the spin to the lattice at room temperature results in short T_1, thereby the resonances for such species are broad beyond detection[4] and can only be observed by decreasing the temperatures (to ~ 4 K) such that T_1 is increased significantly. Conversely, for free radicals (including

[4] This is an important detail that should be taken into account when reading Chapter 4 of this dissertation.

spin labels attached to proteins) the coupling of the spin to the lattice is weak (T_1 is long) and T_2 << T_1, therefore the line broadening in such cases arises from T_2 relaxation. Solution of the Bloch equations in the rotating coordinate system[5], shows that under conditions in which the microwave power is low enough, such that there is no saturation, the average power absorbed by the radical is

$$P(\omega) = \frac{\omega\gamma H_1^2 M_0 T_2}{[1+T_2^2(\omega-\omega_0)^2+\gamma^2 H_1^2 T_1 T_2]} \tag{2.16}$$

Using equation 2.4 to express equation 2.16 in terms of field dependency yields

$$P(\omega) = \frac{H\gamma^2 H_1^2 M_0 T_2}{[1+T_2^2\gamma^2(H-H_0)^2+\gamma^2 H_1^2 T_1 T_2]} \tag{2.17}$$

The absorption curve described by equation 2.17 has a lorentzian lineshape as shown in Figure 2.7A. As will be discussed in section 2.6, for technical reasons the EPR spectra is recorded as the first derivative (depicted in Figure 2.7B) and for such spectra, the peak-to-peak separation in gauss (denoted ΔH_{pp}) is determined by T_2

$$\Delta H_{pp} = \frac{2}{\gamma\sqrt{3}T_2} = \frac{6.56*10^{-8}}{T_2} \tag{2.18}$$

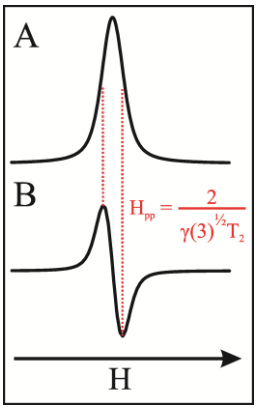


Figure 2.7 Lorentzian lineshape. A) Absorption curve obtained when the microwave power is below saturation levels. B) First-derivative of the absorption curve shown in A. The peak-to-peak separation (ΔH_{pp}) and the equation relating ΔH_{pp} to T_2 are shown.

[5] The complete solution of the Bloch equation can be found in Humphries and McConnell [12].

2.5.2 g-value. The g-factor is a proportionality constant that determines the position of the resonance. Experiments have shown that for a free electron the g value (denoted g_e) is 2.0023, however for a bound electron the g-value can deviate from g_e. As mentioned in section 2.3, an electron possesses an intrinsic spin angular momentum (**S**) and orbital angular momentum (denoted **L**), the latter being due to the motion of the particle through space. The motion of a bound electron through space is determined by the orbital (spatial wave function) of the electron, thus the magnitude of L (hence g) depends on the motion associated with the orbital. For example, the magnitude of L for a particular orbital can be calculated using the following relation:

$$|\mathrm{L}| = \hbar\sqrt{l(l+1)}, \tag{2.19}$$

where l is the angular momentum quantum number. For an electron in the s orbital $l = 0$, thus L = 0 (there is no angular momentum associated with it), however for electrons in the p ($l = 1$) and d ($l = 2$) orbital, $\mathrm{L} \neq 0$. Deviation from free spin value of 2.0023 gives information about the degree of orbital restriction of the unpaired electron. In a way, the g-value can be described as a measurement of the efficiency of the *spin-orbit* coupling. For free radicals, the unpaired electron is often delocalized, therefore its g-value is close to g_e. Conversely, the unpaired electron in transition metal ions and nitroxide spin labels is localized in the d and p orbitals, respectively, therefore those electrons exhibit g values that are different from g_e as shown in Table 2.2. Most importantly, the p and d orbitals are anisotropic (not spherically symmetric), consequently the g value associated with an electron in such orbitals

Table 2.2: Experimentally observed g values in paramagnetic species.

Paramagnetic species	g-value
Fe^{3+}	1.4 – 9.7
Cu^{2+}	2.0 – 2.4
Nitroxide spin labels	2.0020 – 2.0090

exhibits anisotropy[6], such that g will depend on the relative orientations of the electron orbital motion and the external magnetic field. As depicted in Figure 2.8, when the applied magnetic field and the spin are perpendicular to the orientation of the orbital moment L (denoted $L_{\perp}$), there is no contribution of L to the g-value. Conversely, when S and L are aligned with the applied magnetic field ($L_{\parallel}$), S and L couple and the magnitude of L contributes to the g-value ($g_{\parallel}$). Orientations other than the parallel and perpendicular also contribute to the exhibited g-factor, which can be computed using the following equation (Rousseau, 1984):

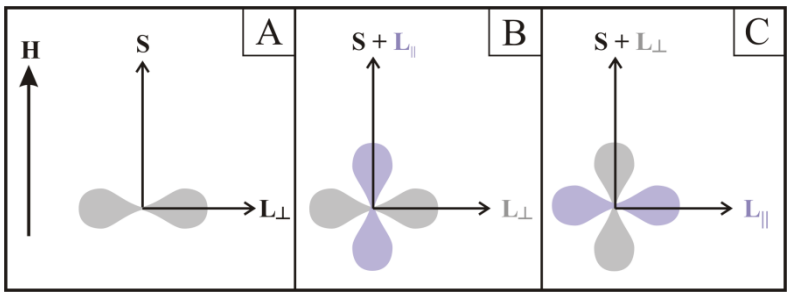


Figure 2.8 Effect of orientation between S and L on the magnitude of spin-orbit coupling. A) When S and L are perpendicular to each other, there is no spin-orbit coupling. B) Coupling between S and the parallel component of the orbital angular momentum. C) System has been rotated so S and $L_{\perp}$ are aligned and give rise to $g_{\perp}$.

$$g_{exhibited} = \sqrt{g_{\parallel}^2 cos^2\theta + g_{\perp}^2 sin^2\theta}, \tag{2.20}$$

where θ is the angle between $L_{\parallel}$ and the applied magnetic field. The principal g-values are often expressed using an axially symmetric coordinate system as represented in Figure 2.9, in which $g_{zz} \neq g_{xx} = g_{yy} = g_{\perp}$; $gzz = g_{\parallel}$. For small radicals tumbling fast in non-viscous solution, the g anisotropy is averaged out, thus the exhibited g value (g_{iso}) represents the average of the principal axis value and is given by the following equation:

$$g_{iso} = 1/3\ (g_{xx}+g_{yy}+g_{zz}) \tag{2.21}$$

in the case of axial symmetry

[6] Anisotropy means is orientation-dependent

$$g_{iso} = 1/3\ (g_{\|} + g_{\perp}) \qquad (2.22)$$

The time scale required to average out the g anisotropy can be calculated by assuming that the correlation time (τ) has to be shorter than the maximal variation in g (Δg):

$$\tau < h/(\Delta g \beta \mathbf{H}) \qquad (2.23)$$

For example, for a maximal variation in g of 0.0053 (as obtained for nitroxides at helical surface sites[7]) at **H** = 0.33T, the correlation time at which the nitroxide has to tumble to average out the anisotropy has to be shorter than 6.5 ns.

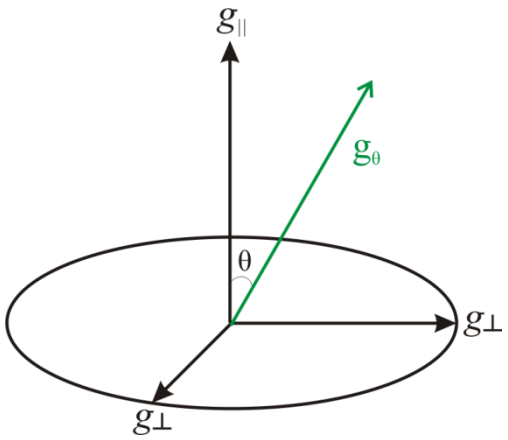


Figure 2.9 Diagrammatic representation of an axially symmetric system. θ is the angle between H and g||.

[7] Columbus et al., 2001 and Columbus and Hubbell, 2004

2.5.3 Hyperfine splitting. The field "experienced" by the electron spin is not only modified by the magnetic moment associated with the orbital (spin-orbit coupling), but it is also modulated by the magnetic moment of nearby nuclei with non-zero nuclear spin ($I \neq 0$). The interaction between the electron spin and nuclear spin is called *hyperfine interaction* and it leads to splitting of the EPR absorption peak into several lines (*hyperfine splitting*). Not all atoms that are of relevance in biological molecules have nuclei with non-zero magnetic moment, for instance, the abundant ^{12}C and ^{16}O atoms[8] have $I = 0$. However, many atoms that are abundant in biological molecules (*i.e.*, ^{1}H and ^{14}N) do possess non-zero nuclei spin, thus unpaired electrons orbiting such nuclei will experience a local field (H_{local}) from the nuclear magnetic moment that can enhance or counteract the applied magnetic field (**H**). As depicted in Figure 2.10, the magnetic moment of the nuclei acts like a bar magnet and produces a magnetic field that can orient parallel or antiparallel to the applied magnetic field, such that when H_{local} align parallel to the applied magnetic field, less magnetic field is required from the magnet to induce the transition between energy levels, therefore the resonance peak (indicated as $\mathbf{H_{S\uparrow I\uparrow}}$ in Figure 2.10) is lowered by a magnitude of H_{local}. Conversely, when the magnetic field of the nuclei aligns antiparallel to **H**, the field of resonance ($\mathbf{H_{S\uparrow I\downarrow}}$) is increased by a magnitude of H_{local}. The hyperfine interaction between the spin of the unpaired electron and

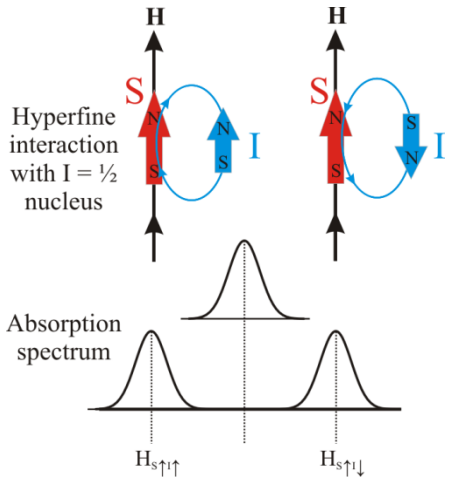


Figure 2.10 Hyperfine interaction. Effect of the orientation of the nuclear spin with respect to the applied magnetic field on the H_{loc} "sensed" by the electron spin. The shifts of the resonance field due to hyperfine interaction are shown in the lower panel.

[8] The natural abundance of ^{12}C and ^{16}O atoms is 98.9 % and 99.76%, respectively. ^{13}C has I = ½, but its natural abundance is 1.1% [10].

the nuclear spin contains both an isotropic component arising due to Fermi contact (*see* below) and anisotropic component arising due to through-space dipolar interaction between S and *I*.

$$H_{loc} = H_{iso} + H_{aniso} \quad (2.24)$$

*2.5.3.1 Isotropic component of the hyperfine interaction (**H_{iso}**).* The isotropic component of the hyperfine interaction arises from a phenomenon known as Fermi contact interaction (first described by Enrico Fermi in 1930[9]), which arises from spin-spin interaction between an electron inside the nucleus and the nucleus. Calculation of the spatial wave function of electrons ($|\psi|^2$) showed that the only orbital with spin density at the nucleus is the s orbital, therefore an unpaired electron density in a s orbital can give rise to Fermi contact interaction. For nitroxide spin labels, the unpaired electron resides mainly in the p orbital of the nitrogen nucleus, thus the spin density at the nucleus from the unpaired electron at the p-orbital (denoted $|\psi(0)|^2$) is zero, however non-zero $|\psi|^2(0)$ is observed experimentally by virtue of the fact that the hybrid orbital exhibit s character. The hyperfine interaction arising due to Fermi contact is **not** orientation-dependent (isotropic) and independent of molecular motion. The magnitude of the contribution from the isotropic hyperfine interaction can be calculated using the following expression (Humphries and McConnell [12])

$$H_{iso} = g\beta g_N \beta_N (\frac{8\pi}{3})|\psi(0)|^2 S \cdot I, \quad (2.25)$$

where g_N and β_N are the nuclear g-value and Bohr magneton, respectively. Equation 2.25 is often represented in the following form:

$$H_{iso} = ha\, S \cdot I, \quad (2.26)$$

[9] Fermi, E. (1930) *Z. Phys* 60: 320

where a is the isotropic hyperfine constant and is usually expressed in MHz (10^6 sec^{-1})

$$a = h^{-1} g \beta g_N \beta_N (\frac{8\pi}{3}) |\psi(0)|^2$$

(2.27)

The isotropic hyperfine component give rise to hyperfine splitting that can easily be predicted and interpreted. The nuclear spin (I) is quantitized and can only adopt certain orientations relative to the magnetic field direction corresponding to m_I = -I, -I+1,…, I–1, I (where m_I is the magnetic momentum of the nucleus). For example, in cases where I = 1, such as in the ^{14}N atom, the possible orientations of I are -1, 0, 1. The unpaired electron "senses" these orientations, thus the EPR resonance peak for such an electron splits into 2I + 1 lines of equal intensities. In cases where the unpaired electron senses the orientations of more than one equivalent nuclei with $I \neq 0$, the resonance peak splits into 2nI + 1 equally spaced hyperfine lines with relative intensities that are given by expansion of the expression $(1+x)^n$. For instance, in a methyl radical ($\bullet CH_3$), the unpaired electron will sense the orientations of 3 equivalent proton nuclear spin (I = ½), so

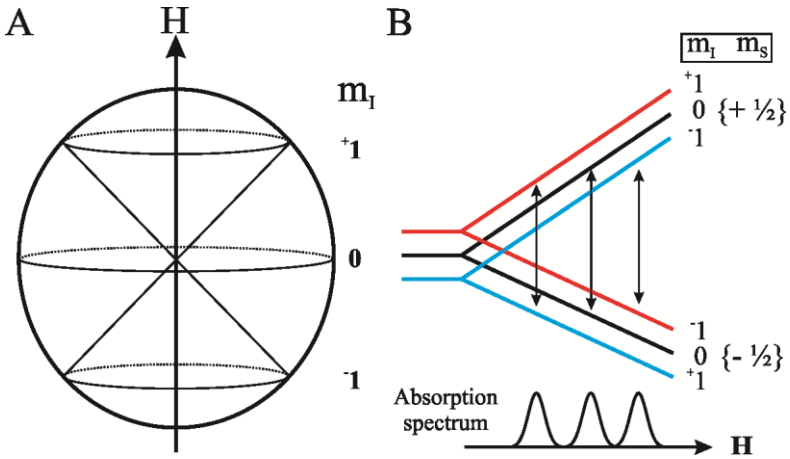


Figure 2.11 Hyperfine interaction in nitroxide spin labels give rise to three peaks in the spectra. A) Quantum mechanically allowed projection of the magnetic moment of the ^{14}N nucleus with respect to H. B) Transitions between energy levels of the electron spin. Selection rules for the allowed transitions: $\Delta m_I = 0$; $\Delta m_S = 1$.

that its resonance peak will split into 4 lines[10] (# lines = 2(3)(½) + 1). In the case of nitroxide spin labels, the unpaired electron senses the nuclear spin of the nearby ^{14}N atom (I = 1), therefore its absorption EPR spectrum consists of 3 peaks as shown in Figure 2.11.

2.5.3.2 Anisotropic component of the hyperfine interaction ($\boldsymbol{H_{aniso}}$). The anisotropic component of the hyperfine interaction arises from through-space[11] dipolar interaction between the magnetic moments associated with the electron and nucleus. In contrast to the Fermi interaction, the dipolar interaction between S and I is: (i) orientation-dependent (anisotropic), (ii) sensitive to molecular motion, and (iii) can have contribution from electron densities in the p, d, and f orbitals. The contribution from the anisotropic hyperfine interaction is described by the following relation (Humphries and McConnell, 1982)

$$H_{aniso} = \left[\frac{S \cdot I}{r^3} - \frac{3(S \cdot I)(I^r \cdot r)}{r^5}\right], \tag{2.28}$$

where r represents the distance between the unpaired electron and nucleus. For nitroxide radicals, such as the spin labels used in the studies discussed on this dissertation, the anisotropic contribution arises from spin density of the unpaired electron in the $2p_z$ orbital of the nitrogen nucleus (as depicted in Figure 2.12A), which is asymmetric as mentioned in section *2.5.2*. To study the magnitude of the orientation-dependence of the applied magnetic field with respect to the z-axis of the orbital in such spin systems, Griffith and coworkers [15], oriented single crystals containing a small amount of nitroxide free radicals with respect to **H**. Those studies revealed that when the applied magnetic field was along the principal axis of the nitroxide (z-axis), the EPR spectra exhibited the <u>maximum</u> splitting (designated $A_{\|}$) between the resonance peaks ($A_{\|}$ = 32 G); conversely, when H was rotated to a plane perpendicular to the principal axis

[10] The neighboring ^{12}C will not split the resonance peak because I = 0.
[11] This implies that the interaction is distance-dependent (*see* equation 2.28).

of the nitroxide, the observed hyperfine splitting (designated $A_{\perp}$) was the smallest (~ 6 G) as depicted in the simulated spectra shown in Figure 2.12B. Figure 2.12C shows the corresponding first derivative spectra of the nitroxide radicals oriented along the z, x, and y, axis, respectively. The absorption and first derivative spectra shown at the bottom of panels B and C of Figure 2.12, respectively, represent the sum of the spectra in all possible orientations in the absence of motion[12]. Notice that for such a spectrum, only the hyperfine extrema ($2A_{zz}$') are resolved. Intermediate orientations other than the parallel and perpendicular also contribute to the observed hyperfine splitting and their contributions can be computed using the following relation

$$A_{\theta} = \sqrt{A_{||}^2 cos^2\theta + A_{\perp}^2 sin^2\theta}, \qquad (2.29)$$

where θ is the angle between the applied magnetic field and the principal axis of the nitroxide. As mentioned in section *2.5.2*, the exhibited g values are also orientation-dependent, thus Griffith et al., 1965 showed that for nitroxides in oriented crystals $g_{xx} = 2.0089$[13], $g_{yy} = 2.0061$, and $g_{zz} = 2.0027$.

[12] This type of spectrum is often called **rigid glass** or **powder spectrum**.

[13] Notice that in the case of g-values, the largest shift in the position of the resonance line towards the lower field is observed when **H** is along the *x* axis of the nitroxide moiety (along the N→O bond).

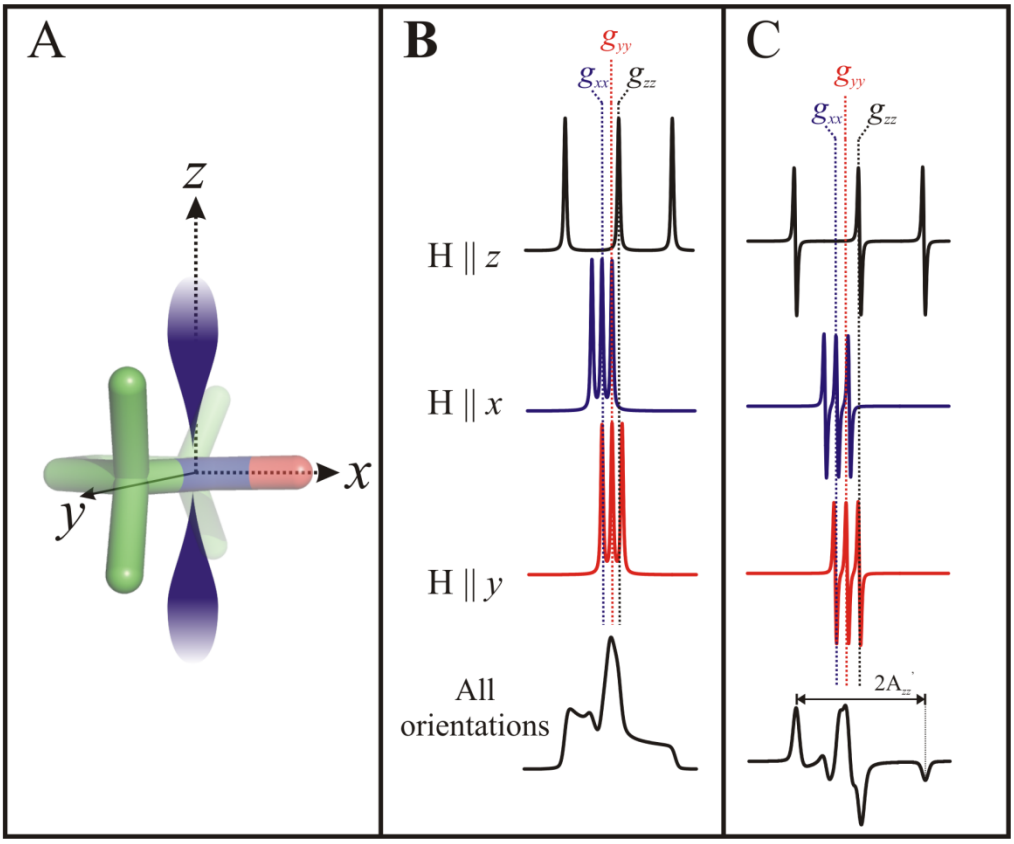


Figure 2.12 Spectral anisotropy of a nitroxide spin label oriented with respect to the applied magnetic field (**H**). A) Stick model of the nitroxide ring indicating the principal axis of the nitroxide radical. The unpaired electron resides primarily in the $2p_z$ orbital of the nitrogen nucleus (shaded region). B) Simulated X-band absorption spectra for a nitroxide radical from a single crystal with the principal axis oriented parallel and perpendicular with respect to the applied **H**. The A- and g- values used in the simulation were: $A_{\|}$= 32 G, $A_{\perp} = A_{xx} = A_{yy}$ = 6 G, g_{xx} = 2.0089, g_{yy} = 2.0060, g_{zz} = 2.0023. The dotted lines indicating the principal g-values are used to help guide the eye. C) Corresponding first-derivative spectra for the data shown in panel B. The powder spectra representing the sum of all the possible orientation are shown at the bottom of panels B and C.

In contrast to the H_{iso} component to the hyperfine interaction, the anisotropic component of the hyperfine interaction is sensitive to molecular motion, so rapid tumbling of free radicals (*i.e.*, a nitroxide spin label tumbling free in non-viscous solution) averages out the contribution from H_{aniso} to zero such that only H_{iso} contributes to the observed anisotropy[14]. In such cases, the

[14] Even for fast tumbling radicals, H_{aniso} contributes to the widths of the resonance lines because motional averaging is never perfect. For such cases, the largest line widths are observed for the high field lines, whereas the smallest line width is observed for the central line (Humphries and McConnell [12]).

observed hyperfine value is the mean of the principal axis value also known as the isotropic splitting constant (a_0) and is expressed as

$$a_0 = \frac{1}{3}\left(A_{xx} + A_{yy} + A_{zz}\right), \tag{2.30}$$

The time scale required to average out the A anisotropy can be calculated by assuming that the correlation time (τ) has to be shorter than the maximal variation in A (ΔA):

$$\tau < \frac{h}{2\pi g \beta \Delta A} \tag{2.31}$$

For example, for a maximal hyperfine anisotropy of ~31 G (as obtained for a nitroxide at surface sites in proteins[15]), the correlation time at which the nitroxide has to tumble to average out the A anisotropy has to be shorter than ~ 1.8 ns. Thus the 0.1 – 1.8 ns time domain is known as the *fast motional regime* in the EPR time scale. As shown above, the g-anisotropy averages out at correlation times shorter than ~ 6.5 ns, thus the A-anisotropy average out at a faster time scale. It should be emphasized that for free radicals bound to *well-ordered* proteins, the tumbling of the radical is slow enough such that the anisotropic contributions of the A and g tensors are significant to the resonant condition, thus the equation 2.3 can be re-expressed in the following form

$$\Delta E = h\nu = g_\theta \beta H + A_\theta m_I \tag{2.32}$$

It is this spectral anisotropy that gives structural and dynamical sensitivity to EPR spectroscopy!

[15] Columbus et al., 2001; Columbus and Hubbell, 2004.

2.5.3.3 Effect of local polarity on the hyperfine interaction. The magnitude of the hyperfine splitting can be influenced by factors that affect the electron spin density distribution (*i.e.*, polarity). Consider the electronic structure of a nitroxide radical shown in Figure 2.13. For such a radical, the unpaired electron is distributed between the p orbitals of the nitrogen and the oxygen atom at a relative distribution that is dependent on the dielectric constant of the solvent. Polar solvents with relatively high dielectric constant (*i.e.*, water) enhance the polar character of the N → O bond, which leads to an increase in the spin density distribution on the nitrogen 2p orbital (equilibrium shifts towards (a)) with a concomitant larger a_0 value. Conversely, non-polar solvents decrease a_0, shifting the equilibrium towards the (b) state. Thus the isotropic hyperfine splitting value can be used as a tool to probe the polarity of the local environment.

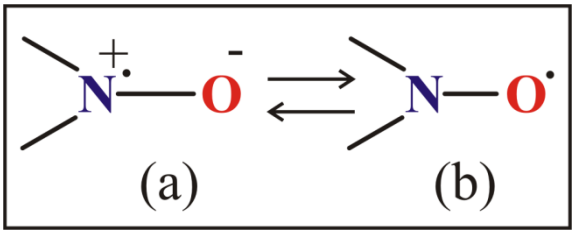


Figure 2.13 Effect of the polarity of the environment on the spin density distribution of the unpaired electron.

2.6 Basic Instrumentation

The aim of the previous sections was to introduce the basic principles of EPR spectroscopy. This section briefly describes the basic function of the various components of the EPR spectrometer. Similar to other forms of spectroscopy, the EPR spectrometer consists of three main components depicted in Figure 2.14, namely an electromagnetic radiation source, sample chamber, and a detector.

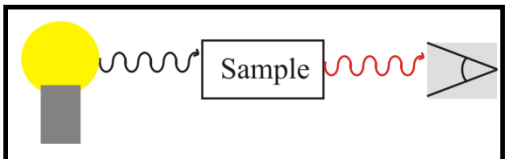


Figure 2.14 Basic components of common spectrometers.

However, there are two main differences between EPR spectroscopy and other forms of spectroscopy. First, due to the difficulty in tuning the frequency source in EPR spectroscopy, the electromagnetic radiation is kept fixed and the magnetic field is scanned. Secondly, most detection systems in other forms of spectroscopy (*i.e.*, optical spectroscopy) measure the amount of radiation that is not absorbed the sample, while the detector in the EPR spectrometer only "sees" the amount of radiation reflected back (*reflection spectrometer*) from the sample chamber. Measuring the amount of radiation getting through the sample is not practical for EPR spectroscopy due the weak signals arising from the small difference in the population between the two energy levels as discussed in section 2.4 and due to the fact that the magnetic dipole transitions have a smaller transition moment than electric dipole transitions detected in other forms of spectroscopy.

Figure 2.15 shows the main components of the EPR spectrometer. The source of the electromagnetic radiation and the detector are located in the rectangular box known as the microwave bridge. The electromagnetic radiation source can be either a klystron vacuum tube, which is an electron tube used to generate or amplify radiation in the microwave region by velocity modulation or a Gunn diode. The microwaves are transmitted through a rectangular tube known as a waveguide (highlighted in Figure 2.15), whose dimensions are in the order of the microwave wavelength (3 cm for X-band frequency). The sample chamber is a resonant cavity, which is a metal box that stores the microwave energy such that at its resonance frequency no signal will be reflected back to the detector. The cavities are characterized by their quality factor denoted as Q, which indicates how efficiently the cavity stores the microwave energy such that as Q increases, the sensitivity of the spectrometer increases. The Q-factor is defined as:

$$Q = \frac{2\pi\,(energy\ stored)}{energy\ dissipated\ to\ the\ side\ of\ the\ walls} \tag{2.33}$$

During resonance, there will be a "standing" wave inside the cavity with the electric and magnetic component exactly out of phase. Since it is the magnetic component of the electromagnetic radiation that drives the electronic transitions in EPR spectroscopy, the cavities are designed such that the sample is placed at the magnetic field maximum and electric field minimum in order to achieve high sensitivity.

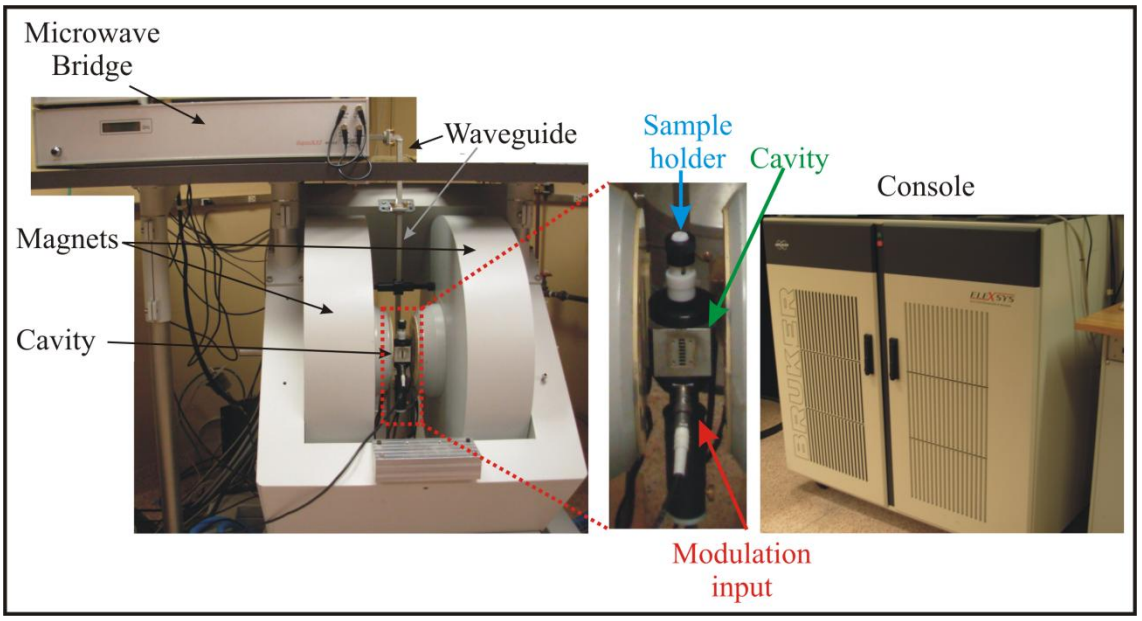


Figure 2.15 Main components of the EPR spectrometer

The magnetic field is produced by the magnet[16], which has the capacity of producing homogeneous and stable magnetic fields in the desired region (*i.e.*, ~3.5 kG for 9 GHz spectrometers and ~ 13 kG for 35 GHz spectrometers). The magnetic field stability is controlled by a Hall-probe attached to a pole piece near the center of the field. The additional component required in EPR spectroscopy is the console, which contains the signal processing and the controllers.

[16] In an electromagnet, a magnetic field is produced by passing a specified electric current through a wire.

2.6.1 Components of the microwave bridge. The microwave radiation is generated in the klystron microwave source (labeled 1 in Figure 2.16). In EPR spectroscopy it is important to be able to adjust the power level, so that enough power is provided to attain good sensitivity, while at the same time the power is kept low enough to avoid saturation. The output microwave power provided by the source cannot be easily controlled, thus a device known as the attenuator (labeled 2 in Figure 2.16) is used to accurately control the microwave power that reaches the sample chamber. The third component in the microwave bridge is the circulator (labeled 3 in Figure 2.16), which directs the radiation coming off the attenuator to the cavity and ensures that only the radiation reflected back from the cavity due to electronic transitions in the sample reach the detector. The reflected microwaves coming off the circulator are detected by the Schottky barrier diode (labeled 5 in Figure 2.16), which converts the microwave power to an electrical current. For optimal sensitivity and quantitative signal intensity, the diode should be operated at power levels greater than 1mW[17], thus in order to insure that the current reaching the detector is at the right level, a reference arm is used to supply the detector with some extra microwave power or "bias" (labeled 6 in Figure 2.16). There is a phase shifter in the reference arm, identified by the symbol φ in Figure 2.16, to ensure that the microwaves provided by the reference arm are in phase with

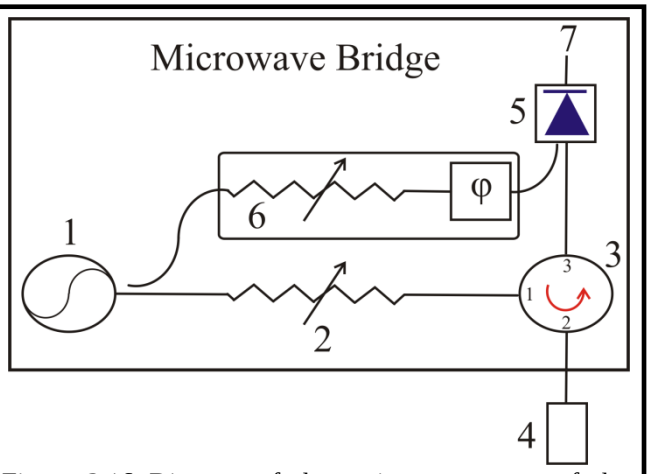


Figure 2.16 Diagram of the main components of the microwave bridge. 1 – Radiation source; 2 – Attenuator; 3 – Circulator; 4 – Cavity; 5 – Diode detector; 6 – Reference arm; 7 – Signal output

[17] At such power level, the detector is called a linear detector because the diode current is proportional to the square root of the microwave power.

the microwave power reflected back from the cavity when the two signals combine in the detector.

2.6.2 Signal channel. As mentioned above, the EPR signal coming from the cavity is small due to the small differences in the relative populations of the + ½ and – ½ electronic states. To enhance the sensitivity of the spectrometer it is necessary to selectively amplify the signal in order to increase the signal-to-noise ratio. EPR spectroscopists use a method known as magnetic field modulation with phase sensitive detection that encodes the desired EPR signal at a particular frequency while filtering out the noise. The magnetic field strength that the sample sees is varied sinusoidally at the modulation frequency of 100 kHz by using modulation coils mounted on the sides of the cavity walls. If there is an EPR signal, the field modulation sweeps through part of the signal and the amplitude of the microwaves reflected back from the cavity is modulated at the same frequency. If an appropriate modulation amplitude is used (*i.e.*, the modulation must be small compared to the intrinsic line width of the signal), the EPR signal will be linear over the set interval (*see* labels a and b in Figure 2.17), and the EPR signal is transformed into a sine wave with amplitude that is proportional to the slope of the signal (*see* left panel in Figure 2.17). The phase sensitive detector filters out all signals other than the signal modulated at 100 kHz by comparing the modulated signal with a reference signal having the same frequency and phase as the field modulation. As a result of the magnetic field

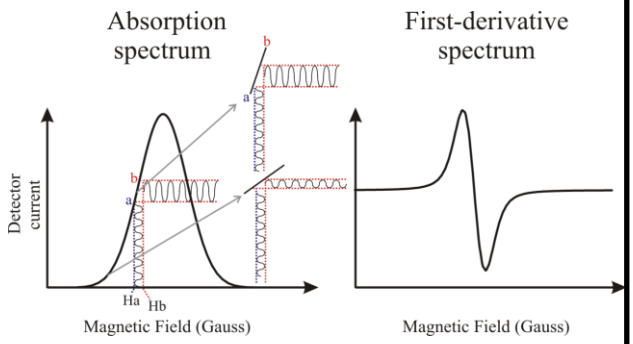


Figure 2.17 Phase modulation. *Left panel*: Absorption lineshape indicating the field modulation. Hb-Ha indicates the modulation amplitude. The effect of the slope of the absorption signal on the amplitude of the sine wave is shown. *Right panel*: First derivarive spectrum produced by the field modulation.

modulation, the absorption spectrum is detected as a first derivative spectrum as depicted in Figure 2.17.

2.7 Introduction to Nitroxides

Most proteins, save for metalloproteins, do not contained unpaired electrons, therefore they are intrinsically "EPR-silent". To study such proteins with EPR spectroscopy it is necessary to introduce an extrinsic paramagnetic sensor. As early as 1964 Burr and Koshland recognized that a "good" reporter group must be sensitive to changes in the environment and its position must be known and controlled by adding a "positioning" group that will guide the sensor to the appropriate site [16]. A year later, McConnell introduced the concept of spin label to describe stable free radicals that can be employed as reporter groups in otherwise "EPR-silent" proteins [2]. The most widely used paramagnetic sensors in EPR spectroscopy are nitroxide-derivatives (Figure 2.18), which contain an unpaired electron in the 2p orbital of the nitrogen. The nitroxide radicals are stable at temperatures up to 80℃ and over a pH range of 3-10. This relatively high stability is achieved by incorporating a quaternary carbon atom at the neighboring bonds such that the radical is protected from disproportionation reactions. In order to limit the flexibility of the nitroxide moiety, hence the unpaired electron, the nitroxide group is incorporated into either a six-membered piperidine or a five-membered pyrrole ring (Figure 2.18A).

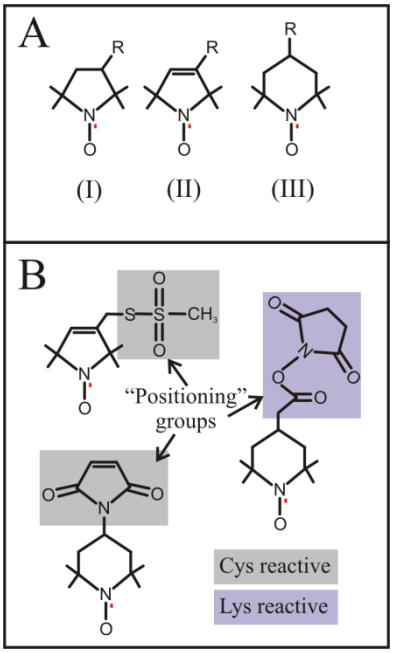


Figure 2.18 Commonly used nitroxides in SDSL. A) Saturated (I) and unsaturated (II) five-membered pyrroline rings. Six membered piperidine ring (III). B) Spin labels commonly used for covalent modification of proteins. The maleimide, methane thio-sulfonate, and hydroxysuccinimide functional groups are highlighted.

Nitroxide spin labels can be used as substrate or ligand analogs. Conversely, spin labels can be designed with a specific "positioning" functional group (shaded groups in Figure 2.18B) for direct and specific attachment to specific groups in the protein of interest. For example, maleimide and methanethiolsulfonate groups (highlighted in gray in Figure 2-18B) can be used to covalently attach nitroxide spin labels to cysteine residues, while hydroxylsuccinimide positioning groups (highlighted in lightblue in Figure 2-18B) can be employed to react with lysine residues[18]. Until the late 1980's, the main limitation for the application of EPR in most proteins was the reliance of naturally occurring cysteine residues and the inability to select specific sites for the incorporation of the paramagnetic sensor. The development of techniques in molecular biology, such as site-directed mutagenesis in the late 1980's, paved the way for the application of EPR spectroscopy to any recombinant protein *via* introduction of spin labels site-specifically using a technique pioneered by Hubbell [3] known as site-directed spin labeling (SDSL). In SDSL, a native residue at a specific site is mutated to a cysteine residue and the cysteine is then reacted with the paramagnetic probe to generate a nitroxide side chain (Figure 2.19). The most widely used nitroxide side chain and the one used for the most part in the work discussed in this dissertation is designated R1, but many others have been developed for specific applications [17-21].

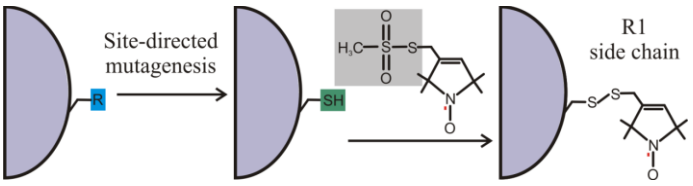


Figure 2.19 Site-directed spin labeling. A selected residue in the protein of interest is mutated to cysteine *via* site-directed mutagenesis and the cysteine-containing protein is reacted with the cysteine-specific methane thiosulfonate derivative to generate a nitroxide side chain. The most widely used nitroxide side chain is R1, which is shown in the right panel.

[18] Attaching to lysine residues is less desirable due to their relatively high abundance in most proteins, thus in most cases, cysteine residues are used as targets for covalent attachments of spin labels.

Several aspects of the EPR spectrum of the nitroxide group have been discussed in the previous sections, such as the hyperfine interaction giving rise to the characteristic 3-line spectrum and the anisotropy of the Zeeman and hyperfine interaction, which confer structural and dynamical sensitivity to EPR spectroscopy. Thus the next sections will focus on the specific effects of motion on the spectral lineshape and the experimental strategies that can be used to extract quantitative and qualitative information from the EPR spectra of spin-labeled proteins.

2.7.1 Effect of rate of motion on the EPR spectra. Conventional CW EPR spectroscopy is useful to extract information on motions occurring in the nanosecond to picosecond time scale. The sensitivity of the EPR signal to motion in this time domain arises from rotational motions that change the spin label orientation with respect to the magnetic field. Based on the time scale required to average out the A- and g- anisotropies (*see* sections 2.5 above), three motional regimes have been defined in EPR spectroscopy: *fast* ($\tau_r \approx 10^{-10} \sim 2 \times 10^{-9}$ seconds), *slow* ($2 \times 10^{9} < \tau_r > 100 \times 10^{-9}$ seconds), and *very slow* ($\tau_r > 100 \times 10^{-9}$ seconds). Motions on the fast and slow motional regime are on the order of T_2 for nitroxide spin labels (T_2 = 15 – 30 ns), hence such motions have effect on the CW spectral lineshape, whereas the very slow motions are on the order of T_1 (μs time scale) and can only be detected using pulsed-EPR techniques [22-23]. For motions faster than about 100 ps, the EPR spectrum will show three sharp lines of equal amplitude and no quantitative information on τ_R can be obtained from such spectra.

2.7.1.1 Fast motional regime (0.1 – 2 ns time scale). When the nitroxide group is tumbling rapidly and isotropically (*i.e.,* free nitroxide tumbling in non-viscous solution), the EPR spectra consist of three narrow lines of almost equal heights as illustrated in Figure 2.20 (top spectrum). The motion is sufficiently fast to average out all the A- and g- spectral anisotropy (*see*

solutions of equations 2.23 and 2.31 above). As the tumbling of the nitroxide group gets slower, there is a differential broadening of the lines corresponding to different nuclear manifolds (+1, 0, -1). This is illustrated in Figure 2.20, which clearly shows that the high field line (denoted -1) gets broader more readily than the other spectral lines as the motion decreases. Eventually when τ_R is longer than ~100 ns, the EPR spectra is less sensitive to motion and resembles a powder spectrum. In the fast motional regime, the tumbling of the protein can be obtained directly from the relative line widths of the spectrum using Redfield's perturbation theory [24], where the correlation time is given by the following relation:

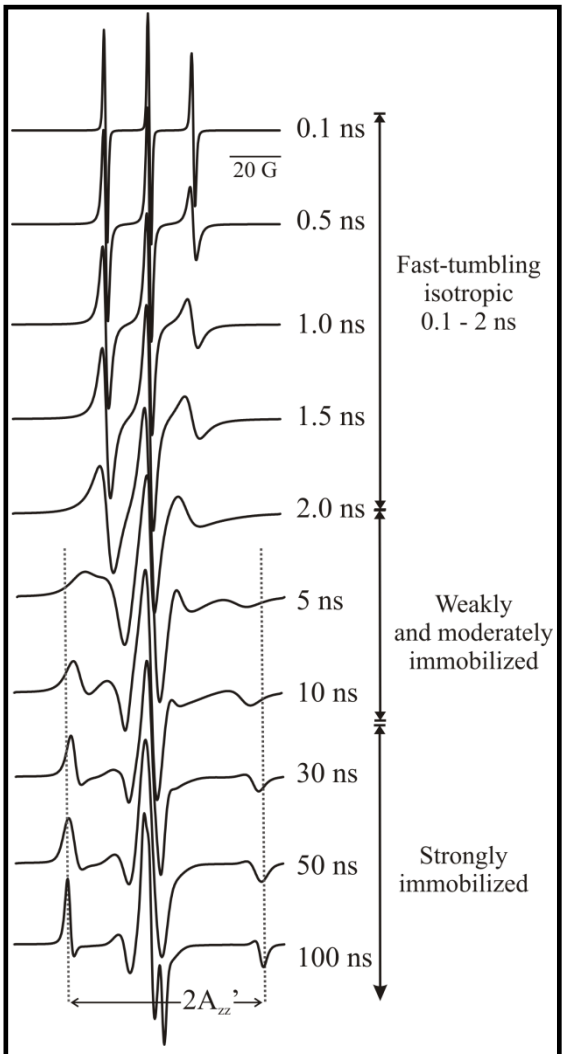


Figure 2.20 The effect of motion of a nitroxide rotating isotropically on the EPR spectra. The corresponding correlation time (τ_R) for each spectrum is indicated. The motional regimes describing the motion are indicated. The well-resolve hyperfine extrema (A'_{zz}) for spectrum in the strongly immobilize motional regime is indicated. Dashed lines are use to guide the eye.

$$\tau_R = 6.5x10^{-10}\Delta H_0\left(\sqrt{\frac{h(0)}{h(-1)}} - 1\right), \qquad (2.34)$$

where $h(0)$ and $h(-1)$ are the heights of the central and high-field, respectively and ΔH_0 is the line width of the central-line in gauss. In equation 2.34, the factor 6.5 x 10^{-10} is derived from the known nitroxide hyperfine and g-value anisotropies [25].

2.7.1.2 Slow motional regime (2 ~ 100 ns time scale). As the motion of the nitroxide is restricted, such as the case of spin labels in well-ordered proteins, the A- and g- anisotropy,

which determine the lineshape and positions of the resonance, respectively, are no longer averaged out and no direct calculation of τ_R from the spectral line widths is possible. Thus to extract motional parameters from such spectra, it is necessary to use a different theory that takes into account the anisotropic contributions from the A- and g- tensors. Such a theory was developed by Freed and required spectral simulations using the stochastic Liouville equation [26]. The theory will not be derived in here, but interested readers are referred to articles by Freed [26], Schneider et al.[27], and Budil et al. [28] for further details on the theory and implementation. As depicted in Figure 2.20, the effective hyperfine splitting of the outer extrema (labeled $2A_{zz}'$) becomes resolved for $\tau_R \geq 10$ ns and becomes increasingly larger as τ_R approaches the rigid limit (~ 100 ns). In addition, the line width of the hyperfine extrema becomes narrower due to T_2 relaxation effects. Freed employed empirical calibrations of spectra to get an estimate of τ_R assuming isotropic motion. Such spectra can be reproduced by slowing down the isotropic motion of a spin label in solution by either decreasing the temperature or by increasing the viscosity of the solution as indicated by the Stokes-Einstein relationship:

$$\tau_R = \left(\frac{4\pi\eta r^3}{3kT}\right), \tag{2.35}$$

where η is the viscosity of the solution, T is the temperature (in Kelvin), r is the hydrodynamic radius of the molecule, and k is the Boltzmann's constant. Such an approach yielded the following empirical equation to calculate τ_R, which assumes an intrinsic line width of 3G:

$$\tau_R = a\left(1 - \frac{A'_{zz}}{A^R_{zz}}\right)^b, \tag{2.36}$$

where a = 0.54 ns, b = -1.36, and A^R_{zz} is the rigid limit value for the hyperfine extrema. In this approach, care should be taken in selecting the appropriate conditions for the rigid limit values,

since A'_{zz} is sensitive to the polarity of the environment as discussed above. A user-friendly simulation program based on the SLE theory was developed by Budil et al. [28] to simulate spectra in the slow motional regime (*see* section 2.7.3 below for additional details).

2.7.1.3 Very slow motional regime (τ_R > 100 ns). When the motion of the nitroxide becomes too slow, the CW EPR spectra is no longer sensitive to its motion. Nevertheless, motions slower than 100 ns are functionally important, thus it is important to develop experimental strategies for direct measurements of such motions in spin-labeled proteins. There are two pulse-EPR based methods that have been developed to study protein motions on the microsecond time domain; namely saturation recovery (SR) EPR and electron-electron double resonance (ELDOR). The details of both methods are beyond the scope of this dissertation, thus the underlying theory will not be discussed here. Interested readers are referred to Bridges et al. [22], Fleissner et al. [23] and references therein for details on the theory and application of such strategies.

2.7.2 Anisotropic spin label motion. In many instances, the spin label cannot freely tumble in all directions in solution (*i.e.,* nitroxide covalently attached to a protein). For such cases, motional averaging occurs only in some directions. The interpretation of anisotropic motion is more easily visualized by considering a lipid spin label diluted into a bilayer as depicted in Figure 2.21A. In such cases, one can design a lipid spin label such that when the nitroxide is inserted in the bilayer, the z-axis of the nitroxide is parallel ($A_{\|}$) to the long molecular axis of the lipid. If the chain rotates rapidly about the long molecular axis, then only the A_{xx} and A_{yy} and g_{xx} and g_{yy} anisotropy will be averaged, so that $A_{\perp} = ½ (A_{xx} + A_{yy})$ and $g_{\perp} = ½ (g_{xx} + g_{yy})$, thus only motion about A_{zz} will contribute to the spectra.

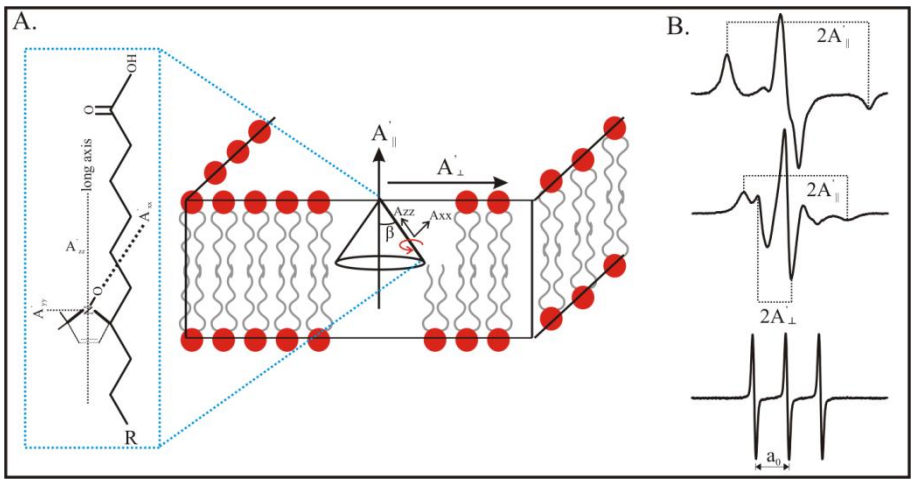


Figure 2.21 Anisotropic motion of a spin label intercalated in a lipid bilayer. A) The nitroxide principal axes are indicated in the spin label shown in the left panel. The z-axis of the intercalated label is directed along the long molecular axis of the lipid as indicated. The axis of motional averaging is parallel (||) to the long molecular axis of the lipid. B) EPR spectrum arising when the lipid spin label rotates rapidly, but with restricted amplitude ($\beta = 0^{\circ}$) along the z-axis (top spectrum). EPR spectrum of a lipid spin label with rapid motion of higher amplitude ($\beta > 0^{\circ}$) is shown in the middle of panel B. The parallel and perpendicular components of the hyperfine are indicated. The isotropic spectrum is shown at the bottom panel.

Such a motion is characterized as a fast motion with restricted amplitude and gives rise to a spectrum that resembles a powder spectrum (top spectrum in Figure 2.21B). Conversely, if the acyl chains can "wobble", the amplitude of the motion is increased and A_{zz} is averaged along with A_{xx} and A_{yy}. For such cases, A_{zz} is reduced from its principal value and A_{xx} and A_{yy} increase to the point where $A_{\perp}$ can be resolved in the spectrum. In the case of fast isotropic motion $A_{||} = A_{\perp}$. For anisotropic motion, the amplitude of the motion is defined by an order parameter (S) and can be obtained with the following equation:

$$S = \frac{A_{||} - A_{\perp}}{A_{zz} - \frac{1}{2}(A_{xx} + A_{yy})} \tag{2.37}$$

It follows that S = 1 for highly ordered systems ($\beta = 0^{\circ}$) and S = 0 for completely isotropic motion ($\beta \geq 54.7^{\circ}$). For intermediate cases, S is related to the angular amplitude of the anisotropic molecular motion:

$$S = \frac{1}{2}(3 < cos^2\beta > -1), \tag{2.38}$$

where the brackets represent a time average over the molecular motion.

The above discussion assumes fast motion of a particular amplitude that can be described by S. Nevertheless, a more complex model is necessary to describe the motion of the nitroxide side chain in spin-labeled proteins for which both the *rate* ($1/\tau$) and *amplitude* of the motion contribute to the observed spectral lineshape. The effect of changes in the rate and order of motion on the EPR spectra are shown in Figure 2.22. By comparing the spectra in Figures 2.20 and 2.22, it is evident that it is difficult to distinguish the contributions from the rate and order to the EPR spectra, thus it is not possible to extract information about the rate and order *directly* from the lineshape. In order to describe the motion of the nitroxide in terms of rate and amplitude, it is necessary to simulate the spectra using the Stochastic Liouville (SLE) equation described by Freed and co-workers [26].

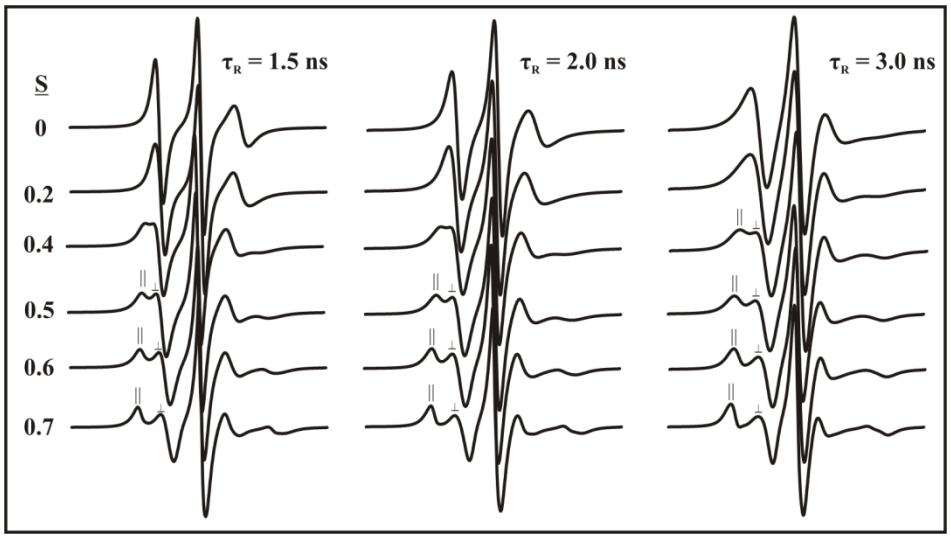


Figure 2.22 Dependence of the X-band EPR spectra on the rate ($1/\tau_R$) and amplitude (S) of motion of the nitroxide side chain. During the simulation, the order parameter was varied from 0 – 0.7 at the three correlation times indicated (1.5, 2, and 3 ns). Notice that for S ≥ 0.4, the parallel (||) and perpendicular (⊥) components of the hyperfine extrema are well-resolved.

2.7.3 EPR spectra simulation using the SLE method. To simulate the spectra using the SLE approach, Budil developed a Windows-compatible program known as NLSL [28]. The program uses a model known as ***m****icroscopic* ***o****rder* ***m****acroscopic* ***d****isorder* (MOMD), which assumes that the nitroxide undergoes a microscopic molecular ordering (*see* below) with respect to a local director (labeled z_D in Figure 2.23), however the local directors are randomly oriented with respect to the laboratory frame (*see* inset in Figure 2.23). In this model, three coordinate systems are employed to describe the motion of the nitroxide in a spin-labeled protein. The first is the *director frame* (x_D, y_D, z_D), which is fixed relative to the structure of the molecule to which the nitroxide is attached. The second coordinate system is the *magnetic frame* (x_M, y_M, z_M), which has been defined such that the z_M lies along the 2p orbital of the nitrogen nucleus, x_M lies along the N-O bond, and y_M lies perpendicular to both x_M and z_M. The magnetic frame is fixed relative to the nitroxide molecule. The third coordinate system is the *rotational diffusion frame* (x_R, y_R, z_R). The relative orientation between the rotational and magnetic frame is specified by the diffusion tilt angles $\Omega_D = (\alpha_D, \beta_D, \gamma_D)$ as shown in Figure 2.23.

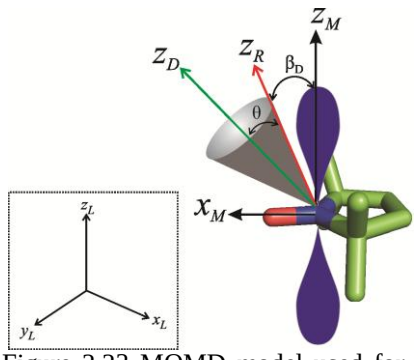


Figure 2.23 MOMD model used for calculation of slow-motion EPR spectra using the NLSL software package. The 3 coordinate systems used to describe the motion of the nitroxide side chain are shown. The β_D angle defines the tilt of z_R with respect to z_M. θ is the angle between z_R and z_D. The order parameter S describes the width of the cone. *Inset* Laboratory frame. z_D for each molecule is randomly oriented with respect to the laboratory frame, thus the system is **macroscopically**

The fitting program includes a mechanism to describe the tendency of the spin label side chain to remain ordered within its local environment by using an ordering potential, which specifies the orientational distribution of z_D in the rotational diffusion frame. The ordering potential is expanded in a series of spherical harmonics,

$$U(\Omega) = -\sum_{L,K} c_{L,K} D^{L}_{0K}(\theta,\phi), \tag{2.40}$$

where Ω is the set of diffusion tilt angles, and $c_{L,K}$ are the coefficients of the potential, which are expressed in units of *k*T. Due to the restrictions in the L and K values, there are only five possible terms of the potential, namely c_{20}, c_{22}, c_{40}, c_{42}, and c_{44}. For the simulations done in this dissertation, only the term c_{20} was used, so that the amplitude of motion (S) was described by the first term of the expansion $D^2_0(\theta)$, whose functional form is ½ (3 $\cos^2\theta$ - 1), thus:

$$S = \frac{1}{2}(3 < cos^2\theta > -1), \tag{2.41}$$

where θ is as defined in Figure 2.23 and the brackets indicate spatial average.

For the calculation of rates of motion, the $\log_{10}$ of the three components of the Cartesian rotational diffusion tensor (R_{xx} ,R_{yy}, R_{zz}) are given in a modified spherical form:

$$\bar{R} = \sqrt[3]{R_{xx}R_{yy}R_{zz}} \tag{2.42}$$

$$N = \left(\frac{R_{zz}}{\sqrt{R_{xx}R_{yy}}}\right) = \frac{R_{||}}{R_{\perp}} \tag{2.43}$$

$$N_{xy} = R_{xx} - R_{yy} \tag{2.44}$$

For isotropic motion $R_{||} = R_{\perp}$, thus $N = 0$ and for an axially symmetric system $N_{xy} = 0$. The effective correlation time (sec) can be calculated from the rbar ($\bar{R}$) value using the following equation

$$\tau = \frac{1}{6x10^{\bar{R}}} \tag{2.45}$$

For example, for rbar = 8.0, the corresponding effective correlation time is 1.67 ns. Additional parameters which are varied during the simulations are the inhomogeneous line width tensor (**W**), which takes into account the anisotropic line broadening due to local electrostatic inhomogeneity and the gaussian inhomogeneous broadening (**gib0** and **gib2**) that arises from unresolved hyperfine interactions between the unpaired electron spin and the surrounding ^{1}H nuclei [29].

Typically during the fitting procedure, the magnetic parameters (A and g tensors) are set to values determined from a spectrum recorded in frozen solution and the dynamic parameters, namely *rbar*, *N*, c_{20}, (α_D, β_D, γ_D) are varied until the simulated spectrum closely resembles the experimental spectrum and the X^2 (as defined in Earle and Budil [29]) has reached its minimum value. For anisotropic motion about the z-axis, typical values for the Euler angles are set to (0°, 36°, 0°), however in a few cases the motion of the nitroxide is best described by x-axis anisotropic motion [30], in which $\beta_D = 90°$. It should be emphasized that some parameters are correlated (*i.e.*, rbar and c_{20}) such that changes in one can be compensated by changing the other, therefore there are multiple equally-satisfying fits to the spectra. As a result of this degeneracy, spectral simulations should be used as approximation to test whether a particular motional model is a plausible model to describe the experimental data, but should never be taken as an empirical observation.

2.8 Applications of SDSL

The spectral anisotropy and its sensitivity to various relaxation pathways are the principal basis for most of the applications of SDSL. The dynamic range of EPR experiments based on lineshape analysis is 0.1 ~ 100 ns, thus functionally important protein motions occurring in this

time domain are expected to contribute directly to T_2 relaxation, and hence determine the EPR spectra of spin-labeled proteins. Indeed, SDSL has been used to map sequence-dependent backbone motions in soluble and membrane proteins [31]. The sensitivity of EPR to slow motions (100 ns and longer) can be extended with various means, one of which is a strategy that was developed as part of this dissertation (*see* Chapter 3).

The EPR spectra of spin-labeled proteins are also sensitive to the presence of multiple paramagnetic species in the sample (*i.e.*, a second bound nitroxide spin label or transition metal), due to changes in the intrinsic relaxation of the nitroxide by the second paramagnetic species. This mechanism is distance-dependent, thus EPR can be employed to determine the accessibility of the nitroxide to the second paramagnetic species and to directly determine inter molecular distances.

Some of the most widely used applications of SDSL for obtaining structural and dynamic information on protein and proteins complexes will be discussed below. Additional applications not discussed here can be found in the following references [22,32].

2.8.1 Solvent Accessibility. Measuring the solvent accessibility of the nitroxide side chain to paramagnetic quenchers is useful for determining the local environment of the nitroxide side chain. Important topographical and structural information can be obtained from such experiments. For example, the immersion of local elements of the peptide backbone within the lipid bilayer in membrane bound proteins can be elucidated. In addition, the local secondary structure of soluble proteins can be determined by doing nitroxide scanning and following the accessibility of the nitroxide to polar and non-polar quenchers (*see* below). The experimental strategy to determine solvent accessibility of the nitroxide involves adding a fast relaxing

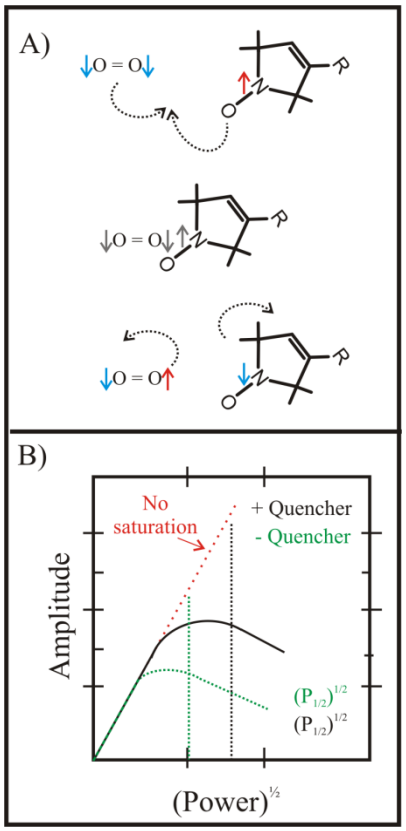


Figure 2.24 Effect of molecular collisions between a nitroxide and an oxygen molecule on the electron spin relaxation (**T_1**). A) The ↑ and ↓arrows represent +½ and – ½ spin states of the unpaired electrons, respectively. The unpaired electrons in oxygen and nitroxide are color-coded blue and red, respectively. Note that during the collision, the unpaired electrons are indistinguishable. B) Power saturation experiment used to determine $P_{1/2}$ of a particular sample. The addition of a quencher relieves saturation, thereby more power is required to saturate the signal.

paramagnetic agent (*i.e.*, O_2 or Ni(II)) to the solution containing the spin-labeled protein, and measuring the effect of the paramagnetic quencher on the EPR spectra or relaxation of the spin label. The underlying principle of the method is illustrated in Figure 2.24A for an oxygen molecule colliding with a nitroxide. Consider a collision between a nitroxide with an unpaired electron in the + ½ state (excited spin state) with an oxygen molecule having the unpaired electrons in the – ½, – ½ state. When the two molecules collide, the three unpaired electrons are indistinguishable due to overlap of the wave functions. When the two molecules separate, there is a ⅔ probability that the nitroxide exchanges its +½ unpaired electron with one of the unpaired electrons of the oxygen molecule in the –½ state (this process is known as Heisenberg Spin Exchange). The collision has relaxed the nitroxide from the +½ state to the –½ state (**this is T_1 relaxation!**). If T_1 of the quencher is faster than the intrinsic T_1 of the nitroxide and the collisions are frequent, the exhibited T_1 for the nitroxide is shortened in comparison with its intrinsic T_1. T_2 is also shortened leading to line broadening. Collisions leading to Heisenberg spin exchange enhance T_1 and T_2 according to the following equation:

$$\Delta T_1^{-1} = \Delta T_2^{-1} = kW_{ex} \tag{2.46}$$

where k is a factor that accounts for the collisions and W_{ex} is the biomolecular collision

frequency. The relaxation enhancement can be measured directly by pulse EPR methods such as saturation recovery [33] or by continuous wave power saturation [34]. In the latter method, the EPR spectra are recorded as a function of microwave power both in the presence and absence of the paramagnetic quencher. At powers below saturation, the amplitude of the signal increases linearly with the microwave magnetic field ($H_1 \alpha P^{1/2}$) until the Boltzmann equilibrium distribution is perturbed (due to inefficient T_1) and the signal between the excited and ground states is perturbed (Figure 2.24B). The addition of fast relaxing quenchers makes T_1 shorter, and hence more efficient, thereby relieving the saturation, such that more power is required to saturate the signal. A fit of the power dependence of the amplitude of the central line yields the parameter $P_{1/2}$, which is the power required to reduce the amplitude to one-half its value in the absence of saturation as depicted in Figure 2.24B. The difference in $P_{1/2}$ values for + and - quencher samples is a measure of the accessibility of the spin label to the quencher.

Normalization of the central linewidth and $P_{1/2}$ to those of a standard sample[19] (usually diphenyl picrylhydrazyl; DPPH) yields a dimensionless accessibility parameter (Π), where

$$\Pi = \frac{\Delta P_{1/2}}{\Delta H_0} \frac{\Delta H_0^{DPPH}}{P_{1/2}^{DPPH}} \tag{2.47}$$

Differential accessibility of the nitroxide side chain to polar (nickel ethylenediaminediacetic acid: NiEDDA) and nonpolar (oxygen) fast relaxing species can be used to measure immersion depth of membrane proteins [35] and to determine the local secondary structure of the backbone [36].

[19] The standard sample DPPH is used to account for inherent differences in the resonators and spectrometers between different laboratories.

2.8.2 Mapping local secondary structure. The environment around the nitroxide side chain is the primary determinant of the nitroxide motion in spin-labeled proteins. For example, in the context of the protein fold, the mobility of the nitroxide side chain located at completely buried sites is expected to be highly constrained due to the high packing density around the nitroxide side chain, whereas a nitroxide located at a partially solvent exposed site is expected to exhibit constrained motion (but of higher mobility relative to a buried site) due to interaction with nearby groups. Conversely, a nitroxide side chain in a solvent exposed site with no tertiary interaction is expected to exhibit relatively high mobility. Indeed, comprehensive SDSL studies [17] in combination with x-ray studies [21,37-38] of spin labeled T4 lysozyme confirmed these expectations. A representative data set from such studies is shown in Figure 2.25. As expected, the EPR spectrum of the solvent inaccessible residue L118R1 in T4 lysozyme exhibit low mobility of the nitroxide side chain due to steric interaction with nearby groups [38]. In contrast, the EPR spectrum of the non-interacting solvent-exposed residue T151R1 reflects fast (~ 1.4 ns) anisotropic motion [21], while the spectrum of residue V75R1 in T4 lysozyme, which is in tertiary contact with a nearby tyrosine [37], exhibits intermediate mobility between buried and solvent exposed sites.

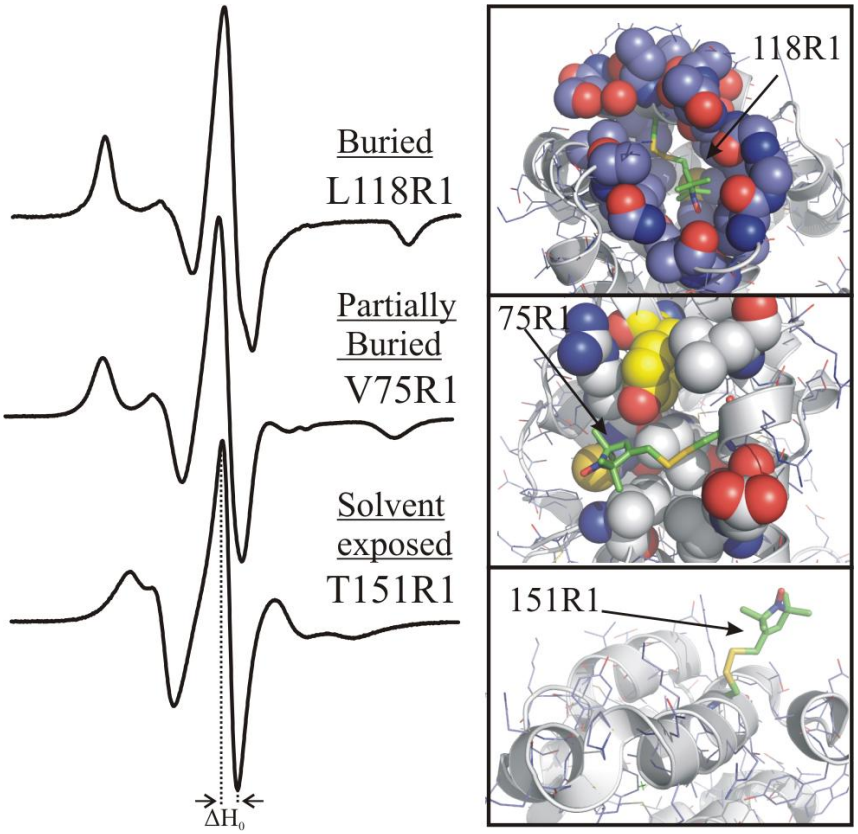


Figure 2.25 The EPR spectra of spin-labeled proteins are determined by the local environment around the nitroxide side chain. *Left panel* EPR spectra of T4 lysozyme at a buried, partially buried, and solvent exposed sites. *Right panel* High resolution structures of the corresponding spin-labeled proteins indicating the position and packing around the nitroxide side chain (stick representation). The interacting residue Y88 in the structure of V75R1 is highlighted in yellow (Langen et al. [37]). The central line width (ΔH_0) used to measure relative mobility is indicated.

The topographical-dependent mobility of the nitroxide side chain provides a means for mapping the local secondary and tertiary fold using SDSL. For the elucidation of secondary structures using SDSL, the strategy involves doing a nitroxide scanning[20] experiment and measuring the mobility of the nitroxide side chain by EPR spectroscopy. The mobility of the nitroxide side chain is expected to be a periodic function of the sequence position for unevenly packed secondary structural elements (*i.e.*, α-helices and β-sheets). The periodic pattern of R1

[20] Nitroxide scanning – a single native side chain is replaced sequentially with a nitroxide side chain.

mobility provides information about the secondary structural element to which the nitroxide is attached to.

The efficacy of this approach was demonstrated in studies done in T4 lysozyme Mchaourab et al. [17] and cellular retinol-binding protein Lietzow et al. [39]. The results of these studies are summarized in Figure 2.27. Panels B and E of Figure 2.26 show the structures of T4 lysozyme and CRBP, respectively with the sites where the nitroxide was introduced indicated by green spheres at the Cα. The corresponding EPR spectra are shown in panels A and D. In these studies, the central line width (*see* ΔH_0 in Figure 2.25) was used as a semi-quantitative measure of mobility Mchaourab et al. [17]. As shown in the ΔH_0 *versus Residue* plots (*see* panels C and F), there is a clear periodicity in the mobility of R1 in α-helices and β-sheets. The mobility of R1 in α-helices exhibits periods of mobility of ~3.6, while the periodicity for β-sheets is 2. These results clearly indicate that the periodicity of mobility of the R1 side chain provides a means to identify sequence-specific secondary structure in spin-labeled proteins. In the same studies, the accessibility of the nitroxide side chain for the same sites to the paramagnetic quenchers oxygen and NiEDDA was analyzed as an independent means to map secondary structure. As shown in Figures 2.26 the periodicity observed from mobility of the nitroxide side chain and from accessibility studies are highly correlated.

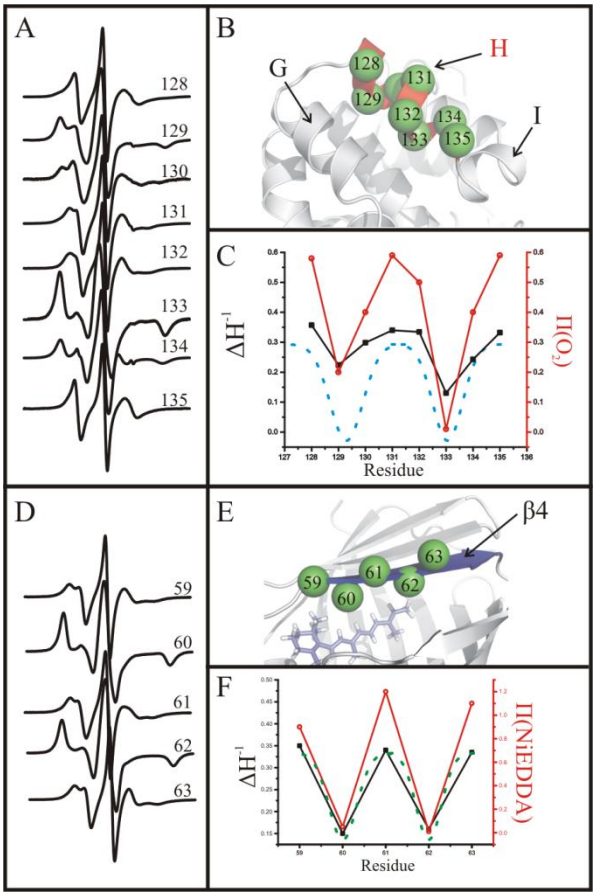


Figure 2.26 Mapping secondary structure with SDSL. Nitroxide scanning studies in T4 lysozyme and CRBP are shown. A) EPR spectra of R1 in α-helix H of T4L. B) Ribbon diagram of T4L showing the sites where the nitroxide was introduced. C) Plot of side chain mobility- and accessibility to NiEDDA vs residue for T4L is shown. D) EPR spectra of R1 in β4 sheet in CRBP. E) Ribbon diagram of CRBP showing the sites where the nitroxide was introduced. F) Plot of side chain mobility- and accessibility to NiEDDA vs residue for CRBP is shown. The periodicities observed in the data are shown in dashed lines.

2.8.3 Mapping tertiary structure with SDSL. In the most common implementation of SDSL, a single nitroxide side chain is introduced and the mobility and accessibility of the nitroxide side chain to the solvent can be studied as discussed above. However, obtaining information on the tertiary fold of a protein by SDSL requires a fundamentally different strategy. To map the tertiary fold of the protein by SDSL one can take advantage of the sensitivity of nitroxide spin labels to the presence of a second paramagnetic species (*i.e.*, paramagnetic metal ion or a second nitroxide label). The experimental strategy relies on the distance-dependent interaction between spins, thus two paramagnetic species can be introduced into a protein of

interest and the type and strength of the interaction between the paramagnetic probes can be analyzed in terms of interspin distance. Depending on the method employed, EPR is capable of measuring interspin distances from 2-80 Å [40-42]. The physical basis for the sensitivity of the nitroxide to a second paramagnetic species is exchange interaction (*J*) arising from overlap of the orbitals of the unpaired electrons, which typically occurs when the interspin distance (r) is 2-8 Å and the dipolar interactions between magnetic moments of the two paramagnetic species (r ~ 8-80 Å). The underlying theory for the study of interspin distance based on exchange interaction will not be discussed here, but interested readers are referred to Fiori et al. [40] and references therein for details on the theory and application of such strategy.

2.8.3.1 *Distance measurements based on dipolar coupling*. The interaction between two spins with r ≥ 8 Å arises because the magnetic dipoles of each spin induce local fields near the second spin. The local fields can add or subtract to the external magnetic field, resulting in further splitting of the resonance lines (this is analogous to the splitting of the resonance line of an unpaired electron due to the local fields induced by nearby nuclei as depicted in Figure 2.10). The strength of the dipolar interaction is distance-dependent (r^{-3}) and results in broadening of the lineshape for distances up to ~ 25 Å. There are two conditions that allow EPR spectroscopists to measure interspin distances in the range of 8 - 25 Å, namely the rigid limit condition, in which all motions are frozen out and the fast motion limit where rapid isotropic tumbling of the protein (τ_R < 6 ns) averages the dipolar interaction and leads to T_2 relaxation effects. In the rigid limit condition, the resonance peaks of the spectrum are inhomogeneously broadened and the broadening can be described by the following equation [43]:

$$\Delta H_{dd} = \pm \frac{3g\beta(3cos^2\theta - 1)}{4r^3}, \tag{2.48}$$

where θ is the angle between the interspin vector and the magnetic field. One of the disadvantages of the rigid limit approach is that it requires frozen solutions. However, Altenbach et al. [41] showed that it is possible to get an estimate of interspin distances in slow tumbling proteins using the rigid limit theory in cases where the nitroxide have fast internal motions (few ns) relative to the protein and the tumbling of the protein fulfills the condition expressed in equation 2.52.

$$\tau_R \geq \frac{h\, r^3}{3\pi g^2 \beta^2} \tag{2.49}$$

For proteins that tumble rapidly (≤ 6 ns), the interspin distance can be measured at physiological temperatures using the Redfield relaxation theory whereby the rotational modulation of the interspin vector results in homogeneous line broadening (ΔH_{dd}) characterized by T_2 relaxation effects. The broadening can be predicted from a particular distance and τ_r using the following relations:

$$\Delta H_{dd} = \frac{h}{g\beta\pi T_{2dd}} \tag{2.50}$$

$$\frac{1}{T_{2dd}} = \frac{3}{20}\gamma^4 \left(\frac{h}{2\pi}\right)^2 \frac{1}{r^6}\tau \left(3 + \frac{5}{1+\upsilon^2\tau^2} + \frac{2}{1+4\upsilon^2\tau^2}\right), \tag{2.51}$$

where T_{2dd} is the transverse relaxation time, γ is the gyromagnetic ratio of the electron, τ is the correlation time of the protein, and υ is the frequency of the microwave radiation. In this experimental strategy, the EPR spectra of single and double R1 mutants are recorded and the changes in the lineshape between single and double-labeled mutants reveal the extent of dipolar broadening from which the distance (r) can be obtained. Using this experimental approach, Machaourab et al. [44] derived interspin distances in T4 lysozyme and demonstrated that the enzyme in solution is in conformational equilibrium between an open and closed state of the

active site cleft, which is a functionally important transition for substrate binding. The magnitude of the open-close transition deduced from interspin distance measurements was ~ 8 Å, which is in agreement with the distance changes expected from the crystal structures solved in the two forms [45].

The experimental approaches discussed above rely on broadening of the spectral lineshape by a second paramagnetic species, thus they are limited to interspin distances up to 25 Å. To measure interspin distance above 25 Å a different approach is required, which does not rely on spectral broadening. Pulse-EPR methods such as double quantum coherence (DQC) and double electron electron resonance (DEER) have been developed for this purpose. The underlying theory of DQC is beyond the scope of this dissertation and will not be discussed in here, but interested readers are referred to Borbat et al. [48] and references therein.

DEER spectroscopy also known as PELDOR (pulse electron electron double resonance) has emerged as a powerful tool in structural biology to map tertiary structure and monitor structural changes in proteins and protein complexes [47]. DEER spectroscopy is well-suited to determine interspin distances in the 20Å - 80Å range [48], which makes it particularly attractive for mapping long-range structural changes. An important feature of the DEER experiment is the capability of obtaining distance distribution, which can be used to extract information about the internal flexibility of the paramagnetic probe [49] as well as backbone flexibility.

In the DEER experiment, the magnetic moments of the unpaired electrons are perturbed by microwave pulses and the response is detected as a delayed microwave signal known as electron spin echo (ESE). In the simplest implementation, the ESE signal can be generated using the pulse sequence depicted in Figure 2.27. The $\pi/2$ (90º) pulse turns the equilibrium net

magnetization from the z-axis into the *xy* plane. Right after the pulse, all the spins are coherent and begin to precess along the *xy* plane. However individual spin packets have difference resonance offsets due to local field inhomogeneities, thus some spins precess faster than others (*see* color-coded arrows in Figure 2.27). A second π pulse (180°) is applied at time τ after the first pulse, which inverts the phase of each spin by 180°. The second pulse places the fast precessing spins "behind" the slow precessing spins. At a time 2τ after the first pulse, the faster and the slower spins are coherent along the *xy* plane giving rise to a net transverse magnetization that can be observed as a signal known as the electron spin echo. This π/2 – τ – π – τ – echo pulse sequence was first introduced by E.L. Hahn in 1950[21], therefore it is known as the Hahn echo pulse sequence.

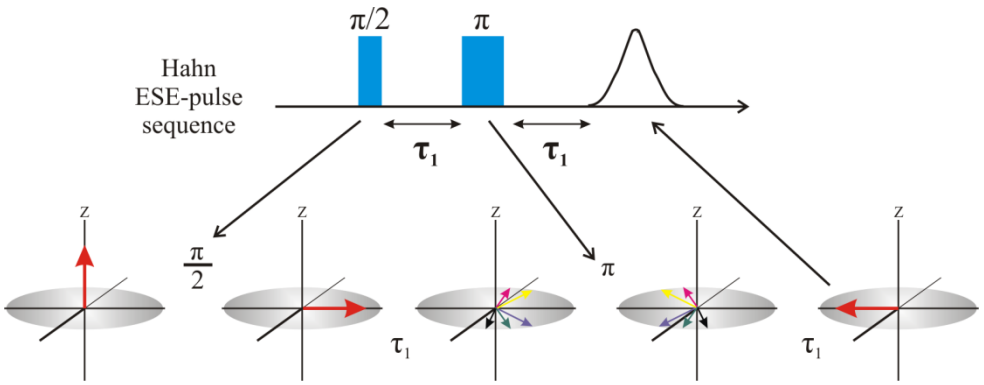


Figure 2.27 Hahn pulse sequence. *Top panel*: Diagrammatic representation of the simplest pulse sequence employed to generate an electron spin echo signal. The pulses are represented by blue bars. τ1 represents time intervals. *Bottom panel*: The changes in the net magnetization due to the applied microwave pulses are shown. The gray oval represents the *xy* plane. Individual spin packets precessing along the *xy* plane at different rates are color-coded for clarity.

Unlike the Hahn echo pulse sequence, the DEER pulse sequence consists of 4 pulses [50], one of which is applied at a different frequency as depicted in Figure 2.28B. In the 4 pulse sequence, the ESE signal of an unpaired electron is observed (observer electron 1 in Figure 2.28A) and the magnetization of the other electron (interacting electron 2 in Figure 2.28A) is

[21] Hahn, E.L. 1950. Spin echoes. *Physical Review* 80:580–594.

flipped 180̊ along the z-axis by a π pulse of different frequency. The change in the magnetization of the interacting electron changes the magnetic field felt by the observer electron, which in turn modulates the ESE signal. The frequency of the modulation is determined by the dipole-dipole coupling (ω_{dd}) of the unpaired electrons, which in turn is proportional to r^{-3} as shown by the following equation [51]:

$$\omega_{dd} = \frac{2\pi g_1 g_2}{g_e^2}(3cos^2\theta - 1)\frac{52.04}{r^3}, \quad (2.52)$$

where g_1 and g_2 are the g values of the two spins, g_e is the g value of the free electron, r is the interspin distance, and θ is the angle between the spin-spin vector and the magnetic field as depicted in Figure 2.28A. The primary data in a DEER experiment is the amplitude of the electron spin echo as a function of time, which has contributions from intra- and random inter-molecular dipolar interactions. The dipolar evolution function shown in Figure 2.28C is obtained after the exponentially decaying background is subtracted to correct for random inter-molecular dipolar interactions. Fourier transformation of the dipolar evolution function generates a dipolar spectrum and either one can be fit to get the interspin distance distribution (*see inset* Figure 2.28C).

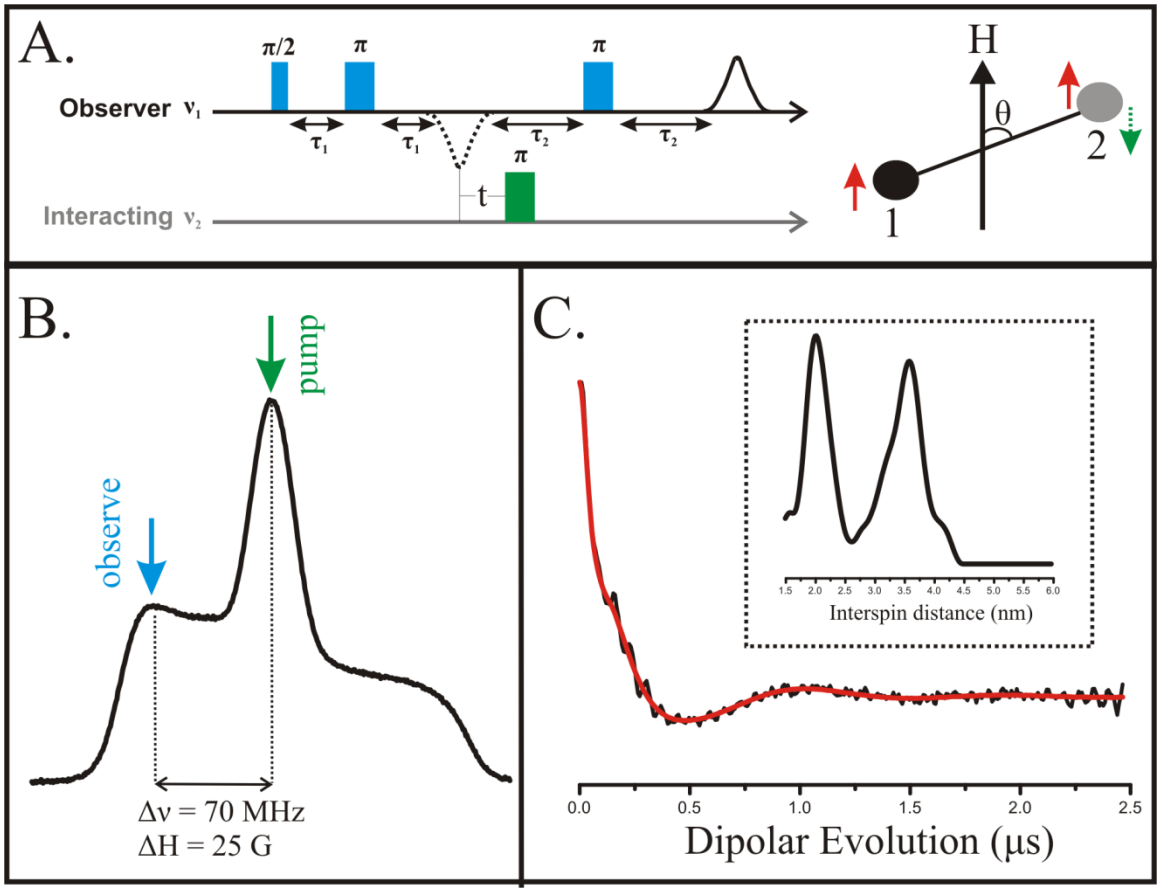


Figure 2.28 DEER pulse sequence. A) Diagrammatic representation of the pulse sequence used in DEER spectroscopy. The observed and pump pulses are represented by blue and green bars, respectively. The undetected echo after the first π pulse on spin 1 is shown in doted lines. A pair of electrons spins coupled via electron dipole-dipole interaction is shown. Notice that the spin of the pump electron (gray) is inverted (*see* green arrow) after the π pulse. θ represents the angle between the interspin vector and the magnetic field (H). B) Absorption spectrum indicating the pump and observed positions. C) Dipolar spectrum and corresponding distance distribution. The fit of the dipolar spectrum is shown in red.

2.8.4 Determination of fast (nanosecond) backbone flexibility with SDSL. As mentioned above, motion of the nitroxide side chain on the nanosecond time scale contribute directly to T_2 relaxation and hence determine the EPR spectra lineshape. For a spin-labeled protein, the overall motion of the nitroxide side chain on the nanosecond time scale (τ_T) is determined by its environment and has contributions from three dynamic modes, namely: overall rotary diffusion of the protein (τ_r), backbone fluctuations (τ_{bb}), and internal motions of the side chain (τ_{int}). Thus, to a good approximation

$$\frac{1}{\tau_T} = \frac{1}{\tau_r} + \frac{1}{\tau_{bb}} + \frac{1}{\tau_{int}} \tag{2.53}$$

To extract information on the contribution of backbone motions to the mobility of the nitroxide side chain, it is necessary to: (1) remove the contributions from the rotational diffusion of the protein and (2) to understand the contribution from the internal motions of the nitroxide side chain. The contribution from τ_r is negligible for membrane proteins and soluble proteins with Mw ≥ 50 kDa. For small proteins, the contributions from τ_r can be usually minimized by increasing the effective viscosity of the solution [17]. Thus for most cases, only τ_{bb} and τ_{int} contribute to τ_T. During the last few years there has been a considerable effort to characterize the internal motions of the nitroxide side chain in soluble [19,20,37] and membrane proteins [52]. These studies have focused on the most widely used R1 side chain shown in Figure 2.19 and reintroduced in panel A of Figure 2.29. In principle, the internal motions of the nitroxide side chain can arise from rotation about its 5 dihedral angles (denoted $X_1 - X_5$ in Figure 2.29), however high resolution structures of spin-labeled proteins in combination with EPR spectroscopy and mutagenesis studies [17-18] have shown that the internal motions of R1 in well-ordered α-helical and loop surface sites are constrained due to *intra*-residue S_δ•••H-C_α interaction. This interaction specifically constrains the motion about X_1 and X_2, and together with the high energy barrier for rotation about X_3 [53], leads to a model for internal motion of R1 known as the "X_4/X_5" model. In this model, rotation about X_4/X_5 dihedral angles gives rise to anisotropic motion of the nitroxide 2p

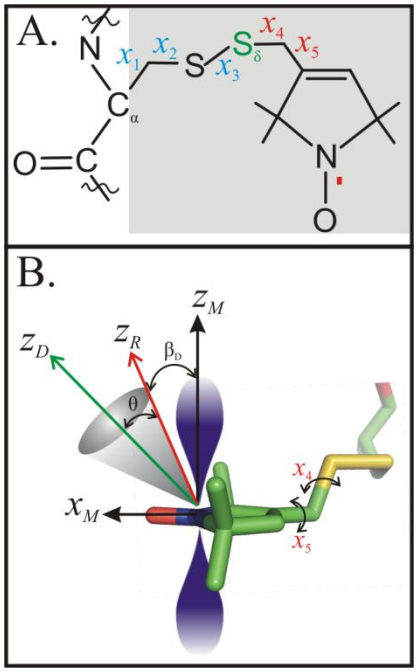


Figure 2.29 R1 side chain and a model of its motion on solvent exposed sites. A) Structure of R1 side chain. The dihedral angles described in the text are shown. B) Coordinate system used to describe the motion of the nitroxide using the MOMD model.

orbital within a cone (Figure 2.29B) that can be described by an order parameter (S) and a correlation time (τ) using the MOMD model described in section 2.7(3). The S_δ•••H-C_α interaction has been observed in many high-resolution structures of spin-labeled proteins [21,30,37], including membrane-bound proteins [52], which suggest that this interaction is a general structural feature that determines the internal motion of the R1 side chain for solvent exposed sites in α-helices and well-ordered loops. *Thus, it is expected that for such sites where the R1 side chain lacks any tertiary interaction, τ_{int} should be constant and independent of position, therefore variations in the order parameter (S) and correlation time (τ) among structurally homologous sites should reflect contributions from local backbone fluctuations.*

The structural origins for the motion of the R1 side chain in β-sheets are less well understood, however SDSL in combination with mutagenesis studies done in cellular retinol-binding protein (CRBP) suggest a model that accounts for the observed motions of R1 in β-sheets [39]. In this model the R1 side chain lacks the S_δ•••H-C_α *intra*-residue interaction observed in α-helices due to steric clashes with neighboring site chains, but its motion is strongly modulated by *inter*-residue interaction. The validity of the model remains to be corroborated *via* determination of high-resolution structures of a β-sheets protein bearing the R1 side chain.

To validate the hypothesis that R1 in solvent exposed sites of α-helices is sensitive to fast backbone motions, Columbus et al. [54] studied the site-dependent variation of the EPR spectra of R1 along the basic leucine zipper (bZip) motif of the yeast transcription factor GCN4. GCN4 was selected as a model system because NMR ^{15}N relaxation experiments identified a gradient of backbone motions on the nanosecond to picoseconds time scale along the DNA binding region [55]. To compare the NMR results with EPR spectroscopy, the R1 side chain was introduced along the same sequence and the mobility of the nitroxide side chain was analyzed in terms of a

normalized version of the inverse of the central line width known as scaled mobility (Ms) and in terms of the MOMD model. Columbus observed a linear gradient in the mobility measured by Ms and rate of motion of R1 (τ^{-1}) along the basic region that mirrors the results observed by NMR studies (Figure 2.30B). In addition, SDSL studies of GCN4 bound to DNA revealed a dampening of the motion of R1 along the same sequence [54]. These results support the view that R1 is a useful sensor to detect backbone motions on the nanosecond to picoseconds time scale. The amplitude of the backbone motions in the basic region of GCN4 was large enough to preclude measurements of S as a function of sequence using R1. To evaluate whether variations of both S and τ reflect contribution of backbone motions to the mobility of R1 in well-ordered protein sequences, R1 was placed at 39 solvent-exposed sites of sperm whale myoglobin and the mobility of R1 was analyzed in terms of scaled mobility (Ms) and the MOMD model as part of this dissertation. The results of these studies are presented in Chapter 4.

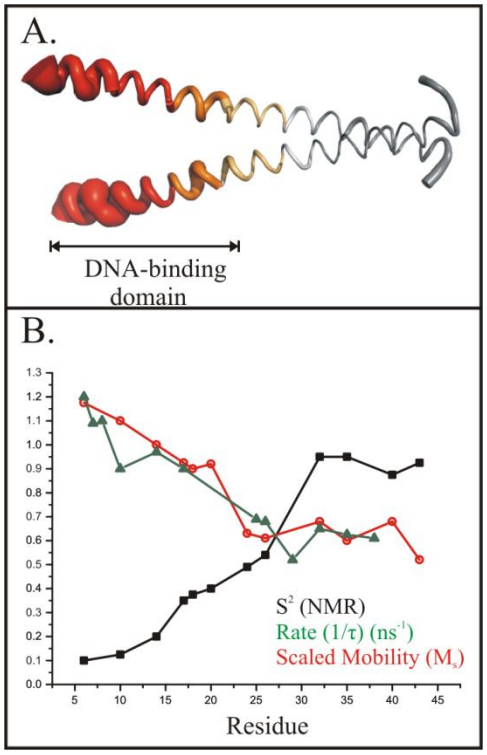


Figure 2.30 Study of fast backbone motions in GCN4 by SDSL. A) Ribbon diagram of the basic leucine zipper motif of GCN4. The width of the backbone is proportional to the mobility as detected by NMR and EPR. B) Plot of the order parameter derived from NMR experiments and rate of motion and scaled mobility of the nitroxide detected by EPR. Strategy to calculate scaled mobility values can be found in Columbus et al. [54]. Individual data points taken from [54-55].

2.8.5 Identifying slow (τ_{ex} > 100 ns) conformational exchange with SDSL. Functionally important protein motions taking place in μs-ms time scale are too slow to produce T_2 relaxation, thus such motions do not contribute directly to the EPR spectra lineshape. As mentioned above, the EPR spectra of spin-labeled proteins is determined by the motion of the nitroxide on the nanosecond time scale, which in turn is determined by it local environment. Since the spin label

is a sensor of the local environment as shown in Figure 2.25A, conformational exchange can be revealed in SDSL in the following two scenarios. In the first scenario, consider a spin-labeled protein that changes its conformation from substate 1 to substate 2 in response to a physical or chemical signal (*i.e.*, substrate binding, phosphorylation, protein-protein interaction) as depicted in Figure 2.31. If the R1 side chain is in a distinct environment in each state, the EPR spectra recorded before and after the signal will reflect changes in the mobility of the nitroxide side chain in the nanosecond time scale as shown in Figure 2.31B, thereby revealing the structural changes in the protein.

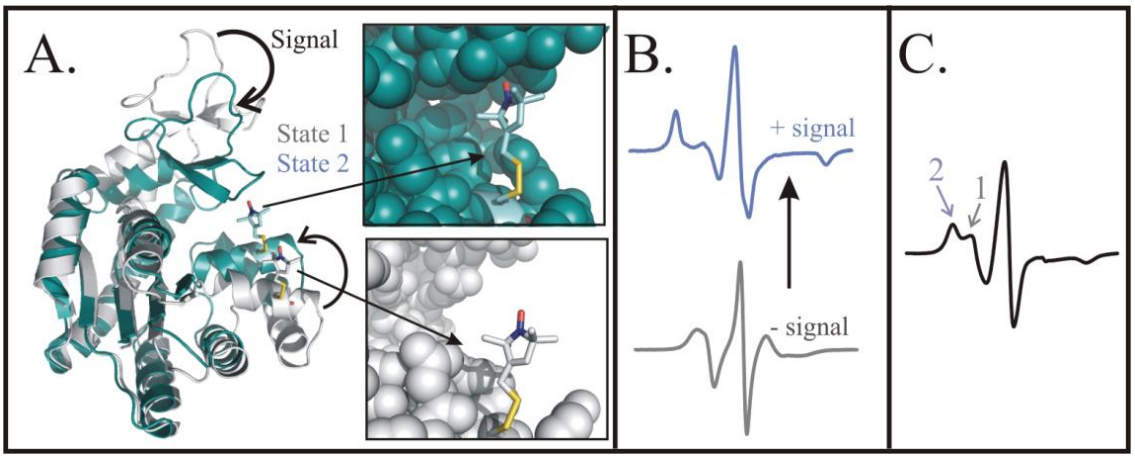


Figure 2.31 Identifying conformational exchange in spin-labeled proteins. A) Ribbon diagram of an arbitrary spin-labeled protein undergoing slow conformational exchange between 2 substates. Each substate is color coded for clarity. The R1 side chain was placed at an arbitrary position that would be sensitive to changes in the local environment. *Inset*: CPK model of the local environment around the R1 side chain on each state. B) In the absence of a pre-existing equilibrium, the EPR spectra of the spin-labeled protein recorded before and after the signal will reflect the local environment of each state. C) EPR spectra observed for a protein in pre-existing equilibrium between multiple states. The two components reflecting each environment are indicated.

In the second scenario, the spin-labeled protein is in a pre-existing equilibrium between substates 1 and 2 and the exchange between the states is slow on the EPR time scale. In this case, the fluctuation gives rise to an EPR spectrum that will be a weighted sum of the spectra shown in panel B of Figure 2.31, with intensities proportional to the relative population of each state as

depicted in Figure 2.31C. Such spectra are known as multicomponent or complex spectra. However, crystal structures of spin-labeled protein [30,37-38] have shown that multicomponent spectra can also arise from the existence of multiple rotamers of the R1 side chain. Thus, to use SDSL as a tool to identify protein conformational exchange in an equilibrium mixture, it is essential to develop experimental strategies to distinguish protein conformational equilibria from R1 rotameric equilibria. In this dissertation, it is shown that these scenarios can be distinguished by the response of the protein to osmolyte-perturbation. The results of this study are presented in the next chapter.

2.9 References

1. Commoner, B., Townsend, J., and Pake, G.E. 1954. Free radicals in biological materials. *Nature*. 174:689-691.
2. Stone, T.J., Buckman, T., Nordio, P.L., and McConnell, H.M. 1965. Spin-labeled biomacromolecules. *Proc. Natl. Acad. Sci.* 54:1010-1017.
3. Todd, P.A., Cong, J., Levinthal, C., and Hubbell, W.L. 1989. Site-directed mutagenesis of Colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. *Proteins J*. 6:294-304.
4. Rousseau, D.L. 1984. *Structural and resonance techniques in biological research.* Academic Press. London ; New York.
5. Augusto, O. and Vaz, S.M 2007. EPR spin-trapping of protein radicals to investigate biological oxidative mechanism. *Amino Acids* 32:535-542.
6. Eaton, G.R., Eaton, S.S., Barr, D.P., and Weber, R.T. 2010. *Quantitative EPR*. Springer-Verlag/Wien. Germany : New York.

7. Carrington, A., and McLachlan, A.D. 1967. *Introduction to magnetic resonance with application to chemistry and chemical physics*. Harper & Row. New York.
8. Marsh, D. 1981. Electron spin resonance: Spin labels. Grell, E. *Molecular Biology, Biochemistry, and Biophysics: Membrane Spectroscopy*. (pp. 51-142) Springer-Verlag. Berlin
9. Poole, C.P. 1983. *Electron spin resonance: A Comprehensive Treatise on Experimental Techniques*. Dover publications. Mineola, New York.
10. Levine, I.N. 2002. *Physical Chemistry*. Fifth edition. McGraw Hill: New York.
11. F. Block. 1946. *Phys. Rev*. 70: 460-527.
12. Humphries, G.M.K., and McConnell H.M. 1982. Nitroxide spin labels. *Methods of Experimental Physics*. 20:53-109.
13. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., and Skelton, N.J. 1996. *Protein NMR spectroscopy: Principles and practices*. Academic Press. New York.
14. Fermi, E. 1930. "Uber die magnetischen Momente der AtomKerne" *Z. Phys*. 60:320-333.
15. Griffith, O. H., Cornell, D.W. McConnell, H.M. 1965. Nitrogen hyperfine tensor and g tensor of nitroxide radicals. *J. Chem. Phys* 43: 2909-2910.
16. Burr, M., and Koshland, D.E. 1964. Use of "reporter groups" in structure-function studies in proteins. *Proc. Natl. Acad. Sci.* 52:1017-1024.
17. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35:7692-7704.
18. Mchaourab, H.S., Kálai, T., Hideg, K., and Hubbell, W.L. 1999. Motion of spin-labeled side chains in T4 lysozyme. Effect of side chain structure. *Biochemistry* 38:2947-2955.

19. Columbus, L., Kálai, T., Jeko, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in alpha-helices:analysis by variation of side chain structure. *Biochemistry* 40:3828-3846.
20. Fleissner, M.R. 2007. X-ray structures of nitroxide side chains in proteins: a basis for interpreting distance measurements and dynamic studies by electron paramagnetic resonance. PhD thesis, University of California, Los Angeles.
21. Fleissner, M.R., Cascio, D., and Hubbell, W.L. 2009. Structural origins of weakly ordered motion in spin-labeled proteins. *Prot. Sci.*. 18:893-908.
22. Bridges, M.D. Hideg, K., and Hubbell, W.L. 2010. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson*. 37:363-390.
23. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011 Structure and dynamics of a conformationaly constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246
24. Redfield, A.G. 1965. The theory of relaxation processes. *Adv. Magn. Reson.* 1:1-32.
25. Knowles, P.F., Marsh, D., and Rattle, H.W.E. 1976. *Magnetic Resonance of Biomolecules: An introduction of the theory and practice of NMR and ESR in biological systems*. John Wiley and Sons. London: New York.
26. Freed, J.H. 1976. Theory of Slow Tumbling ESR Spectra for Nitroxides: *Spin Labeling Theory and Applications*, L.J. Berliner, ed. Academic Press, New York (pp. 53-132).
27. Schneider, D.J., and Freed, J.H. 1989. *Spin labeling theory and applications*, eds. L.J. Berliner, J. Reuben, Plenum Press, New York (pp 1-76).

28. Budil, D.E., Lee, S., Saxena, S., and Freed, J.H. 1996. Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquard algorithm. *J. Magn. Reson. Series A* 120:155-189.

29. Earle, K.A., and Budil, D.E. 2006. Calculating slow-motion ESR spectra of spin-labeled polymers: *Advance ESR methods in polymer research*. Schlich S. ed. John Wiley and Sons. London: New York. (pp 53-83).

30. Gou, Z., Casio, D., Hideg, K., and Hubbell, W.L. 2008. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Prot. Sci.* 17:228-239.

31. Columbus, L and Hubbell, W.L. 2002. A new spin on protein dynamics. *TiBS* 27:288-295.

32. McCoy, J. and Hubbell, W.L. 2011. High-pressure EPR reveals conformational equilibria and volumetric properties in spin-labeled proteins. *Proc. Natl. Acad. Sci.* 108:1331-1336.

33. Pyka, J., Ilnicki, J., Altenbach, C., Hubbell, W.L., and Froncisz, W. 2005. Accessibility and dynamics of nitroxide side chains in T4 lysozyme measured by saturation recovery EPR. *Biophys. J*. 89:2059-2068.

34. Altenbach, C., Froncisz, W., Hemker, R., Mchaourab, H., and Hubbell, W.L. 2005. Accessibility to nitroxide side chains: absolute Heisenberg exchange rates from power saturation EPR. *Biophys. J.* 89:2103-2112.

35. Oh, K.J., Zhan, H., Cui, C., Hideg, K., Collier, R.J., and Hubbell, W.L. 1996. Organized of diphtheria toxin T domain in bilayers: a site-directed spin labeling study. *Science* 273:810-812.

36. Hubbell, W.L., Mchaourab, H., Altenbach, C., and Lietzow, M.A. 1996. Watching proteins move using site-directed spin labeling. *Structure* 4:779-783.

37. Langen, R., Oh, K.J., Cascio, D., and Hubbell, W.L. 2000. Crystal structures of spin labeled T4 lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry*, 39:8396-8405.

38. Guo, Z., Cascio, D., Hideg, K., Kálai, T., and Hubbell W.L. 2007. Structural determinants of nitroxide motion in spin-labeled proteins: Tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. *Prot. Sci.* 16:1069-1086.

39. Lietzow, M.A., and Hubbell, W.L. 2004. Motion of spin label side chains in cellular retinol-binding protein: Correlation with structure and nearest-neighbor interactions in an antiparallel β–sheet. *Biochemistry* 43:3137-3151.

40. Fiori, W.R., Miick, S.M., and Millhauser, G.L. 1993. Increasing sequence length favors α-helix over 3_{10}-helix in alanine-based peptides: Evidence for a length-dependent structural transition. *Biochemistry* 32:11957-11962.

41. Altenbach C., Oh, K.J., Trabanino, R.J., Hideg, K., and Hubbell, W.L. 2001. Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry* 40:15471-15482.

42. Altenbach, C., Kusnetzow, A.K., Ernst, O., Hofmann, K.P., and Hubbell, W.L. 2008. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci.* 105:7439-7444.

43. Fajer, P.G. 2000. Electron spin resonance spectroscopy labeling in peptides and protein analysis. *Encyclopedia of Analytical Chemistry*. Meyers, R.A. ed. John Wiley & Sons, Chichester. (pp 5725-5761).

44. Mchaourab, H., Oh, K.J., Fang, C.J., and Hubbell, W.L. 1997. Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition stidued by site-directed spi labeling. *Biochemistry* 36:307-316.

45. Dixon, M.M., Nicholson, H., Shewchuk, L., Baase, W., and Matthews, B.W. 1992. Structure of a hinge-bending bacteriophage T4 lysozyme mutant, Ile3-->Pro. *J. Mol. Biol*. 227:917-933.

46. Borbat, P.P, and Freed, J.H. 2007. Measuring distances by pulse dipolar ESR spectroscopy: Spin-labeled histidines kinase. *Methods in Enzymology* 423:52-116.

47. Jeschke, G., and Polyhash, Y. 2007. Distance measurements on spin-labeled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys*. 9:1895-1910.

48. Borbat, P.P., Davis, J.H., Butcher, S.E., and Freed, J.H. 2004. Measurement of large distance in biomacromolecules using double-quantum filtered refocused electron spin echoes. *J. Am. Chem. Soc. Comm*. 126:7746-7747.

49. Fleissner, M.R., Brustad, E.M., Kálai, T., Altenbach, C., Schultz, P.G., and Hubbell, W.L. 2009. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci.* 106:21637-21642.

50. Pannier, M., Veit, S., Godt, A., Jeschke., G., and Speiss, H.W. 2000. Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn. Reson*. 142:331-340.

51. Jeschke G. 2002. Distance measurements in the nanometer range by pulse EPR. *Chem. Phys. Chem*. 3:927-832.

52. Kroncke, B.M., Horanyi, P.S., and Columbus, L. 2010. Structural origins of nitroxide side chain dynamics on membrane protein α-helical sites. *Biochemistry* 49:10045-10060.
53. Fraser, R.R., Boussard, G., Saunders, J. K., Lambert, J.B., and Mixan, C.E. 1971. Barriers to rotation about sulfur-sulfur bond in acyclic disulfides. *J. Am. Chem. Soc.* 93:3822.
54. Columbus, L., and Hubbell., W.L. 2004. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* 43:7273-7287.
55. Bracken, C., Carr, P.A. Cavanagh, J., and Palmer, A.G. 1999. Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA. *J. Mol. Biol*. 285:2133-2146.

Chapter 3: Osmolyte-perturbation reveals conformational equilibria in spin-labeled proteins

3.1 Summary

Recent evidence suggests that proteins at equilibrium can exist in a manifold of conformational substates, and that these substates play important roles in protein function. Therefore, there is great interest in identifying regions in proteins that are in conformational exchange. Electron paramagnetic resonance (EPR) spectra of spin labeled proteins containing the nitroxide side chain (R1) often consist of two (or more) components that may arise from slow exchange between conformational substates (lifetimes >100 ns). However, crystal structures of proteins containing R1 have shown that multicomponent spectra can also arise from equilibria between rotamers of the side chain itself. In this report, it is shown that these scenarios can be distinguished by the response of the system to solvent perturbation with stabilizing osmolytes such as sucrose. Thus, site directed spin labeling (SDSL) emerges as a new tool to explore slow conformational exchange in proteins of arbitrary size, including membrane proteins in a native-like environment. Moreover, equilibrium between substates with even modest differences in conformation is revealed, and the simplicity of the method makes it suitable for facile screening of multiple proteins. Together with previously developed strategies for monitoring ps-ns backbone dynamics, the results presented here expand the time scale over which SDSL can be used to explore protein flexibility.

3.2 Introduction

It is now clear that structural fluctuations occurring on the picosecond to millisecond time scale are intimately involved in function, save for proteins with a purely structural role (*see* Chapter 1 of this dissertation). The elucidation of the molecular mechanisms of protein function thus requires information on protein motions taking place over a wide range of time and length scales, and it is essential to develop experimental strategies that allow one to observe both fast (nanosecond) backbone motions and conformational exchange (microsecond to millisecond) in systems at equilibrium. NMR spectroscopy has been particularly effective in this regard in studies of small proteins in solution [1], but equivalent measurements on high molecular weight soluble proteins, membrane bound proteins, and transient complexes between them remain challenging, particularly when it is desired to screen a large number of such proteins.

Site directed spin labeling (SDSL) has the potential to provide information on protein motions occurring on the above time scales without limitations regarding the size or complexity of the system. Moreover, soluble or membrane proteins can be investigated under strictly physiological conditions in a native-like environment. In SDSL, a nitroxide side chain is introduced in a site specific manner. The most widely used is that designated R1, although many others have been developed [2,3]. Motions of the R1 nitroxide group on the ps-ns time scale result in magnetic relaxation that determines the EPR spectrum. Thus, backbone fluctuations that occur on this time scale can contribute directly to the EPR spectrum, and SDSL has been shown to provide a measure of fast backbone dynamics [4,5]. This capability should be particularly valuable for locating dynamically disordered domains within proteins, and for identifying natively unfolded proteins and monitoring their interactions with binding partners [6,7].

While fast backbone motions are directly revealed in the EPR spectrum, detection of conformational exchange on the μs-ms time scale requires a fundamentally different approach, because such motions are too slow to produce relaxation effects that reveal themselves in the spectra. Rather, equilibria between *conformational* substates in exchange on this time scale are revealed as multicomponent ("complex") EPR spectra, provided that R1 is in a region where it has distinct interactions with the local environment in the different substates as depicted in Figure 3.1A. Examples of complex spectra and their analysis in terms of a two-state conformational equilibrium model are given in Kusnetzow *et al.* [8] and Crane *et al.* [9].

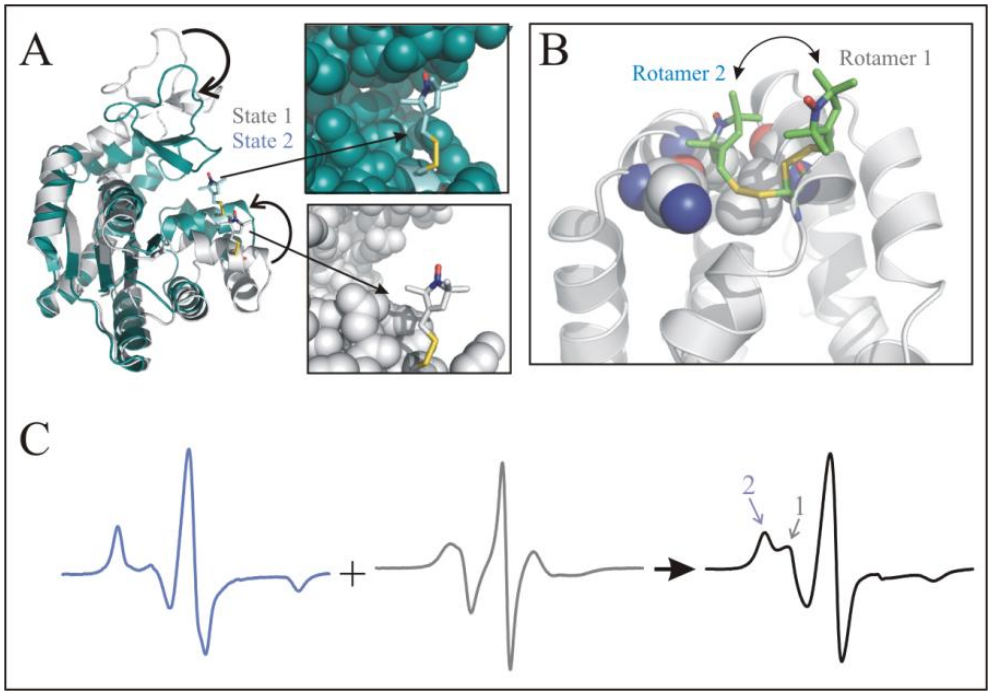


Figure 3.1: Origins of complex EPR spectra. A) Conformational equilibria of an arbitrary protein between two conformational substates. Each substate is color coded for clarity. *Insets*: CPK model of each conformational substate showing the environment around the R1 side chain. The R1 side chain in the state 2 is in tertiary interaction with the protein environment. B) Ribbon diagram of T4 lysozyme mutant T115R1 (pdb: 2ou8) showing rotameric equilibrium of R1. Rotamer 2 is in tertiary interaction with residues shown in CPK. C) The resulting complex EPR spectra from R1 rotameric or conformational equilibria is a weighted sum of the individual spectrum from each R1 or protein state.

Combined X-ray crystallography, mutagenesis and EPR studies revealed that complex spectra can also arise from the existence of multiple rotamers of R1 at a given site [10-12] (Figure 3.1B). As shown in Figure 3.1C, the complex EPR spectrum is a weighted sum of the spectra corresponding to each rotameric or conformational state. Therefore, to use SDSL as a tool to positively identify conformational exchange in an equilibrium mixture, it is essential to develop experimental strategies that can be used to distinguish protein conformational equilibria from R1 rotameric equilibria as origins of complex spectra.

One approach that can be used for this purpose is to measure the response of the spin labeled protein to solvent perturbation by osmolytes, which are small organic molecules that either stabilize or destabilize proteins with respect to unfolding. Protecting osmolytes increase the stability of proteins, and include carbohydrates, free amino acids and their derivatives, and methylamines [13,14]. It is well established that protecting osmolytes are excluded from the surface of proteins (preferential hydration), thereby raising the chemical potential of the protein in relation to the area exposed to solvent [15] (Figure 3.2A), although the relationship may not be simple [16]. Because the unfolded state has a larger area in contact with solvent than the native state, it is destabilized relative to the native state. On the other hand, denaturing osmolytes (e.g. urea) are preferentially accumulated at the surface and have the opposite effect [17].

The molecular origin of the osmolyte effect on protein stability appears to be a result of interaction of the osmolyte with the peptide backbone, which is unfavorable (relative to water) for protecting osmolytes and favorable for denaturing osmolytes [18,19]. It follows that osmolytes should also drive folding of partially unfolded states of proteins (Figure 3.2B), and in

fact, such solutes drive a largely unfolded sequence in the transporter BtuB to a folded state [20-21]. In addition, recent studies have also shown that osmolytes attenuate structural fluctuations of proteins in their native state [22-25].

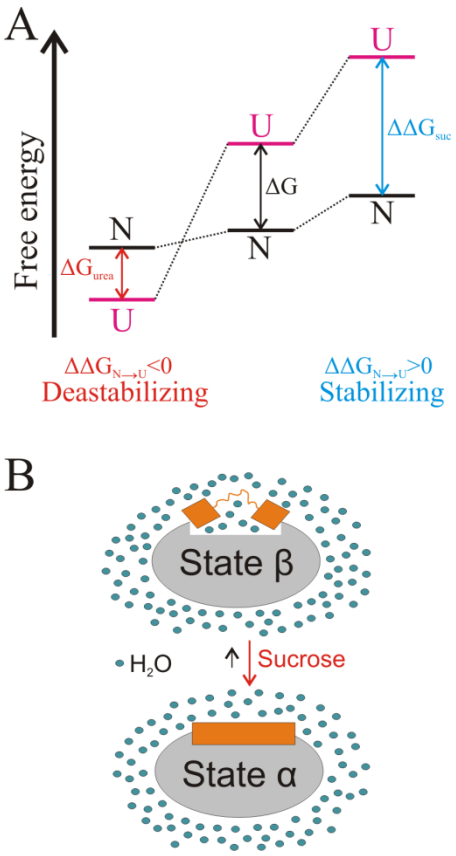


Figure 3.2: Effect of osmolytes of the conformational equilibria between substates with different solvent-exposed surface areas. A) Relative free energy of the native (N) and unfolded (U) states in the presence of destabilizing/stabilizing osmolytes. B) Effect of stabilizing osmolyte such as sucrose on the conformational equilibrium between a partially folded state (β) and a more compact folded state (α).

In general, different conformational substates of a protein are expected to have different solvent accessible surface areas, even if the structural differences are more subtle than those between the folded and unfolded states. If so, osmolytes should shift the equilibrium among substates and have measurable effects on complex EPR spectra of R1. In contrast, complex spectra arising solely from rotameric equilibria of R1 are expected to be relatively insensitive to osmolytes due to the small difference, if any, in solvent exposed area of the protein for different rotamers. Thus, the response of multicomponent spectra to osmolyte perturbation may provide the desired tool to distinguish R1 rotameric equilibria from protein conformational equilibria; this study evaluates this proposal. The experimental strategy is to examine the effect of osmolyte perturbation on multicomponent EPR spectra of R1 located at sites where the origin can be assigned with confidence to side chain rotamers, and at sites in proteins known to be in

conformational exchange. As will be discussed below, examples of these situations are found in T4 lysozyme (T4L), rat intestinal fatty acid-binding protein (rIFABP), and apomyoglobin (Myb).

3.3 Results

3.3.1 Separation of Osmotic from Viscosity Effects. To examine the osmolyte effect by SDSL, sucrose was selected as a well-characterized stabilizing osmolyte known to be preferentially excluded from the protein surface [26,27], and known to increase stability and reduce conformational fluctuations in proteins [23-25]. A complication arises in assigning observed effects of sucrose to osmolarity, because sucrose also increases the viscosity of the solution, slowing the overall rotational diffusion rate of the protein. For small proteins (MW<50 kDa) the rotary motion contributes substantially to nitroxide magnetic relaxation and hence the X-band EPR spectrum. Ideally, one would want to compare the EPR spectrum in sucrose solution with that of the same sample obtained in a reference medium of the same effective viscosity, but with no osmolarity.

Solutions of Ficoll 70 are of high viscosity but low osmolarity, and should serve as a suitable reference medium. Moreover, Ficoll 70 appears to be inert with respect to protein interactions [28-29]. For the purpose of this study, the effective viscosities of Ficoll 70 and sucrose solutions are considered to be matched at concentrations that result in equal protein rotational correlation times measured by a spin label rigidly attached to a well ordered protein (*i.e.,* a spin label with no motion relative to the protein). Previous studies have shown that the spin label 1 – oxyl -2,5,5 - trimethyl - 2,4 diphenyl - 2,5 - dihydro- 1*H* – pyrrol – 3 - ylmethyl methanethiosulfonate shown in Figure 3.3A give a nitroxide side chain (denoted R8 hereon) that

is fully immobilized with respect to the protein when reacted with a surface cysteine residue [30]. Thus, this side chain was introduced at site D72 in T4L which is in a well-ordered region of the protein with minimal contributions from local backbone fluctuations [3]. To a good approximation, the EPR spectrum of 72R8 is determined by the rotational motion of the protein as a whole. For EPR spectra of a nitroxide whose motion is close to the rigid limit, such as the R8 side chain, the rotational correlation time (τ_R) can be estimated as:

$$\tau_R = a\left(1 - \frac{A_{zz}'}{A_{zz}^{R}}\right)^{b}, \qquad (3.1)$$

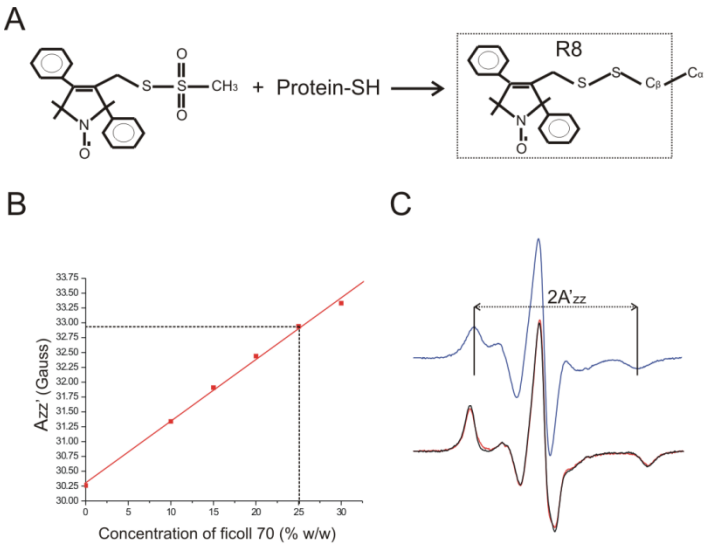


Figure 3.3: Calibration of the effective viscosity of Ficoll solutions. A) Structure of the R8 side chain used to estimate the rotary diffusion rate of T4L. B) Plot of hyperfine splitting (A_{zz}') values for the side chain at site 72 in T4L at different concentrations of Ficoll 70. The black dashed lines mark the A_{zz}' determined for a solution containing 30% w/w sucrose and the concentration of Ficoll 70 that gives the same value, and hence the same rotational diffusion rate for T4L C) EPR spectra of the side chain at site 72 in T4L compared in 30% w/w sucrose (black) and 25% w/w Ficoll 70 (red). The EPR spectrum of 72R8 in buffer (blue) is shown for comparison.

where A_{zz}' is the observed outer hyperfine splitting, A_{zz}^{R} is the value of hyperfine splitting in the absence of protein rotational motion, and a and b are constants [31]. Thus, τ_R is measured by A_{zz}' and the concentration of Ficoll 70 that gives the same A_{zz}' as a 30% w/w sucrose solution at 295 K has an equal effective viscosity with respect to protein rotary diffusion. EPR spectra were recorded in the presence of 30% w/w sucrose (~0.988M) and as a function of Ficoll 70 in the range of 0 – 30 % w/w (Figure

3.3B). The hyperfine splittings values were determined by individually fitting the low field and the high field peaks (*see* Figure 3.3C) to a mixture of gaussian and lorentzian lineshapes using the Xepr program (Bruker, Germany) and by measuring the magnetic field separation between the low field and high field peaks to get $2A_{zz}'$. The hyperfine splitting A_{zz}' in 30% w/w sucrose was 32.92 ± 0.02 G. A plot of A_{zz}' vs Ficoll 70 concentration (Figure 3.3B) showed that a solution containing 25% w/w Ficoll 70 has a hyperfine splitting value of 32.94 ± 0.02 G, identical to that of sucrose within the experimental error. An overlay of the EPR spectra of 72R8 in 30% w/w sucrose and 25% w/w Ficoll 70 shows them to be essentially identical (Figure 3.3C). For the purpose of this study it should be pointed out that the other soluble proteins studied here (e.g., myoglobin and rIFABP) have a similar size and shape compared to T4L, thus the same concentration of Ficoll 70 can be used to match effective viscosities to that in sucrose solution for all cases. Thus, osmotic effects alone due to sucrose can be isolated by comparing the EPR spectrum of an R1 mutant in a 30% sucrose solution to that in a 25% Ficoll 70 solution.

Although Ficoll 70 is a suitable viscosity agent, it can also act as a crowding agent favoring the states of a protein with the smallest excluded volume [32-34]. Such effects are generally detected *via* stabilization of the protein against denaturation and *via* changes in the secondary structure [33], the effect being due to the larger excluded volume of the denatured state relative to the native state [35]. To determine if Ficoll has significant crowding effects on T4L, apomyoglobin, and apo-rIFABP (MW<20 kDa), denaturation studies were carried out in buffer, in sucrose, and in Ficoll (Figure 3.4). In addition, changes in the secondary structures were monitored via far-UV CD spectroscopy (Figure 3.4). Thermal denaturation of T4L and apomyoglobin, and guanidine denaturation of apo-rIFABP were essentially identical with and

without 25% Ficoll. On the other hand, 30% sucrose produced substantial increases in the stability of each protein (Figure 3.4A-C), as expected for a protecting osmolyte. Addition of Ficoll 70 did not induce any changes in the secondary structures of T4L, apo-rIFABP, and apomyoglobin (Figure 3.4D-F). The above results demonstrate that crowding effects of Ficoll can be ignored in the systems and sites studied here. Further evidence showing a lack of crowding effects from Ficoll 70 will be provided below with respect to apomyoglobin.

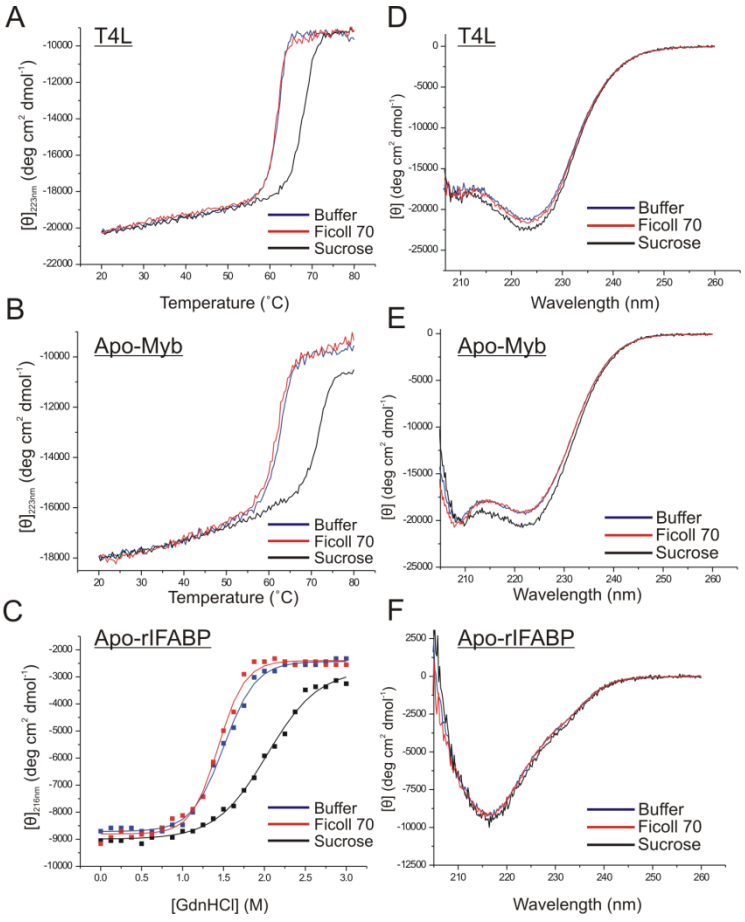


Figure 3.4: Effect of sucrose and Ficoll 70 on the stabilities and secondary structures of T4L WT*, apomyoglobin WT, and apo-rIFABP WT. A) Thermal denaturation of T4L. B) Thermal denaturation of apomyoglobin. C) Chemical denaturation (GdnHCl) of apo-rIFABP. The chemical denaturation data were fit using a linear extrapolation method of Santoro and Bolen [36]. D) Far-UV CD of T4L. E) Far-UV CD of apomyoglobin. F) Far-UV CD of apo-rIFABP. In each plot, data from buffer (blue), 30% w/w sucrose (black), and 25% Ficoll 70 (red) are compared. WT* is T4L (C54T/C97A).

3.3.2 Contribution of Protein Rotary Diffusion to Nitroxide Motion in T4L and the Effect of Sucrose and Ficoll on R1 Internal Motion. In general, at non-interacting, solvent-exposed helical sites (i.e., where interactions of the nitroxide with the protein are absent), the internal motion of R1 is weakly ordered, giving rise to single-component EPR spectra characteristic of z-axis anisotropic diffusion [3,37]. The motion of the nitroxide in this state has contributions from R1 internal motions and the overall rotary diffusion of the protein. Increasing the viscosity of the solution with sucrose or Ficoll reduces the contribution from protein rotary diffusion, but is not expected to influence the internal motion of R1 [38,39]. To demonstrate this point, sites 72 and 131 in T4L were selected for introduction of R1. Both sites are located at solvent exposed surfaces of helices and both have single component EPR spectra reflecting the weakly ordered state [3,37,39]. Moreover, both are located in a stable protein fold with no evidence for multiple conformational substates in equilibrium on the μs-ms time scale [40], although low amplitude backbone fluctuations on the ps-ns time scale might exist [4].

The EPR spectra of 72R1 and 131R1 in buffer alone are shown in Figure 3.5 (row a), and are similar to those published previously [41]. For each mutant in either sucrose (row b) or Ficoll (row c), the spectra clearly reflect the change in motion due to attenuation of rotary diffusion. In order to completely eliminate the protein rotary diffusion, two immobilization strategies were utilized, both of which generate chemical crosslinks *via* native lysine residues. There are 13 native lysines in T4L, and multiple points of attachment and strong immobilization of the protein are anticipated, as has been observed in a number of proteins [42]. In one strategy, the protein was covalently attached to a CNBr modified Sepharose solid support, which occurs by reaction of ε-amino groups of lysine with the activated substrate to form a Sepharose-bound isourea [43];

this method has been used previously to immobilize T4L for SDSL studies [44]. In the glutaraldehyde crosslinking strategy, a Schiff's base is formed between ε-amino groups of lysine and aldehyde functions on monomeric and polymeric forms of glutaraldehyde [45]; the reaction has been used to immobilize proteins for NMR [46] and EPR studies [47]. The EPR spectra of 72R1 and 131R1 immobilized by these methods are shown in Figure 3.5 (rows d and e). In each case, a small spectral component corresponding to an immobilized state (arrow) was observed. This is very likely due to a fraction of the nitroxide being immobilized by direct interaction with the matrix near a site of attachment with Sepharose, or with another T4L molecule near a site of crosslinking in the case of glutaraldehyde. Replacing K135 near 131R1 with alanine reduces the amount of the immobilized component, supporting the above interpretation (*see* legend of Figure 3.5).

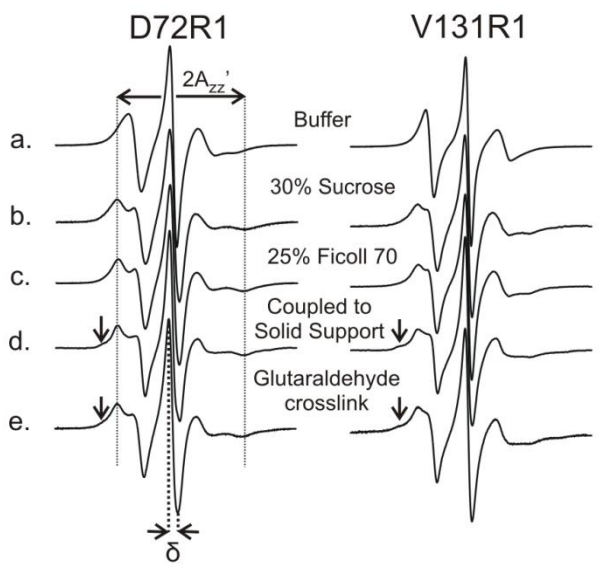


Figure 3.5: Contribution of protein rotational diffusion to the motion of 72R1 and 131R1. EPR spectra of residues 72R1 and 131R1 in T4L recorded under the indicated conditions. The central line width (δ) and the splitting of the resolved hyperfine extrema ($2A_{zz}'$) are defined on the spectra of 72R1. The arrows identify a minor spectral component. In the case of 131R1, neighboring lysine K135 was replaced with alanine for the spectra in (d) and (e) to avoid coupling near the nitroxide side chain. Previous studies have shown that this *i*+4 mutation had little effect on the EPR spectrum of 131R1 in solution [39].

As shown in Figure 3.5, the EPR spectra of T4L 72R1 in sucrose and in Ficoll 70 are indistinguishable from one another and from the major components of the spectra for the proteins immobilized on Sepharose or by crosslinking; the same is true for 131R1. The inverse of the central line width (δ^{-1}) and the splitting of the hyperfine extrema ($2A_{zz}'$) indicated in Figure 3.5

are measures of nitroxide mobility [5]. For either residue, the maximum variation in these quantities is only about 15% for δ^{-1} and 4% for A_{zz}'.

To evaluate if either sucrose or Ficoll 70 influences the internal motions of R1, the spectra of residue 72R1 attached to the CNBr modified Sepharose support were recorded over a range of sucrose concentrations (0-30%) and with 25% Ficoll 70. Figure 3.6 shows that neither sucrose (up to 30%) nor Ficoll (25%) have a substantial influence on the internal motion of R1; δ^{-1} and A_{zz}' between buffer and 30% sucrose differ by only 8% and 2%, respectively. The difference between buffer and 25% Ficoll is negligible. These results demonstrate that: (i) 30% sucrose and 25% Ficoll indeed remove the effect of protein rotary diffusion in a similar fashion and (ii) neither sucrose nor Ficoll substantially influence the internal motion of the R1 side chain.

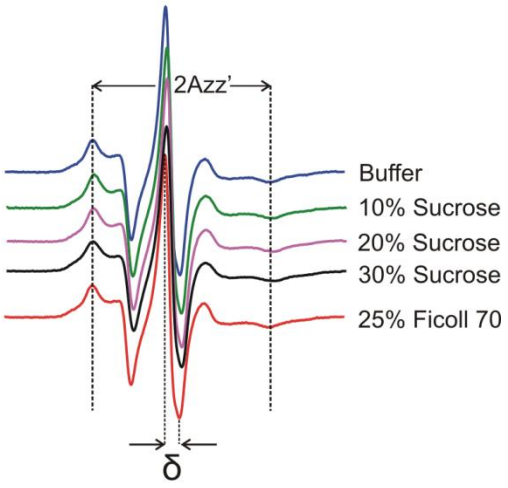


Figure 3.6: Effect of sucrose and Ficoll 70 on the internal motions of R1 at residue 72R1 of T4L. EPR spectra of 72R1 attached to the Sepharose solid suppport in buffer and the indicated concentrations of sucrose and Ficoll 70.

To further evaluate if R1 internal modes are largely unaffected by osmotic perturbation in stable protein folds (i.e., those without conformational fluctuations), the motion of R1 was explored at additional solvent-exposed and partially buried sites in T4L, where the nitroxide exhibits a single dynamic mode (teal spheres in Figure 3.7A). The EPR spectra of some of these mutants in 30% sucrose were previously reported [11,12,37,39] and the spectra obtained in this study are similar. Panels B and C of Figure 3.7 show a comparison of the spectra for each mutant

in Ficoll and in sucrose, and it is evident that there are few differences in the spectra of R1 (if any at all). These mutants represent a range of protein local structures and side chain dynamics. The results show that the internal motion of R1 is not affected by osmolyte perturbation or microscopic viscosity in the presence of osmolytes.

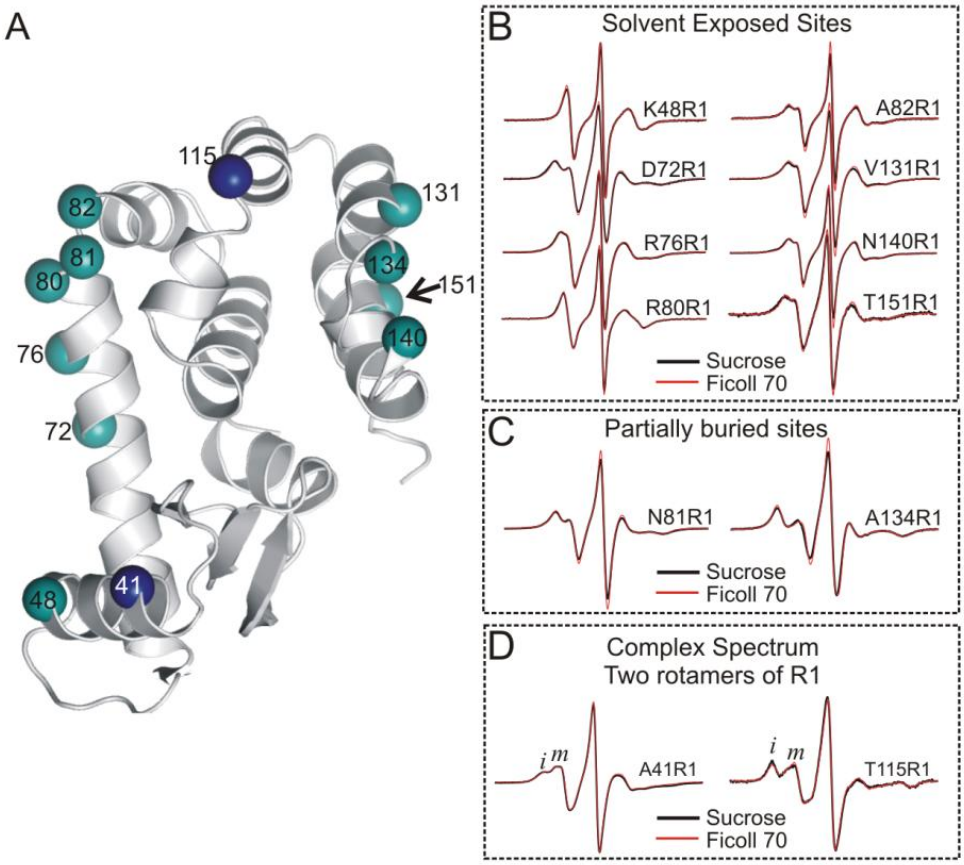


Figure 3.7: Effect of osmolyte perturbation on the motion of R1 in ordered regions of T4L. A) Ribbon diagram of T4L (pdb: 3lzm) showing sites where R1 was introduced. The spheres are color-coded to indicate sites where R1 has single-component (teal) or multicomponent spectra arising from two rotamers of R1 (dark blue). B) and C) EPR spectra of R1 at the indicated sites corresponding to cyan spheres in A). In each case the spectra in 30% w/w sucrose and in 25% w/w Ficoll 70 are superimposed. D) EPR spectra of residues 41R1 and 115R1 in sucrose and Ficoll 70. Letters *i* and *m* identify components of the spectra that correspond to relatively immobile and mobile states of R1.

3.3.3 Effect of Osmolytes on R1 Rotamer Equilibria. Earlier studies based on X-ray crystallography and mutagenesis indicated that the complex spectra of T4L mutants 41R1 and 115R1 (dark blue spheres in Figure 3.7A) arise from the presence of two rotamers of the R1 side chain [11,12]. The spectra of 41R1 and 115R1 in Ficoll are shown in Figure 3.7D (red trace) in which two components are resolved, corresponding to relatively immobile (i) and mobile (m) states of the side chain. The spectra in sucrose (black traces) are essentially identical, indicating that rotameric equilibria of R1 at these sites are insensitive to osmolyte perturbation.

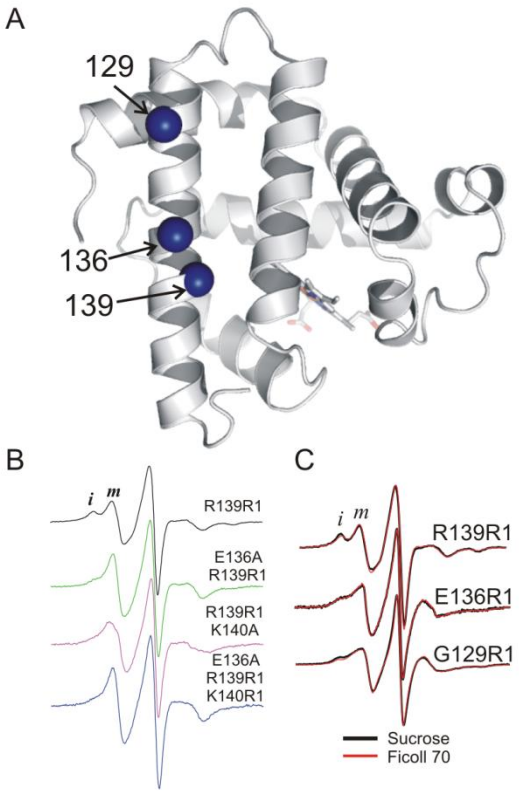


Figure 3.8: Effect of osmolyte perturbation on the motion of R1 in ordered regions of holo-Myb. A) Ribbon diagram of holo-Myb (pdb: 2mbw) showing spheres at sites where R1 was introduced. B) Effect of mutations of neighbor residues on the EPR spectra of 139R1. C) The corresponding EPR spectra of 129R1, 136R1, and 139R1 in 30% w/w sucrose and 25% w/w Ficoll 70 are superimposed. Letters *i* and *m* identify components of the spectra that correspond to relatively immobile and mobile states of R1.

Although structural data are not available, it is likely that the two-component spectra of the 129R1, 136R1, and 139R1 mutants of holo-Myb (Figure 3.8A) also arise from multiple rotamers with differential interactions in the structure, rather than from multiple protein conformations. Indeed, mutation of nearby residues E136 and K140 to Ala strongly reduces the immobilized component of the 139R1 spectrum (Figure 3.8B). Moreover, residues 129, 136, and 139 are located in the N-terminal region of helix H, which has been identified as one of the most stable regions of the holo-Myb molecule

[48] and apparently does not undergo conformational exchange in solution [49]. As shown in Figure 3.8C, there is no significant difference between the spectra of 129R1, 136R1 and 139R1 in Ficoll (red trace) and sucrose (black trace). Collectively, the above results show that rotameric equilibria of R1, like the internal motion, are insensitive to osmolyte perturbation.

3.3.4 Effect of Osmolytes on Conformational Equilibria. To study the effect of osmolyte perturbation on proteins with previously characterized conformational exchange, apo-rIFABP and apomyoglobin were selected as model systems. For each of these proteins, experimental evidence suggests that the conformational exchange occurs on the μs-ms time scale in discrete regions of the molecules [49,50]. Apo-rIFABP provides an example of low amplitude fluctuations of a short helix [50], while in apomyoglobin, a short region (corresponding to helix F in the holo form) is believed to undergo conformational exchange as inferred from NMR line broadening, but no direct information is available on the structures of the substates involved [49].

3.3.4.1 Apo-Rat Intestinal Fatty Acid-Binding Protein. rIFABP is a β-clam protein composed of 10 anti-parallel β-sheets and two α-helices that binds fatty acids in an internal cavity formed by opposing and roughly orthogonal β sheets (Figure 3.9) [50]. The specific function of the rIFABP is unclear, but it is thought the protein may play important roles in the intracellular solubilization, transport, and targeting of fatty acids to specific organelles and metabolic pathways [51]. NMR studies have identified three distinct regions of the apo-protein

that undergo conformational exchange on the μs-ms time scale [50] and that exhibit structural heterogeneity in solution (colored teal in Figure 3.9). These flexible regions are thought to constitute a dynamic portal, which regulates the entry and exit of the fatty acids [52,53]. Figure 3.10A shows a superposition of two of the NMR derived structures from the ensemble, with sequences identified to be in conformational exchange highlighted (*i.e.,* helix αII, and the βC-βD and the βD-βE turns).

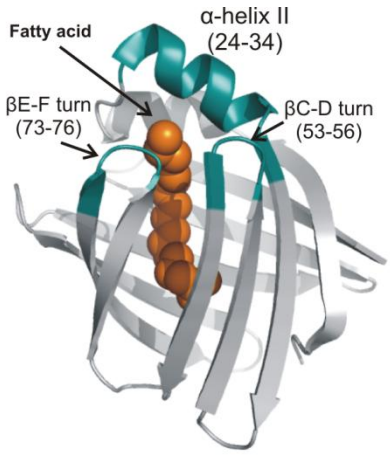


Figure 3.9: High-resolution structure of **holo**-rIFABP (pdb: 2ifb). The palmitate ligand is shown in cpk representation. The regions identified as flexible in the absence of the fatty acid are highlighted in teal.

Based on this information, R1 was introduced at sites throughout αII (24-35), in the βC/βD turn (55 and 56) and in the βE/βF turn (73-75) (Figure 6B). As a reference, R1 was also introduced at sites 14-18 in helix αI, which is spatially adjacent to αII, but not thought to be in conformational exchange [50]. The EPR spectra recorded in Ficoll 70 at the sites investigated are shown in Figure 3.10 (panels C and D, red traces). With few exceptions, the EPR spectra in αII, the βC/βD and the βE/βF turns are complex, with components corresponding to relatively immobile and mobile states of R1. The nitroxide mobility for R1 residues in αII, measured by the

inverse central line width (δ^{-1}), reveals the basic periodicity expected for a helix (Figure 3.11). Although there are an insufficient number of sites in αI to judge periodicity, the mobility is consistent with expectations based on the crystal structure (pdb: 1ifc), (i.e., solvent exposed residues 15R1 and 16R1 have relatively high mobility, while the buried or partially buried residues 14R1, 17R1 and 18R1 are immobilized).

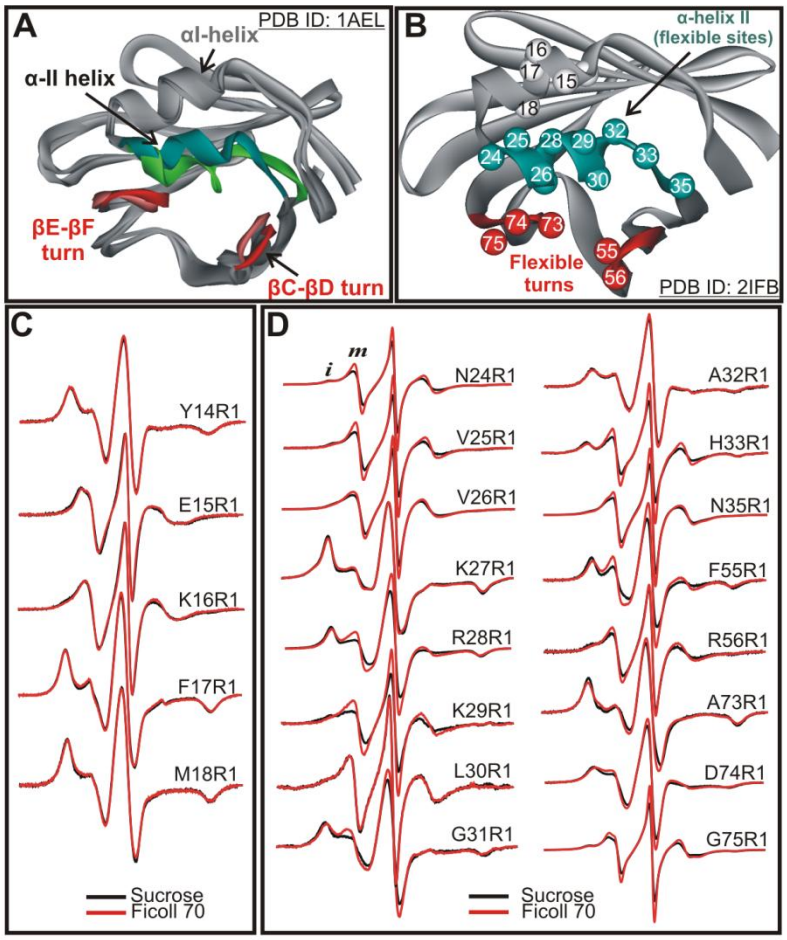


Figure 3.10: Effect of osmolyte perturbation on the motion of R1 in ordered and flexible regions of apo-rIFABP. A) Ribbon diagram of apo-rIFABP showing 2 of the 20 NMR structures consistent with structural constraints (Hodsdon *et al.* [50]). The flexible regions are highlighted (green and blue for α-helix II and red for the βC-βD and βE-βF turns). B) Ribbon diagram of the holo protein showing the positions where R1 was introduced. C) EPR spectra of R1 at sites in α-helix I in 30% w/w sucrose and 25% w/w Ficoll. D) EPR spectra of R1 at sites in the flexible regions of apo-rIFABP in 30% w/w sucrose and 25% w/w Ficoll. Letters *i* and *m* identify components of the spectra that correspond to relatively immobile and mobile component of the spectra.

The spectra of R1 at sites within helix αI show no response to osmotic perturbation (Figure 10C), consistent with the single conformation of αI determined by NMR relaxation [50]. On the other hand, the majority of the residues in the αII helix and in the βC/βD and βE/βF turns showed substantial sensitivity to osmolyte perturbation (Figure 10D, black traces). In each case, sucrose increased the fraction of immobile components relative to the more mobile ones, reflecting a shift in the equilibrium between the states that give rise to these components. Similar effects were observed when the stabilizing osmolytes betaine and sarcosine were used as the osmotic perturbants (*see* Appendix).

To evaluate these results more quantitatively, some of the spectra showing significant differences between Ficoll 70 and sucrose were fit to a two-component model to estimate the relative populations and the apparent equilibrium constant K = [*m*]/[*i*] (*see* Appendix). In Ficoll 70, K ranged from 0.41 (at 31R1) to 2.33 (at 56R1). Addition of sucrose decreased K by 25 – 57%, corresponding to free energy shifts ($\Delta\Delta G^{o}$ = ΔG_{suc}° - ΔG_{fic}°) of +169 to +492 cal/mole. In addition to the changes in K predicted by the osmolyte effect, sucrose also caused small changes in the effective correlation time (τ) and order parameter (S) of the mobile component, but essentially no change in those parameters for the immobile component (*see* Appendix).

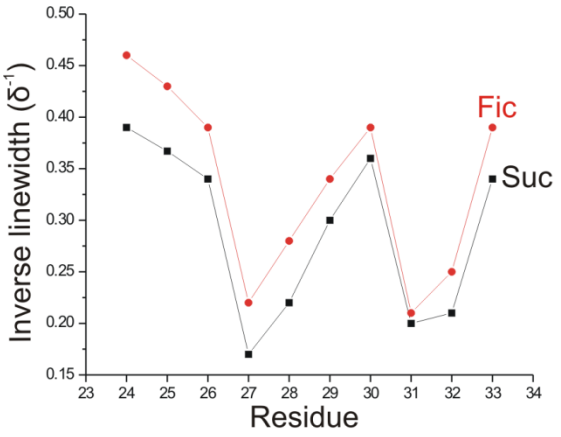


Figure 3.11: Inverse central line width *versus* sequence for R1 residues in the αII-helix of apo-rIFABP.

3.3.4.2 Apo-Myoglobin. The holo- and apo- forms of myoglobin have similar tertiary structures in solution [49,54] with the exception of residues 82-101, which form the F helix and flanking residues in the holo form (pdb: 2mbw), but are apparently in conformational exchange on the μs-ms time scale in the apo-protein [55]. To evaluate the effect of osmolyte perturbation on this flexible region of apomyoglobin, the R1 side chain was introduced at five sites within the region, and the spectra were recorded in sucrose and in Ficoll 70. As a series of controls, the R1 chain was also incorporated at sites in the helices A, B, D, E, and in the N-terminal region of helix H (Figure 3.12A), which are not thought to be in conformational exchange in apomyoglobin [49,55]. The corresponding EPR spectra are shown in Figures 3.12B and C. The motion of all R1 residues located in ordered regions of apomyoglobin were insensitive to osmolyte perturbation (Figure 3.12B). The spectra of R1 at every site located within the helix F region have two dominant components, reflecting relatively immobilized (*i*) and highly mobile (*m*) states of the R1 side chain. The highly mobile state is similar to the EPR spectrum of R1 in unstructured protein sequences [56-57]. Thus, these results suggest equilibrium between a folded and a partially or fully unfolded state of helix F of apomyoglobin[1].

In contrast to the results for R1 in the ordered regions, sucrose has a significant effect on the relative populations of *i* and *m* states of R1 along helix F, substantially reducing the more mobile state with a concomitant increase in the immobilized population (Figure 3.12C). Qualitatively, the above results are similar to the effect observed in the flexible αII helix in rIFABP, where sucrose shifts the equilibrium toward the more immobilized state(s). The spectra

[1] Additional insights into the conformational properties of helices F of apomyoglobin will be discussed in detail in chapter 5 of this dissertation.

for 84R1, 87R1, 89R1 and 91R1 in sucrose and in Ficoll were fit to a two-component model (*see* Appendix). The apparent equilibrium constants K = [*m*]/[*i*] in Ficoll ranged from 0.20 to 0.35. Addition of sucrose decreased K by 34 - 46% (*see* Appendix), corresponding to free energy differences ($\Delta\Delta G^{o} = \Delta G_{suc}^{\circ} - \Delta G_{fic}^{\circ}$) of +236 to +348 cal/mole, similar to the range in apo-rIFABP.

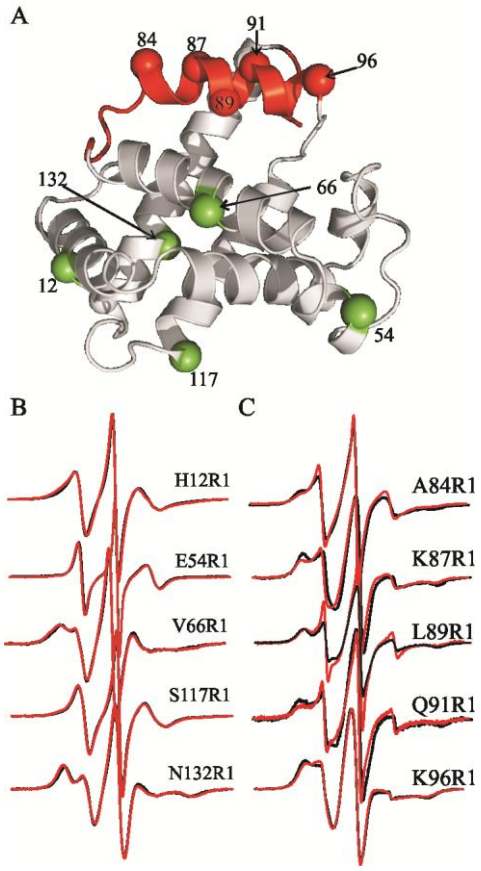


Figure 3.12: Effect of osmolyte perturbation on the motion of R1 in ordered and flexible sequences of apo-Myb. A) Ribbon diagram of holo-Myb (pdb: 2mbw) showing spheres at the sites where R1 was introduced. The sequence in conformational exchange in the apo form is highlighted in red. B) EPR spectra of R1 at sites in ordered regions of apo-Myb in 30% w/w sucrose and 25% w/w Ficoll 70. C) EPR spectra of R1 at sites in helix F in 30% w/w sucrose and 25% w/w Ficoll 70 are compared. Letters *i* and *m* identify components of the spectra that correspond to relatively immobile and mobile component of the spectra

Because the above results suggest equilibrium between a folded and a partially or fully unfolded state of helix F in apomyoglobin, Ficoll 70 could, in principle, exert crowding effects, although it has essentially no effect on the thermal stability of apomyoglobin or on the far UV circular dichroism spectra (Figure 3.4). To directly examine whether Ficoll influences the localized conformational isomerization of helix F, apomyoglobin with R1 at sites 87, 89, or 91 was immobilized on a Sepharose solid support in which any crowding effects of Ficoll could be isolated. As shown by the data in Figure 3.13, the addition of Ficoll had no effect on the spectra, while the addition of sucrose showed the same effect as in solution relative to Ficoll 70. These results combined support the contention that Ficoll 70 exerts little crowding effects on the conformational flexibility of helix F in apomyoglobin.

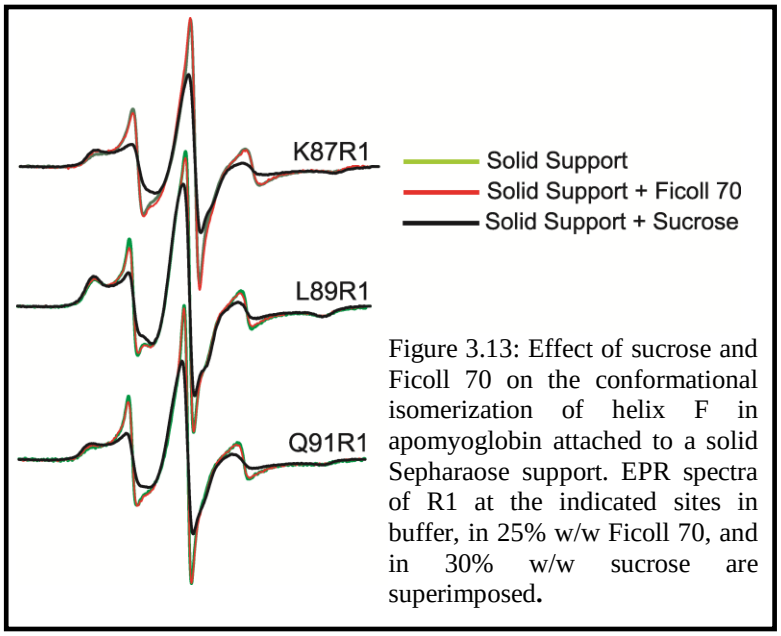


Figure 3.13: Effect of sucrose and Ficoll 70 on the conformational isomerization of helix F in apomyoglobin attached to a solid Sepharaose support. EPR spectra of R1 at the indicated sites in buffer, in 25% w/w Ficoll 70, and in 30% w/w sucrose are superimposed.

3.4 DISCUSSION

3.4.1 Suitability of Ficoll 70 as an Agent to Selectively Reduce Rotational Diffusion. In the studies of the small (15 – 19 kDa) model proteins investigated here (*i.e.,* T4L, rIFABP, and Myb), a 25% Ficoll 70 solution was used to match the effective viscosity of a solution with that of a 30% sucrose solution. Although Ficoll can exert molecular crowding effects [33,34] it does not appear to do so for the proteins investigated in this study (Figure 3.4). Indeed, at concentrations up to 25%, Ficoll has no measurable effect on the conformational equilibria of helix F residues in apomyoglobin (Figure 3.13). However, if Ficoll 70 or another similar agent is employed to increase viscosity in other systems, it is important to test for potential crowding effects, particularly in proteins that have a tendency to oligomerize. For sufficiently high molecular weight proteins (>50 kDa) and membrane associated proteins, it is unnecessary to employ an agent to match the effective viscosity of sucrose due to the already slow diffusion of the protein. Because the Ficoll molecule (~70 kDa) is considerably larger than each of the model proteins, the microscopic viscosity at the protein surface in a Ficoll solution is essentially the same as in pure buffer. Thus, one would not expect Ficoll to influence the internal motion of R1, and the data presented in Figures 3.5 and 3.6 strongly support this conclusion. Together, the results demonstrate that for the small proteins investigated here, Ficoll 70 can be employed as a simple viscosity agent to reduce the protein rotational diffusion rate with no additional effects on the protein or R1 side chain motion.

3.4.2 Effect of Sucrose on the Dynamics of R1 in Well-Ordered Protein Sequences. In comparing the effect of sucrose with that of Ficoll in order to isolate osmolyte perturbation effects, it is assumed that there is no effect of microscopic viscosity on the internal dynamics of

R1, at least in the narrow range of viscosities studied. As discussed previously, the microscopic viscosity near the protein surface in 25% Ficoll is essentially the same as buffer, and consequently has no effect on R1 internal dynamics. Sucrose at 30% also has little effect on R1 internal motions as illustrated by the data in Figure 3.6. This is not surprising if one considers the nitroxide as a small molecule, diffusing on the surface of the protein. Stokes-Einstein behavior predicts that $\frac{\partial \tau}{\partial \eta} \alpha r^3$, where τ, η and r are the correlation time, the viscosity of the medium, and the hydrodynamic radius of the diffusing species, respectively. Since the radius of the protein is at least an order of magnitude larger than that of the nitroxide, the effect of viscosity is ≈ 1000 times smaller on the R1 internal motion compared to the rotational diffusion of the protein.

In principle, an increase in the solution viscosity could damp the rate of ns backbone fluctuations ("slaved processes"; Beece *et al.* [58]; Fenimore *et al.* [59]); such effects would be manifested as an increase in the central line width (δ) of the EPR spectrum. However, the EPR spectra of R1 at solvent exposed sites in ordered sequences throughout T4L, apo-rIFABP, and apomyoglobin show subtle difference between sucrose and Ficoll 70 (Figures 3.7B, 3.10C, and 3.12B) solutions that differ in microscopic viscosity by about a factor of 3. The amplitude of backbone fluctuations or the change in viscosity might be too small to yield an observable effect in ordered sequences. For complex spectra of R1 at sites in conformational exchange, the rate of motion ($1/\tau$) of the most mobile component is reduced to some extent by sucrose, as determined by spectral simulations (Tables A.1 and A.2 in appendix section). It should be noted that central

line width[2] (δ) for all sites in the α-helix II of apo-rIFABP are quantitatively broader in sucrose compared to those in Ficoll 70 as shown in Figure 3.11. While it is possible that the apparent reduction of rate of the more mobile state is the results of sucrose damping flexible backbone motions, it is also possible that it arises from heterogeneity of the mobile state (i.e., the "mobile" state could actually consist of a manifold of states with similar, but not identical, dynamic properties). Population shifts among states in this heterogeneaous mixture due to osmolyte perturbation would appear as changes in rate, and hence as changes in $1/\tau$ when interpreted in the simple two-state model. The frequency of slow conformational exchange is also predicted to be slaved to solvent fluctuations [59] and hence to be dependent on microscopic viscosity. However, if this effect is present, it cannot be detected because the exchange rate is slow on the EPR time scale and therefore will not affect the spectra of the individual states involved.

3.4.3 Osmolyte Perturbation on R1 Rotameric Equilibria and Protein Conformational Equilibria. Complex EPR spectra reflecting multiple dynamic modes of the R1 side chain are very common in SDSL studies. Prior to this study, determining the origins of complex EPR spectra was challenging because multiple dynamic modes of the nitroxide can result from simple R1 rotameric equilibria in a static protein structure, but also from conformational exchange. The data presented in this report suggest a simple solution to this problem, namely the response of the protein to osmolyte perturbation as reflected in the EPR spectrum. Although there are a limited number of examples, the data in Figures 3.7D and 3.8 show that rotameric equilibria of R1 are insensitive to osmolyte perturbation, as anticipated from the small differences in solvent exposed

[2] The central line width is primarily determined by the rate of motion of the nitroxide (Columbus et al., 2004).

area between rotamers of R1. On the other hand, R1 located at many sites within regions that are known to undergo μs–ms conformational exchange show substantial sensitivity to osmolyte perturbation (Figures 3.10D and 3.12C).

The osmolyte-induced shifts in conformational equilibria in apo-IFABP and apomyoglobin are small because they involve localized changes in relatively short helices where the difference in solvent exposed area between the conformers is not large. Bolen and coworkers have shown that for global unfolding the major contribution to osmolyte stabilization is the positive free energy of transfer of the peptide group from water to osmolyte solution [18], which is about 56 cal/mole for 1M sucrose [60] (≈ 30% w/w). Consider a simple two-state conformational fluctuation of a short helix of 10 residues, approximately the length of the helices examined in rIFABP and Myb, where state (i) is folded and state (m) is partially unfolded. If the transition (i)→ (m) resulted in the exposure of 5 additional peptide groups to solvent, the difference in the standard state free energy change for the transition in water and sucrose, $\Delta G^{o}_{sucrose} - \Delta G^{o}_{water}$ would be 5x56 = 280 cal/mole from the peptide contribution alone, within the range 169 – 492 cal/mole estimated for the apo-rIFABP and apomyoglobin transitions based on a two-state model. The conclusion that the peptide backbone dominates the osmolyte effect is based on global unfolding where the contribution from side chains to the free energy is small [18,61]. However for local structural changes the side chain contribution could be substantial and the above calculation would only be an order-of-magnitude estimate.

3.4.4 Comparison with other EPR methods for detection of conformational substates in dynamic equilibrium in spin-labeled proteins. The data presented in this report show that

osmolyte perturbation in combination with SDSL can be used to identify localized and subtle conformational fluctuations in proteins that have substates with μs or longer lifetimes. In addition to solvent perturbation with osmolytes, three additional strategies have been recently developed to identify protein conformational equilibria on the μs time domain *via* SDSL, namely; high-pressure EPR [62], saturation recovery (SR) EPR [63], and pulsed electron-electron double resonance (ELDOR) [64]. In particular, the main advantages of high pressure EPR is the capability of (a) providing quantitative information of the volumetric properties of substates in exchange and (b) identifying the existence of low-lying (< 5 %) excited substates that are more disordered than the ground state [62]. Among these strategies, SR and ELDOR provide a direct measure of exchange rates between substates in the μs time domain [63,64]. The major advantage of the osmotic-perturbation method introduced here is its simplicity; a protein fold can be sampled with a small number of R1 sensors on the outer surface where structural perturbations are minimal, and the response of the spin labeled protein to sucrose or other osmolytes then is examined with a conventional EPR spectrometer without the need of additional instrumentation. The aforementioned strategies combined with earlier methods for sampling fast backbone dynamics [5] expands the capability of SDSL for exploring protein dynamics in the ps to millisecond time scale. At the present stage, the osmolyte perturbation strategy is qualitative. For example, it is not possible to determine the extent of conformational space sampled in the exchange process, nor can one deduce with certainty the number of such states involved. Nevertheless, a qualitative identification of the regions of a protein that are flexible and knowledge of the time scales of that flexibility are often what is needed to test models and to design additional experiments to determine the magnitude of changes involved using other

methods (such as direct distance measurements using DEER spectroscopy [65]). Osmolyte perturbation is especially well-suited for identifying conformational exchange in high molecular weight and membrane bound proteins for which the application of NMR methods is particularly challenging. It should be noted that in such cases the need for Ficoll as a reference state is unnecessary because the rotational diffusion is intrinsically slow.

3.5 Appendix

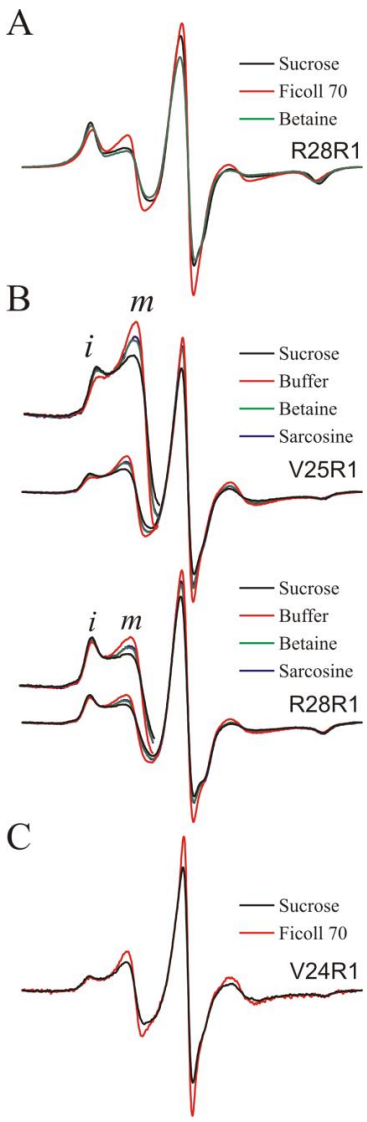


Figure 3.A.1: Stabilizing osmolytes sucrose, betaine, and sarcosine exert similar effects on the conformational exchange of α-helix II in apo-rIFABP in solution and immobilized on solid support. EPR spectra of R28R1 in sucrose, Ficoll 70, and betaine *in solution*. B) EPR spectra of residues V25R1 and R28R1 immobilized on a solid support at the indicated conditions. The low field sections of the spectra were amplified for clarity. C) EPR spectra of V24R1 immobilized on a solid support in sucrose and Ficoll 70.

Simulations of EPR spectra for R1 in apo-rIFABP and apo–Myb

Apo-rIFABP

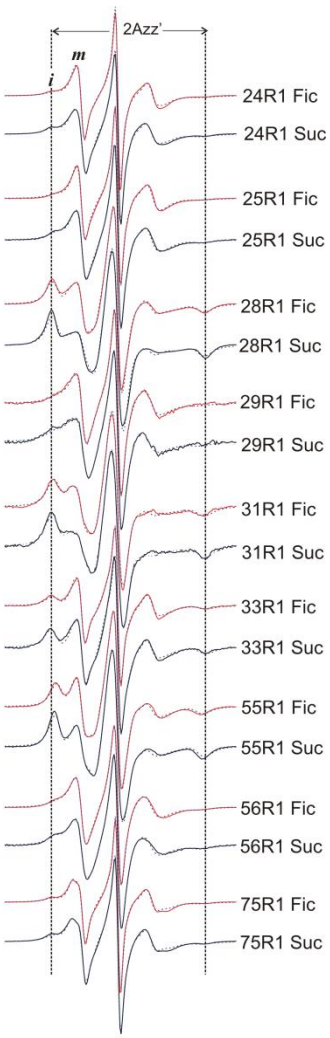


Figure 3.A.2: EPR spectra and 2 component MOMD fits of residues showing significant osmolyte perturbation effects in apo-rIFABP. The black and red traces represent the spectra in 30% w/w sucrose and 25% w/w Ficoll 70 respectively. The dashed blue lines are the best fits.

Apomyoglobin

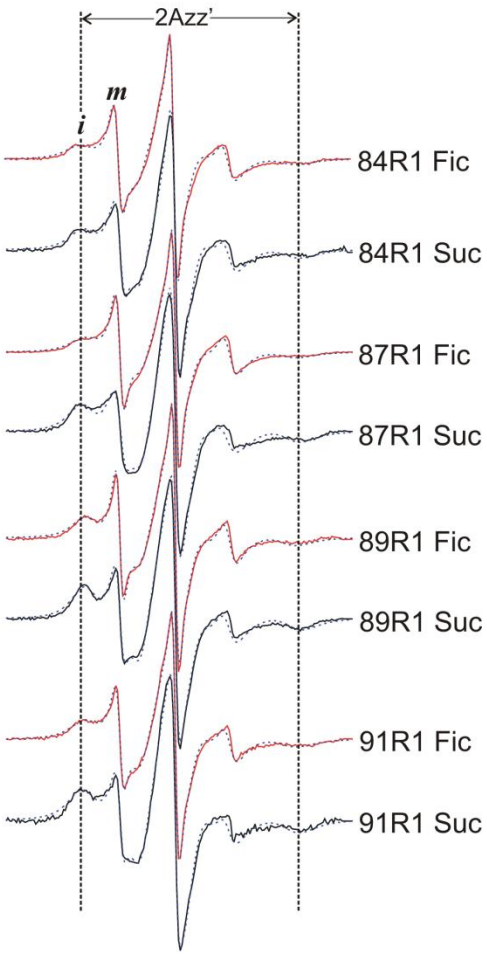


Figure 3.A.3: EPR spectra and 2 component MOMD fits of residues showing significant osmolyte perturbation effects in apomyoglobin. The black and red traces represent the spectra in 30% w/w sucrose and 25% w/w Ficoll 70 respectively. The dashed blue lines are the best fits.

Table 3.A.1: Dynamic parameters and relative proportions of the mobile (m) and immobilized (i) components for residues in the flexible regions of apo-rIFABP estimated from spectral simulations. (S = 30% sucrose, F = 25% Ficoll 70)

Residue	**Solute**	**% *i***	**τ_c (ns)**	**% *m***	**τ_c (ns)**	**S_m [(b)]**	**K [(a)]**
24R1	F	38	7.0	62	1.4	.24	1.63
	S	45	7.4	55	1.9	.25	1.22
25R1	F	34	7.0	66	1.9	.18	1.94
	S	42	7.0	58	2.2	.18	1.38
28R1	F	68	21.0	32	1.8	.32	0.47
	S	76	21.0	24	1.9	.37	0.32
29R1	F	38	5.9	62	2.5	.18	1.63
	S	53	6.5	47	3.0	.21	0.89
31R1	F	71	21.0	29	2.4	.31	0.41
	S	78	21.0	22	2.4	.38	0.28
33R1	F	56	21.0	44	1.6	.23	0.78
	S	73	21.0	27	1.6	.26	0.37
55R1	F	66	21.0	34	1.9	.21	0.52
	S	80	21.0	20	1.9	.23	0.25
56R1	F	30	7.0	70	2.5	.13	2.33
	S	50	7.0	50	2.5	.17	1.00
75R1	F	35	7.9	65	1.7	.31	1.86
	S	42	7.9	58	1.7	.32	1.38

[a] $K = [m]/[i]$

[b] S_m = order parameter for the *m* component

Table 3.A.2: Dynamic parameters and relative proportions of the mobile (m) and immobilized (i) components for residues in the flexible regions of apomyoglobin estimated from spectral simulations. (S = 30% sucrose, F = 25% Ficoll 70)

Residue	Solute	% *i*	τ_c (ns)	% *m*	τ_c (ns)	S_m [(b)]	K [(a)]
84R1	F	75	7.0	25	1.4	0.11	0.33
	S	82	7.0	18	1.6	0.12	0.22
87R1	F	74	5.9	26	1.3	0.08	0.35
	S	82	6.6	18	1.8	0.13	0.22
89R1	F	80	7.5	20	1.5	0.06	0.25
	S	86	7.5	14	2.0	0.06	0.16
91R1	F	83	5.9	17	1.3	0.11	0.20
	S	90	6.3	10	1.4	0.11	0.11

[a] $K = [m]/[i]$

[b] S_m = order parameter for the *m* component

3.6 References

1. Mittermaier, A., and Kay, L.E. 2006. New tools provide new insights in NMR studies of protein dynamics. *Science* 312: 224-228.
2. Mchaourab, H.S., Kálai, T., Hideg, K., and Hubbell, W.L. 1999. Motions of spin-labeled side chains in T4 lysozyme: Effect of side chain structure. *Biochemistry* 38: 2947-2955.

3. Columbus, L., Kálai, T., Jekö, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in α-Helices: Analysis by variation of side chain structure. *Biochemistry* 40: 3828-3846.
4. Columbus, L. and Hubbell W.L. 2002. A new spin on protein dynamics. *TIBS* 27: 288-295.
5. Columbus, L., and Hubbell, W.L. 2004. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* 43: 7273-7287.
6. Fink, A.L. 2005. Natively unfolded proteins. *Curr. Opin. Struct. Biol*. 15: 35-41.
7. Sugase, K., Dyson, H.J., Wright, P.E. 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447: 1021-1025.
8. Kusnetzow, A., Altenbach, C., and Hubbell W.L. 2006. Conformational states and dynamics of rhodopsin in micelles and bilayers. *Biochemistry* 45: 5538-5550.
9. Crane J.M., Suo, Y.Y., Lilly, A.A., Mao, C.F., Hubbell, W.L., and Randall, L.L. 2006. Sites of interaction of a precursor polypeptide on the export chaperone SecB mapped by site-directed spin labeling. *J. Mol. Biol.* 363**:** 63–74.
10. Langen, R., Oh, K.J., Cascio, D., and Hubbell, W.L. 2000. Crystal structures of spin labeled T4 lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry* 39: 8396-8405.
11. Guo, Z., Cascio, D., Hideg, K., Kálái, T., and Hubbell, W.L. 2007. Structural determinants of nitroxide motion in spin-labeled proteins: Tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. *Protein Sci.* 16: 1069-1086.

12. Guo, Z., Cascio, D., Hideg, K., and Hubbell, W.L. 2008. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci.* 17: 228-239.

13. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. and Somero, G.N. 1982. Living with water stress: evolution of osmolyte systems. *Science* 217: 1214-1222.

14. Yancey, P.H. 2005. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *The J. Exp. Biol.* 208: 2819-2830.

15. Arakawa, T., and Timasheff, S.N. 1985. The stabilization of proteins by osmolytes. *Biophys. J.* 47: 411-414.

16. Auton, M. Ferreon A.C., Bolen, D.W. 2006. Metrics that differentiate the origins of osmolyte effects on protein stability: a test of the surface tension proposal. *J. Mol. Biol.* 361: 983-92.

17. Timasheff, S., and Xie, G. 2003. Preferential interactions of urea with lysozyme and their linkage to protein denaturation. *Biophys. Chem*. 105: 421-448.

18. Liu, Y. and Bolen, B.D. 1995. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* 34: 12884-12891.

19. Bolen, W.D., and Rose, G.D. 2008. Structure and energetic of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem*. 77: 339-362.

20. Fanucci, G.E., Lee. J.Y., and Cafiso, D.S. 2003. Spectroscopic evidence that osmolytes used in crystallization buffers inhibit a conformational change in a membrane protein. *Biochemistry* 42: 13106-13112.

21. Kim, M., Xu, Q., Fanucci, G.E., and Cafiso, D.S. 2006. Solutes modify a conformational transition in a membrane transport protein. *Biophysical J*. 90: 2922-2929.

22. Butler, S. L., and Falke, J.J. 1996. Effects of protein stabilizing agents on thermal backbone motions: A disulfide trapping study. *Biochemistry* 35: 10595-10600.

23. Kendrick, B.S., Chang, B.S., Arakawa, T., Peterson, B., Randolph, T.W., Manning, M.C., and Carpenter, J.F. 1997. Preferential exclusion of sucrose from recombinant interleukin-1 receptor antagonist: Role in restricted conformational mobility and compaction of the native state. *Proc. Natl. Acad. Sci.* 94: 11917-11922.

24. Cioni P., Bramanti, E., and Strambini, G.B. 2005. Effects of sucrose on the internal dynamics of azurin. *Biophys. J*. 88: 4213-4222.

25. Chen, L.Y., Ferreira, J.A.B., Costa, S.M.B., Cabrita, G.J.M., Otzen, D.E., and Melo E.P. 2006. Compaction of ribosomal protein S6 by sucrose occurs only under native conditions. *Biochemistry* 45: 2189-2199.

26. Lee, J.C., and Timasheff, S. 1981a. The stabilization of proteins by sucrose. *J. Biol. Chem*. 256: 7193-7201.

27. Lee, J.C., Gekko, K., and Timasheff, S. 1981b. Measurements of preferential solvent interaction by densimetric techniques. *Methods Enzymol*. 61: 26-49

28. Luby-Phelps, K., Castle, P.E., Taylor, D.L., and Lanni, F. 1987. Hindered diffusion of inert tracer particles in the cytoplasm of mouse 3T3 cells. *Proc. Natl. Acad. Sci*. 84: 4910-4913.

29. Zhou, H.X., Rivas, G., and Minton, A.P. 2008. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys*. 37: 375-397.

30. Sale, K., Sár, C., Sharp, K.A., Hideg, K., and Fajer, P.G. 2002. Structural determination of spin label immobilization and orientation: A Monte Carlo minimization approach. *J. Mag. Res*. 156: 104-112.

31. Freed, J.H. *Theory of slow tumbling ESR spectra for nitroxides. In Spin Labeling: Theory and Applications*, L.J. Berliner, ed. New York, Academic Press, pp. 53-132.

32. Wenner, J.R., and Bloomfield, V.A. 1999. Crowding effects on EcoRV kinetics and binding. *Biophysical J*. 77: 3234-3241.

33. Stagg, L., Zhang, S.Q., Cheung, M.S., and Wittung-Stafshede, P. 2007. Molecular crowding enhances native structure and stability of α/β protein flavodoxin. *Proc. Natl. Acad. Sci.* 104: 18976-18981.

34. Roque, A., Ponte, I., and Suau, P. 2007. Macromolecular crowding induces a molten globule state in the c-terminal domain of histone H1. *Biophys. J*. 93: 2170-2177.

35. Minton, A.P. 1981. Excluded volume as a determinant of macro-molecular structure and reactivity. *Biopolymers* 20: 2093–2120.

36. Santoro, M.M., and Bolen, D.W. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry* 27:8063-8068.

37. Fleissner, M., Cascio D., and Hubbell, W.L. 2009. Structural origins of weakly ordered nitroxide motion in spin labeled proteins. *Prot. Sci.* 18:893-908.

38. Timofeev, V.P., and Tsetlin, V.I. 1983. Analysis of mobility protein side chains by spin label technique. *Biophys. Struct. Mech*. 10: 93-108.

39. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chain in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35: 7692-7704

40. Anderson, D.E., Lu, J., McIntosh, L., and Dahlquist, F.W. 1993. The folding, stability, and dynamics of T4 lysozyme: a perspective using nuclear magnetic resonance. *In NMR of Proteins*, G.M. Clore, and A.M. Gronenborn, eds. Boca Raton, Fla, CRC, pp. 258-304.

41. Liang, Z., Lou, Y., Freed, J., Columbus, L., and Hubbell, W. 2004. A multifrequency electron spin resonance study of T4 lysozyme dynamics using the slowly relaxing local structure model. *J. Phys. Chem. B.* 108: 17649-176559.

42. Wilchek, M., and Miron, T. 2003. Oriented versus random protein immobilization. *J. Biochem. Biophys. Methods*. 55: 67-70.

43. Lowe, C.R., and Dean, P.D.G. 1974. *Affinity chromatography*. John Wiley & Sons Ltd., London.

44. Guo, Z. 2003. "Correlation of spin label side-chain dynamics with protein structure: studies of T4 lysozyme with site-directed mutagenesis and x-ray crystallography" Ph.D. thesis, University of California, Los Angeles.

45. Migneault, I., Dartiguenave, C., Bertrand, M.J., and Waldron, K.C. 2004. Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *BioTechniques* 37: 790-801.

46. Persson, E., and Halle, B. 2008. Nanosecond to microsecond protein dynamics probed by magnetic relaxation dispersion of buried water molecules. *J. Am. Chem. Soc*. 130: 1774-1787.

47. Thomas, D.D., and Hidalgo, C. 1978. Rotational motion of the sarcoplasmic reticulum Ca^{2+}-ATPase. Proc. Natl. Acad. Sci. 75:5488-5492.

48. Cavagnero, S., Thériault, Y., Narula, S.S., Dyson, H.J., and Wright, P. 2000. Amide proton hydrogen exchange rates for sperm whale myoglobin obtained from ^{15}N-^{1}H NMR spectra. *Prot. Sci.* 9: 186-193.

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

50. Hodsdon, M.E. and Cistola, D.P. 1997. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N NMR relaxation and ^{1}H exchange. *Biochemistry* 36: 2278-2290.

51. Falomir-Lockhart, L.J., Laborde, L., Kahn, P.C., Storch, J., and Córsico, B. 1996. Protein-membrane interaction and fatty acid transfer from intestinal fatty acid-binding protein to membranes. Support for a multistep process. *J. Biol. Chem*. 281:13979-13989.

52. Hodsdon, M.E. and Cistola, D.P. 1997. Discrete backbone disorder in the nuclear magnetic resonance structure of the apo intestinal fatty acid-binding protein: Implications for the mechanism of ligand entry. *Biochemistry* 36: 1450-1460.

53. Eads, J., Sacchettini, J.C., Kromminga, A., and Gordon, J.I. 1993. *Escherichia coli*-derived rat intestinal fatty acid-binding protein with bound myristate at 1.5 Å resolution

and I-FABP$^{Arg106\rightarrow Gln}$ with bound oleate at 1.74 Å resolution. *J. Biol. Chem.* 268: 26375-26385.

54. Lecompte, J.T., Kao, Y.H., and Cocco, M. J. 1996. The native state of apomyoglobin described by proton NMR spectroscopy: The A-B-G-H interface of wild-type sperm whale apomyoglobin. *Proteins J.* 25: 267-285.

55. Eliezer, D., Yao, Y., Dyson, H.J., and Wright, P. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Struc. Biol.* 5: 148-155.

56. Qu, K. B., Vaughn, J. L., Sienkiewicz, A., Scholes, C. P., and Fetrow, J. S. 1997. Kinetics and motional dynamics of spin-labeled yeast Iso-1-cytochrome *c*: 1. Stopped-flow electron paramagnetic resonance as a probe for protein folding/unfolding of the C-terminal helix spin-labeled at cysteine 102. *Biochemistry* 36:2884−2897.

57. Langen, R., Cai, K., Altenbach, C., Khorana, H.G., and Hubbell, W.L. 1999. Structural features of the C-terminal domain of bovine rhodopsin: a site-directed spin-labeling study. *Biochemistry* 38:7918-7924.

58. Beece, D., Einstein, L., Frauenfelder, H., Good, D., Marden, M.C., Reinisch, L., Reynolds, A.H., Sorensen, B., and Yue, K.T. 1980. Solvent viscosity and protein dynamics. *Biochemistry*. 19: 5147-5157.

59. Fenimore, P.W., Frauenfelder, H., McMahon, B.H., and Parak, F.G. 2002. Slaving: Solvent fluctuations dominate protein dynamics and functions. *Proc. Natl. Acad. Sci.* 99: 16047-16051.

60. Auton, M., and Bolen, W. D. 2004. Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale. *Biochemistry* 43: 1329-1342.

61. Wang, A., and Bolen, W.D. 1997. A naturally occurring protective system in urea-rich cells: Mechanism of osmolyte orotection of proteins against urea denaturation. *Biochemistry*. 36: 9101-9108.

62. McCoy, J. and Hubbell, W.L. 2011. High-pressure EPR reveals conformational equilibria and volumetric properties in spin-labeled proteins. *Proc. Natl. Acad. Sci.* 108:1331-1336.

63. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* 37:363-390.

64. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011 Structure and dynamics of a conformationaly constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246.

65. Jeschke, G. 2002. Distance measurements in the nanometer range by pulse EPR. *ChemPhysChem*. 3: 927-932.

Chapter 4: Mapping the molecular flexibility of metmyoglobin by SDSL and correlation with structural properties of the protein

4.1 Summary

Identifying protein flexibility on the picoseconds to milliseconds and longer time scales is essential in the understanding of the molecular mechanism of protein function. In the last chapter a new strategy was reported for identifying slow exchange in spin-labeled proteins. In the study presented here, the newly developed strategy was employed along with other methods developed for sampling nanosecond backbone motions to study the molecular flexibility of myoglobin with SDSL. The results presented here revealed sequence-dependent changes in the rate and amplitude of nanosecond local backbone motions in stable helices of myoglobin and suggest that these motions may be determined to some extent by structural properties of the protein. Osmotic-perturbation studies showed that the protein exhibits limited conformational flexibility in the microsecond to millisecond time. Interestingly, the sequences that exhibit high flexibility on the nanosecond time scale do not correspond to the regions in conformational exchange on the microsecond to millisecond time scale.

4.2 Introduction

The protein structures derived from x-ray crystallography and NMR spectroscopy provide a static snapshot of the lowest energy ground state conformation of the molecule. This prevalent static picture that often emerges from structural biology is however insufficient to understand the molecular mechanism of protein function given that proteins in solution exist in thermodynamic equilibrium between multiple conformational and statistical substates, some of which are required for function (*see* Chapter 1). Thus to obtain a more complete picture on the dynamic modes relevant for function it has been essential the application of solution techniques such as NMR relaxation experiments [1], hydrogen/deuterium exchange [2], and fluorescence [3]. SDSL have also been particularly effective in providing structural and functionally relevant dynamic information of soluble and membrane-bound proteins [4-5]. For example, conformational switching on light-activated rhodopsin required for visual signal transduction was first identified by SDSL [6]. In this case, conformational exchange between substates was detected *via* changes in the mobility of the R1 side chain upon light activation [6-7], while the magnitude of the structural change was elucidated by interspin distance measurements on the dark and light-activated states [8-9]. In the previous chapter of this dissertation it is shown that exchange between conformational substates in dynamic equilibrium in R1-labeled proteins can be identified by the response of spin-labeled proteins to osmotic perturbation. Additional strategies that can be used to identify conformational exchange in spin-labeled proteins have been recently reported [10-12].

In addition to identifying conformational exchange, SDSL have been shown to be useful in detecting collective and local backbone fluctuations on the nanosecond time scale in membrane and soluble proteins [13-14]. Unlike conformational exchange, the nanosecond

motions of the backbone contribute directly to the EPR spectra. In this case, the relative contribution from fast backbone motions (τ_{bb}) to the spectra can be elucidated provided that the relative contributions from other dynamic modes (*i.e.*, internal motions of the R1 side chain [τ_{int}] and rotational diffusion of the protein [τ_r]) are known. As discussed in the previous chapter, the contribution from τ_r for membrane proteins and for soluble proteins with $M_W \geq 50$ kDa to the spectrum is negligible, thus for such cases only τ_{int} and τ_{bb} determine the spectrum, while for small proteins, the contributions from τ_r can be minimized by increasing the effective viscosity of the protein solution [15] or by selectively immobilizing the protein on a solid surface (*see* Chapters 3 and 7). Previous SDSL studies in combination with high-resolution structures of spin-labeled proteins [16-18], variation of nitroxide side chain structure [19], mutational analysis [15,20], and quantum mechanical calculations [21] have provided a quantitative model that accounts for the contribution from τ_{int} to the EPR spectrum of R1 in solvent-exposed α-helical and loop sites in soluble and membrane-bound proteins thereby allowing functionally relevant backbone fluctuations to be extracted and quantified (refer to section 2.84 in chapter 2 for more details on the structural origins for the internal motion of R1). Based on these studies it was found that for R1 at solvent exposed sites, where the nitroxide lacks any tertiary interaction**, the contribution from τ_{int} should be independent of position, thus variations in the mobility of the nitroxide side chain among topologically equivalent sites reflect differential contributions from local backbone fluctuations (τ_{bb}) on the nanosecond time scale**. This capability of SDSL to extract backbone motions on spin-labeled proteins was validated using the well-characterized DNA-binding domain of GCN4 [14].

Although the aforementioned study revealed a gradient in the scaled mobility and rates of motion of R1 in GCN4 that is in excellent agreement with NMR relaxation data, the high

mobility of the basic sequence of GCN4 in the free form (*i.e.*, not bound to DNA) did not allow for a sequence-dependent measurement of the relative amplitude of the backbone motions with SDSL [14]. Thus, to show the capability of SDSL to evaluate the sequence-dependent changes in the relative rates and amplitudes of motion of the backbone in well-ordered protein sequences, it is necessary to extend the previous study to proteins that are more ordered than that of the basic sequence of GCN4. Myoglobin is an ideal model system for this purpose as it has long served as a paradigm for the study of the connection between structure, dynamics, and function. To date, over 200 high-resolution structures of the molecule has been solved in different space groups, including an NMR-derived structure [22], while spectroscopic methods such as fluorescence [23], H/D exchange data [24], and NMR spectroscopy [25] have provided a wealth of information on the molecular flexibility of the protein in solution.

In the study presented here, the sequence-dependent changes in the mobility of the R1 side chain at solvent-exposed surface sites of myoglobin were evaluated. In addition, given that recent studies have suggested that backbone motions can be directly encoded in the three dimensional structure of the protein [26-27], the possible link between structural properties of myoglobin with the local fast backbone motions detected by SDSL was investigated. Finally, the osmotic-perturbation strategy introduced in the preceding chapter was used to detect slow conformational exchange. It should be emphasized that even though the holomyoglobin molecule contains a paramagnetic ion (*see* below), the EPR signal of the iron center is not observed in the studies presented here. As discussed in section 2.5 of chapter 2, the intrinsic short T_1 relaxation (~ ps) of the metal at room temperature, broadens its EPR resonance beyond detection.

4.2.1 *Sperm whale myoglobin.* Myoglobin, which primary function is to store oxygen in cardiac and skeletal muscles, was the first structure to be seen at atomic resolution thanks to the

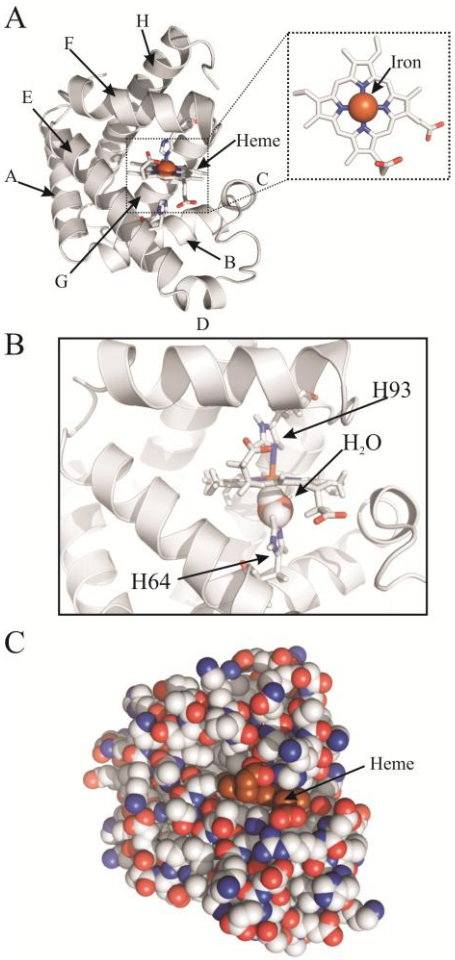


Figure 4.1: Structure of sperm whale myoglobin. A) Ribbon model of myoglobin (pdb: 2mbw) showing the location of the helices. *Inset*: Stick representation of heme. B) Heme pocket. The proximal and distal histidines are highlighted. The ligand (water) is shown in cpk (pdb:1cq2). C) CPK model of myoglobin

groundbreaking work by Kendrew and colleagues [28]. As depicted in Figure 4.1, the myoglobin molecule consists of a single polypeptide chain of 153 amino acids[1] and an iron atom complexed to protoporphyrin IX (*see* inset in Figure 4.1A), which is required for ligand binding. Approximately 70% of the protein is folded into 8 α–helices (denoted A-H), while the rest of the chain forms the interhelical turns. The protein is covalently bonded to the porphyrin ring *via* a histidine-iron bond at the fifth coordination position of the metal center (Figure 4.1B). The ligand-binding site in the molecule is located at the opposite surface of the heme plane near histidine 64, which is known as the distal site. The iron center in the porphyrin ring can exist in either ferrous (Fe^{2+}) or ferric (Fe^{3+}) oxidation states, the former of which is the only one capable of binding oxygen. In solution, the protein rapidly (~ hrs) oxidizes to metmyoglobin (Fe^{3+}; [29]), turning into a physiologically inactive form whereby water is the primary ligand in the distal site as depicted in Figure 4.1B[2]. Earlier studies have shown that although the free heme group can bind oxygen, the myoglobin scaffold provides higher

[1] Recombinant sperm whale myoglobin expressed in *E. coli* contains an N-terminal methionine, which is denoted as M0.

[2] In the oxidized metmyoglobin form, the water ligand can be ionized to OH^- or exchanged with other anions ligands such as cyanide and azides.

selectivity for oxygen binding and reduces the rate of autooxidation of the iron center from ~ 1 s to hours [30].

As depicted in Figure 4.1C, the CPK model of the high resolution structure shows no clear pathway for ligand migration from the solvent to the binding pocket, suggesting that protein motions are required for function. Indeed, elegant experiments done by Austin and coworkers [31] revealed that protein fluctuations between conformational and statistical substates are required for ligand binding [32]. More recent time-resolved x-ray crystallography studies have provided additional insights into the migration pathway and structural changes required for ligand migration inside the protein [33-34].

4.3 Results

To map the molecular flexibility of myoglobin, the nitroxide side chain R1 was placed, one at a time, at 39 individual solvent-exposed sites throughout α-helices A to H, A/B turn, E/F turn, and G/H turn (Figure 4.2).

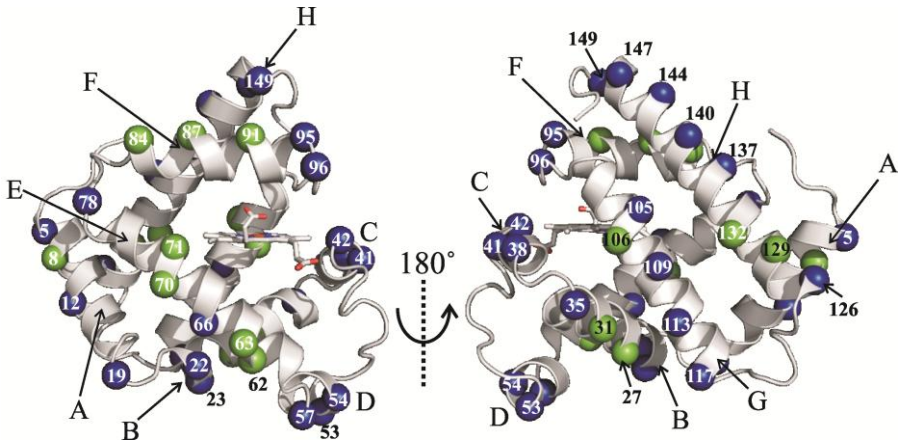


Figure 4.2: Ribbon diagram of the myoglobin molecule (pdb: 2mbw) showing the sites where the R1 side chain was introduced. The blue spheres represent the sites where the EPR spectra reflect a single component and the site is defined as "non-interacting" (*see* text below).

Solvent-exposed sites in the holomyoglobin structure were selected in order to minimize any structural perturbation by the paramagnetic probe. Previous studies have shown that introduction of R1 at surface sites causes little perturbation in soluble [16-17,35-36] and membrane [18] proteins. In addition, mutations of the surface residues selected here were not expected to impair heme binding to the apo protein [37-38]. Indeed, the spin-labeled proteins were fully reconstituted with heme as judged by A_{409}/A_{280} ratio (*see* Appendix).

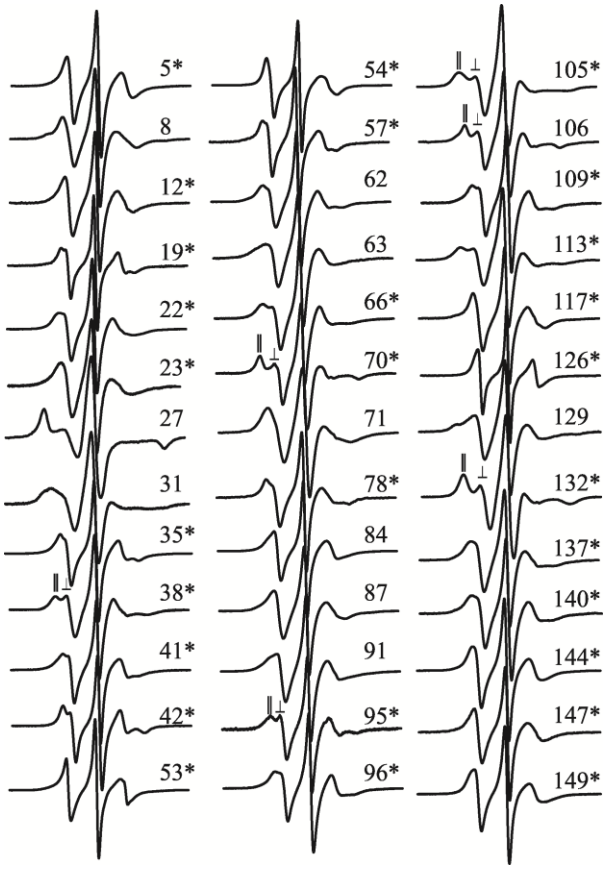


Figure 4.3: EPR spectra of R1 mutants of metmyoglobin. The asterisks identify the sites reflecting a single dynamic state of R1 (*see* text). The symbols ⊥ and ∥ identify the well-resolved hyperfine components arising due to anisotropic motion of the nitroxide with $S \geq 0.3$ (Columbus et al. [14]).

The EPR spectra of the 39 spin-labeled mutants are shown in Figure 4.3. It is interesting that the spectra observed for the 39 sites in myoglobin exhibit almost the complete range of R1 mobility that is often encountered in spin-labeled proteins: from broad spectra indicating

constrained motion of R1 arising from tertiary interaction (*e.g.*, D27R1) to spectra exhibiting high amplitudes and narrow line widths indicating fast ns motion of the nitroxide side chain (*e.g.*, A53R1). Interestingly, the EPR spectra of ~ 65% of the sites (blue spheres in Figure 4.2) studied here are consistent with a simple anisotropic motion of the nitroxide (highlighted with asterisks on Figure 4.3) characteristic of R1 at solvent-exposed sites where the nitroxide lacks any tertiary interaction with neighboring residues[3]. The relatively high frequency of sites reporting a single environment around the R1 side chain suggests that most of the sequence is well-ordered in the microsecond to millisecond time scale. Additional evidence that support the aforementioned argument will be shown below.

4.3.1 *Characterization of fast (nanosecond) backbone motions in myoglobin.* To monitor the relative contribution from τ_{bb} to the EPR spectra it should be emphasized that sites lacking tertiary interaction are ideal for monitoring backbone dynamics on the nanosecond time scale given that the quantitative model used to account for the contribution of τ_{int} is most appropriate for sites reflecting a single dynamic state of R1. Thus the rest of the discussion in this section will be based on the data from the 28 sites showing simple spectra. Although spectral simulation based on the MOMD model can be used to extract information on the rate ($1/\tau$) and order of motion (S) of the nitroxide for a particular site, it has been shown [14] that for large-scale studies, the relative mobility can be measured by an empirical "scaled mobility" parameter, which is determined from the spectral central line width (M_s) and is proportional to τ^{-1} (Figure

[3] A "non-interacting" site is defined as a surface site for which R1 has a single component EPR spectrum reflecting anisotropic motion wherein the order parameter (S) of the nitroxide does not exceed that for the purely internal motion, *i.e.*, $S \approx 0.5$.

4.4A). Thus M_s measurements were used in this work to probe sequence-specific changes in mobility of R1.

As shown in Figure 4.2, the sites with EPR spectra reflecting a simple anisotropic motion of the nitroxide are distributed throughout the A-H helices, A/B turn, and E/F turn, thus the mobility of R1 at each site gives valuable information on local backbone flexibility of the different structural elements of the molecule. The average scaled mobility per helical segment is shown in Figure 4.4B and clearly reveals sequence-specific differences in the mobility of R1 for topologically equivalent sites. For instance, the spectra of R1 at surface sites on the short D-helix reflect high mobility with an average M_s value 30% higher than the rest of the molecule (Figure 4.4B). It should be noted that some regions that may be more flexible than the D-helix sequence (*e.g.*, CD loop) were not studied in this work. As shown in Figure 4.4B, the average mobility of the nitroxide side chain for sites on the C-helix was the second highest suggesting that the short helices C and D are the most flexible helices of the holomyoglobin molecule on the nanosecond time scale. Interestingly, helix D exhibits relatively high B-factor values in the high-resolution structure [39]. In addition, crystallographic studies of holo myoglobin at multiple temperatures (80–300 K) revealed a considerable volumetric expansion of the sequence corresponding to helix C, the CD turn, and helix D as the temperature was raised [40], which supports the contention that this sequence is the more plastic region of the molecule.

Conversely, the mobility of the nitroxide side chain for sites within helices B, E, and G (M_s = 0.60) are below average ($\langle M_s \rangle$ = 0.70; red line in Figure 4.4B) suggesting a relative reduction in the contribution from τ_{bb} to the spectral line shape. As shown in Figure 4.4B, the lowest average scaled mobility value is observed for the B-helix suggesting that this sequence is the most rigid part of the molecule on the nanosecond time scale. As depicted in Figure 4.4C,

the high-resolution structure shows low solvent accessibility (mean solvent-accessible surface area = 18%) and extensive contact of the B-helix with several regions of the protein including helices D, E, and G. In fact surface residues D27R1 and R31R1 reflect reduced mobility due to tertiary interaction with the protein environment (Figure 4.3)[4]. In addition to the high packing density around the B-helix, there is an important stabilizing hydrogen bond between the buried histidine 24 on the B-helix with histidine 119 [41-42]. Thus it is likely that the low mobility of the nitroxide side chain at sites within the B-helix are a reflection of restricted nanosecond motion of the sequence due to high packing and extensive interaction of the helix with the protein environment. Additional evidence that suggests a correlation between the extent of local backbone motions, as judged by the mobility of the R1 side chain, with structural properties of the protein will be discussed below.

Based on the EPR spectra of the 28 sites reflecting single dynamic state of R1, the scaled mobility values were mapped onto the structure of the holomyoglobin molecule. As depicted in Figure 4.4D, save for the short helices C and D, the lowest mobility of the nitroxide side chain is generally observed near the center of the helices (Ms ~ 0.66), while the loops and helix termini generally exhibit an increase in the Ms values (Ms ~ 0.78). The higher mobility of R1 towards the ends of some of the helices suggests dynamic fraying in these regions. In this regard, it is pertinent to mention that hydrogen/deuterium exchange experiments showed evidence of dynamic fraying near the end of some of the helices [24].

[4] Residues D27R1 and R31R1 were not used in the calculation of average Ms value for the B-helix since both are multicomponent spectra.

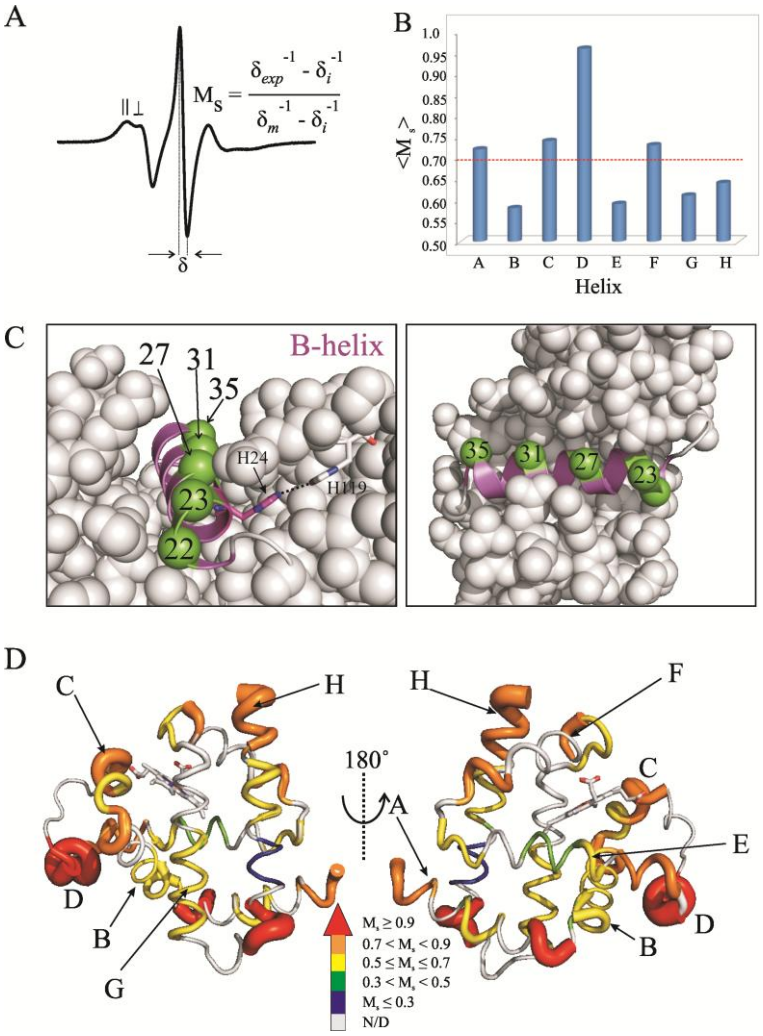


Figure 4.4: Sequence-dependent scaled mobility values in myoglobin. A) EPR spectrum indicating the central linewidth, which is used to calculate the scaled mobility (δ_i = 8.0 and δ_m = 2.1 G) [14]. B) Average scaled mobility value per helix. C) Side view and top view of the local packing around helix B. The spheres at the Cα indicate the sites where R1 was introduced. The hydrogen bond between H24 and H119 is shown in yellow. D) "Putty" representation of the myoglobin molecule. The color and width of the backbone is proportional to the Ms values (*see* legend).

4.3.2 Correlation between scaled mobility values with structural properties of the protein. In order to determine whether the extent of local backbone fluctuations, as judged by the mobility of the R1 side chain, is influenced by structural properties of the proteins (*e.g.*, solvent exposure of the backbone), the fraction of surface buried on a per residue basis was calculated using the high resolution structure of the holo protein (pdb: 2mbw) and the software Getarea 1.1 [43] with a probe radius of 1.4 Å. The average fraction of surface buried for a particular residue is defined as the solvent inaccessible area for a particular residue in the context of the three dimensional structure normalized by the solvent inaccessible area of the same residue in the tripeptide Gly-X-Gly in a random coil conformation (*see* methods section in chapter 8). The average fraction of surface buried taken over a sliding window of 6-9 residues around the nitroxide side chain was found to have a statistically significant anti-correlation ($R^2 = 0.71$) with the M_s values (Figure 4.5A).

The M_s values for the R1 side chain in α-helices can either reflect rigid body motions (*i.e.*, axial helix rocking motions) or local segmental fluctuations [13]. It is expected that for short helices, such as helices C and D of myoglobin, where the M_s values and surface burial are similar within the segment, the relative mobility of the nitroxide may reflect collective modes of motions (*i.e.*, motions of the helix as a whole). However, for long helices where there is a variability of local surface burial and M_s values, the site-to-site differences in mobility of the R1 side chain may reflect local segmental motions as those previously observed in GCN4 [14].

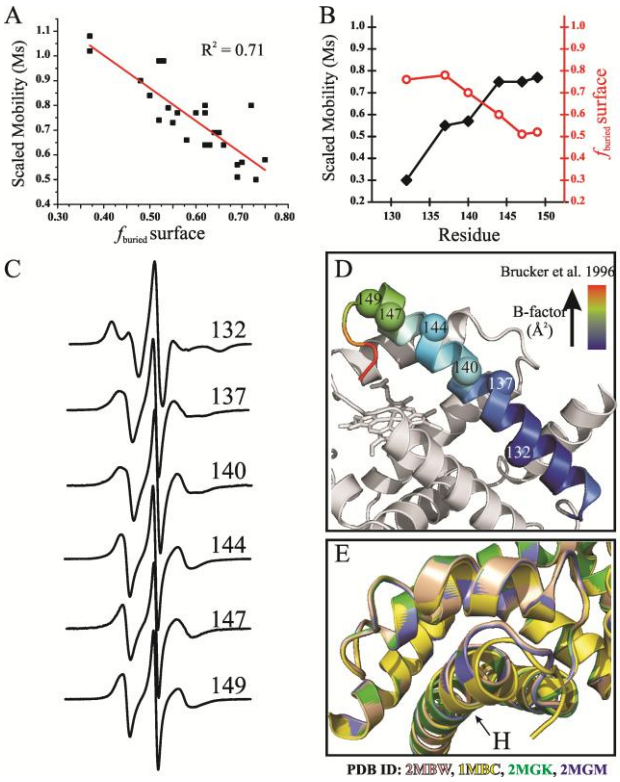


Figure 4.5: The sequence-dependent changes in the mobility of the nitroxide side chain correlates with structural properties of the protein. A) Correlation between scaled mobility and fraction of surface buried. B) Gradient of scaled mobility and fraction of surface buried observed for residues within the H helix. C) EPR spectra of R1 at sites within H-helix. D) B-factor values for the Cα atoms in the high-resolution structure (pdb: 2mbw; Brucker et al.[47]). E) Overlay of high-resolution structures of myoglobin showing structural plasticity near the c-terminal region of the H-helix. The pdb identification numbers are shown.

Indeed, this appears to be the case in the long helix H sequence. As shown in Figure 4.5B-C, a gradient of increasing mobility is observed from residue 132R1 to residue 149R1. The local fraction of surface buried along the same sequence (red trace on Figure 4.5B) shows a decreasing gradient likely due to the reduction in contact near the C-terminal end of the helix, in which only 25% of the helix surface interacts with other structural elements of the protein. Interestingly, the high-resolution structure of myoglobin shows a gradient of B-factor values along the same sequence (Figure 4.5D) with above average flexibility near the C-terminal end of the H-helix [39,48]. In addition, different high-resolution structures of myoglobin show plasticity beyond residue 146 (Figure 4.5E), thus it is not surprising that the mobility of R1 near the C-terminus of the H-helix reflects higher mobility than for R1 located at sites near the center of the helix. The aforementioned observations suggest that the gradient of R1 mobility observed within the long H-helix likely arise from local segmental motions on the nanosecond time scale that might be determined by the extent of local contact with other structural elements of the protein.

4.3.3 Correlation between rate and order of motion with structural properties of the protein. Even though the measurements based on the central line width values are a practical approach to evaluate *relative* mobility of R1, M_s does not give quantitative information about the amplitudes and rate of motion of the nitroxide side chain. To evaluate if there is a correlation between the rate ($1/\tau$) and amplitude (S) of motion of R1 with the segmental fraction of surface buried (f_{buried}), the EPR spectra of 13 of the sites reflecting a single dynamic state of R1 were fit to the MOMD model[5] (*see* Appendix). The correlations between rate of motion and the order parameters obtained from the fits with f_{buried} are shown in Figure 4.6. The analysis revealed that there is a statistically significant anti-correlation ($R^2 = 0.84$) between the rate of motion and the f_{buried} similar to that observed with M_s (Figure 4.5A). This finding was not surprising in lieu of previous studies, which have shown that, in principle, the M_s value is a primary measure of the relative rate of motion of R1 at helix surface sites [14]. The analysis also showed that there is a statistically significant correlation ($R^2 = 0.50$) between the order parameter, which describes the amplitude of motion of the nitroxide side chain, and f_{buried}. The latter finding suggests that the amplitude of the local backbone fluctuations can be in fact determined to some extent by the local packing. Although the correlation of order *versus* f_{buried} is statistically significant, it is weaker compared to the correlation of rate *versus* f_{buried}. The possible interpretations for the relative differences in the correlation between rate and order with the fraction of surface buried will be

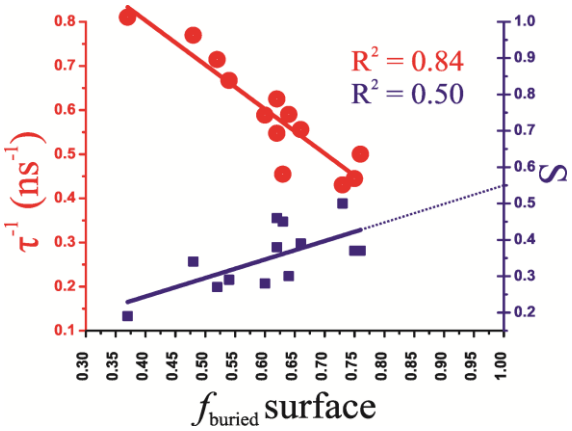


Figure 4.6: Correlation between the rate ($1/\tau$) and order of motion (S) with the fraction of surface buried are shown in red and blue, respectively.

[5] The rate and motion of the more mobile component in the spectrum of residue K62R1 was also included in the analysis.

discussed below (*see* Discussion). Remarkably, extrapolation of the linear fit of the order parameter *versus* f_{buried} suggests that for a solvent exposed site with f_{buried} = 1[6], the order parameter would be 0.55 ± 0.08, which is close to the order parameter observed at sites reflecting pure internal motion of R1 (*i.e.*, no contribution from τ_{bb}; [13,19]).

4.3.4 Mapping slow conformational exchange in the holomyoglobin molecule. To identify slow conformational exchange (τ_{exch} > 100 ns) in the holomyoglobin sequence, the osmotic-perturbation strategy introduced in Chapter 3 of this dissertation was employed. As discussed in Chapter 3 of this dissertation, osmotic-perturbation is a simple strategy that can be used in combination with SDSL, which relies on the response of the protein to solvent perturbation by stabilizing osmolytes such as sucrose. It has been established that stabilizing osmolytes shift protein conformational equilibria towards the least solvent-exposed state [49-50], thus the response of the spin-labeled protein to osmotic-perturbation can be used to identify localized and subtle conformational exchange provided that the solvent accessibilities are different in the states.

To map the conformational flexibility of holomyoglobin *via* osmotic-perturbation SDSL, the EPR spectra of R1 mutants of a set of sites showing single and multiple components were recorded in a solution containing either 30% w/w sucrose or Ficoll 70 (Figure 4.7). As discussed in the preceding chapter, Ficoll 70 is not an osmolyte and is used to match the effective viscosity of the protein solution such that the contribution of τ_r to the spectra is the same as the spectra recorded in sucrose solutions. As mentioned above, the EPR spectra of most sites in myoglobin reflect a single environment around the nitroxide side chain, indicating a single protein

[6] In reality a surface site cannot have f_{buried} = 1 since a segmental fraction of surface buried of 1 would imply that the sequence is completely buried.

conformation on those sites. On the other hand, the EPR spectra of 11 of the 39 residues studied reflect multiple dynamic states of the nitroxide (Figure 4.7A), which may arise from rotameric equilibria of R1 or protein conformational exchange (*see* Chapter 3). Interestingly, the multicomponent EPR spectrum of residue 71R1 is broader than a typical spectrum of R1 in a helix surface site; this is likely due to exchange interaction between the R1 side chain and the iron atom in the porphyrin ring due to their close proximity (*see* Discussion). The multicomponent spectra of residues 8R1, 27R1, 71R1, 84R1 and 129R1 (Figure 4.7A) showed little sensitivity to osmotic-perturbations suggesting that the two components likely arise from rotameric equilibria of R1 and not from conformational exchange. These results are in agreement with earlier studies because these residues are located in helices A, B, E, G, and H, which have been identified as well-ordered sequences with high protection factors [24].

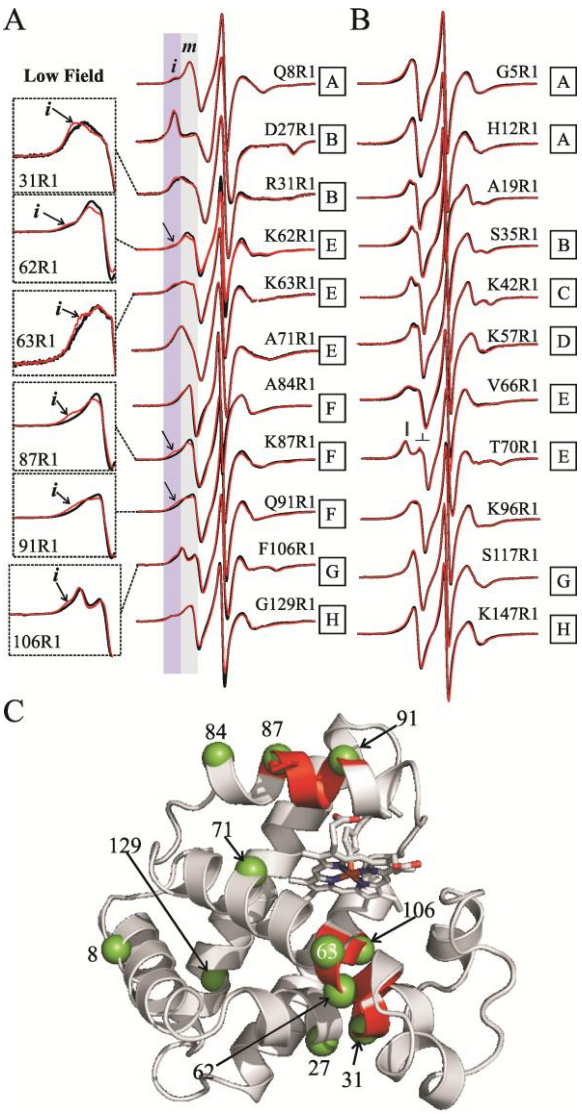


Figure 4.7: Osmolyte-perturbation revealed localized conformational exchange in discrete regions of metmyoglobin. A) Osmolyte-perturbation on sites reflecting multiple components. The low field region of the spectra was amplified for clarity. B) Osmolyte-perturbation on sites reflecting a single component. C) Ribbon model of myoglobin highlighting the regions in conformational exchange. The spheres at the Cα identify the sites where the spectra reflect two components.

On the other hand, the EPR spectra of residues K62R1 and K63R1, localized in the first turn of the E-helix, showed a subtle, but noticeable

increase in the population of the more immobile component in sucrose as shown in Figure 4.7A suggesting that the multicomponent spectra for both sites arise from conformational exchange. The osmotic-perturbation effect observed for residues 62R1, 63R1 and 106R1 is smaller than anticipated for a region in conformational exchange. It should be noted that for residue K62R1 and 106R1 the population giving rise to the immobile component is low (~ 10%) as judged by the relative spectral intensity (*see* arrow in Figure 4.7A) and spectral simulation (*see* Discussion), thus the subtle sucrose effect observed in this region may be a result of the low population of the second state. Interestingly, the EPR spectra of residues V66R1 and T70R1, which are located in the central region of the E-helix, reflect a single component and no response to osmotic perturbation (Figure 4.7B) suggesting that the conformational exchange process detected in residues 62R1 and 63R1 is limited to the first turn of helix E.

The EPR spectra of residues 31R1, 87R1, and 91R1 reflect multiple spectral components and showed significant sensitivity to osmotic-perturbation (Figure 4.7A). In each case, the addition of sucrose increased the intensity of the more immobile component (arrows in Figure 4.7A), reflecting a shift in the equilibrium between the substates that give rise to these components. Residues 87R1 and 91R1 are located in the center of the F-helix, which is directly linked to the heme group and has been shown to have low protection factors [24]. Remarkably, the spectra of residues A84R1 and K96R1, both of which are located at the ends of the F-helix (K96R1 reflecting a single component) are insensitive to osmotic-perturbation (Figure 4.7), suggesting that the conformational exchange process detected in residues 87R1 and 91R1 is localized to the central region of the helix.

The conformational exchange detected in residues 31R1 and 106R1 was not anticipated based on solution studies, which revealed that these sequences are well-ordered [24-25]. The

possible origins of the conformational exchange detected in the N-terminal region of the E-helix, the central region of the F-helix and in residues 31R1 and 106R1 will be discussed below (*see* Discussion)

As expected, the EPR spectra of sites showing a single dynamic state of the nitroxide are insensitive to osmotic-perturbation (Figure 4.7B), which is consistent with a single conformational state of the sequence. The features that might be interpreted as multiple components in the spectra of residues 42R1, 66R1, and 70R1 (Figure 4.7B) are the resolved parallel ($A_{\|}$) and perpendicular ($A\perp$) hyperfine components characteristic of an anisotropic motion of the nitroxide with $S \geq 0.3$ [14]. The above results combined show that holomyoglobin is well-ordered with limited conformational flexibility on the microseconds to milliseconds time scale observed in discrete regions of the molecule (Figure 4.7C).

4.4 Discussion

The primary goals of the work presented here were to (1) map the molecular flexibility (ns time scale and longer) of holomyoglobin with SDSL and (2) study whether there is a correlation between backbone flexibility and the structural properties of the protein. To this end, T_2-based line shape analysis was used to study the contribution from fast (ns) backbone fluctuations to the spectra, while the osmotic-perturbation strategy developed as part of this dissertation, was employed to identify slow (μs-ms) exchange between conformational substates. To minimize any structural perturbation by the paramagnetic probe, the R1 side chain was placed only at solvent-exposed sites throughout the molecule.

The utility of SDSL-EPR to detect ns backbone fluctuations has been demonstrated in earlier studies on the relatively unstable helices of GCN4 [14], which have been extensively

characterized by NMR [63]. Based on the NMR studies, the helices of GCN4, in the DNA-free form, sample both helical and non-helical conformations on the ps-ns time scale. The motions involve transient rupture of backbone hydrogen bonding and therefore the backbone experiences a large amplitude motion that is reflected in extensive averaging of the R1 magnetic tensors which give rise to a relatively sharp and narrow EPR spectral lineshape [14]. In addition to GCN4, the SDSL-EPR method has been employed to identify disordered sequences in several other proteins, including membrane proteins [13,64-65].

In contrast to these highly flexible sequences that give rise to sharp resonance lineshapes, the EPR spectra of R1 on the solvent exposed surfaces of well-ordered stable helices, as judged, for example, by high protection factors in hydrogen/deuterium exchange data, exhibit broader lines reflecting anisotropic constrained motions on the ns time scale ($1 \leq \tau \leq 4$ ns). Nevertheless, there are variations in the motion from site to site as shown in Figure 4.3 (*see* sites highlighted with asterisk). A question arises as to the origin of these variations. On one hand, the site-to-site variations in the spectra of R1 could reflect subtle differences in the local backbone motion. However, in principle, site-specific differences in the local interactions of the R1 ring with nearby side chains could also account for variations in EPR spectra of the kind shown in Figure 4.3 [56-57]. Extensive mutagenesis experiments have ruled out the latter as the source of variation for R1 at solvent exposed sites of the kind analyzed here [15,17].

4.4.1 Analysis of fast backbone motions and correlation with structural features of the protein. In order to obtain information on the relative contribution from fast backbone motions to the EPR spectra of spin-labeled myoglobin mutants, the mobilities of sites whose spectra reflect a single dynamic component were compared. Only sites lacking tertiary interactions with the environment (*i.e.*, sites reflecting a single spectral component) were selected for the analysis

since it is essential that the contribution from the internal motions of the side chain (τ_{int}) to the spectra is constant among all sites such that the site-to-site variations in the relative mobilities can be assigned to distinct local backbone motions in the nanosecond time scale.

The mobility of the nitroxide side chain, as measured by a normalized version of the central line width, for the 28 sites reflecting single dynamic component revealed region- and sequence-specific differences in the mobility of R1 for topologically equivalent sites (Figures 4.4B). Specifically, the short helices C and D, and the termini of helix H exhibit the highest mobility suggesting high flexibility on the nanosecond time scale. Interestingly, the regions where the R1 side chain exhibit the highest flexibility on the nanosecond time scale correspond to those regions where x-ray crystallography shows above average flexibility as judged by B-factor values and structural heterogeneity observed in different crystal forms of the molecule [39,48] (Figure 4.5D-E). Remarkably, the more mobile regions correspond to sequences where the local packing is relatively low as judged by solvent accessibility measurements. Conversely, regions with low solvent accessibilities exhibit relatively low mobility (Figures 4.4C) suggesting that the local packing may be a determinant of the extent of nanosecond backbone fluctuations inferred from the EPR spectral lineshape. Indeed, a statistically significant anticorrelation was observed between the scaled mobility values and a structural property of the protein, namely the local fraction of surface buried (Figure 4.5A). The rate and amplitude of motion of the nitroxide extracted from spectral simulation also showed a statistically significant correlation with the local fraction of surface buried (Figure 4.6), suggesting the interesting possibility that the rates and amplitudes of motion of the backbone on the nanosecond time scale in myoglobin are influenced by the three dimensional structure of the protein. In this regard, it is worth mentioning that previous studies have established a possible link between physical properties of proteins

with dynamics. For example, independent studies done by Mittermaier et al. [53] and Goodman et al. [24] found a weak, but statistically significant correlation between the NMR-derived N-H order parameter from a database of proteins with the degree of solvent exposure and local packing. Zhang and Brüschweiler [27] have shown that fast backbone motions, as described by the order parameter, can be accurately predicted from x-ray and NMR-derived structures by calculating local contacts experienced by the amide proton and carbonyl oxygen of the peptide bond. Another model that accurately predicts order parameters from the local contacts with the Cα atom has also been reported [55]. Analogous observations based on B-factor values from x-ray crystal structure were reported by Halle [26], who found that protein flexibility, as defined by the atomic mean square displacement in crystal structures, can be accurately predicted by the atomic packing density around the Cα atoms. These independent studies support the contention that local backbone flexibility can be directly encoded or influenced by the three dimensional structures of proteins. It should be noted that an alternative approach to studying the link between dynamics with the structural properties of proteins (*e.g.*, local packing) is to compare the backbone flexibility of wild-type and mutant versions of the same protein for which the local packing density has been modified. Results from such a study will be shown in chapter 6 of this dissertation.

In regards to the correlation observed in this work between the rate ($1/\tau$) and order of motion (S) of the nitroxide with the local fraction of surface buried, it is interesting that the correlation with the rate was stronger than that of the order parameter (R^2 = 0.84 *versus* 0.50), the latter of which describes the amplitude of motion. The origins for the differences in the correlation are unclear, but it is possible that the rate of motion is more sensitive to more types of backbone modes than the order parameter. For example, previous studies have suggested that the

average orientation of the nitrogen 2p orbital is roughly perpendicular to the helical axis (*see* review by Columbus and Hubbell, [13]), which in turn provides high sensitivity to helix rocking motion. Under this scenario, the differences in the order parameter between topologically equivalent sites are a measure of the differences in the angular amplitude of the helix rocking motions.

However, it should be emphasized that the correlation time (τ) and the order parameter are anti correlated in the spectra simulation [19]. For example, a equally satisfying fit can be obtained by lowering the correlation time with a concomitant increase in the order parameter, thus the solutions obtained from the simulations are not unique. In order to assess the level of uncertainty of each parameter, the MOMD fits for several of the sites were repeated multiple times (data not shown). From the multiple fits, a change in the correlation time of ~ 0.1 ns could be compensated with a maximum change in the order parameter of 0.05. Given this level of uncertainty, it is possible that the rate and amplitude of motions are correlated with the local fraction of surface buried to a similar extent. It should be noted that the scaled mobility value accounts for relative changes in both the rate and order of motion [14] and has the advantage of being an *empirically-determined* measure of relative motion that does not require time-consuming spectral simulations.

Regardless of the difference in the level of correlation between the rate and order with the local fraction of surface buried, an important finding from this study is that there is a correlation between the mobility of the nitroxide at topologically equivalent sites with the local packing. Interestingly, extrapolation of the linear fit of order *versus* f_{buried} to $f_{\text{buried}} = 1$ (Figure 4.6), yields an order parameter of 0.55 ($\pm$ 0.08), which is close to that observed for at sites where the spectrum reflects pure internal motion of R1 [19]. In addition, the intercept in the linear fit of

rate *versus* f_{buried}, is 1.2096 ns^{-1}, which corresponds to a correlation time of 0.83 ns. This value is close to those observed on the flexible basic helical sequence (residues 6-20) of the transcription factor GCN4 in the free form [14] for which the fraction surface buried ~ 0. Interestingly, a comparable correlation between scaled mobility and f_{buried} is observed if the scaled mobility values of R1 at non-interacting surface sites of T4 lysozyme are added to the Ms *versus* f_{buried} plot (Figure 4.8). These findings support the contention that the correlations observed in this work between the mobility of R1 with local packing in myoglobin are not spurious, but actually reflect a link between structural properties of proteins and local dynamics. One important outcome of this study is that these observations would not be expected if the mobility of R1 at helical surface sites was determined by interaction of the nitroxide ring with neighboring residues as suggested by molecular dynamics simulations [56-57], thus the data shown in this report supports the interpretation that differences in rate and amplitude of motion of R1 at non-interacting surface sites indeed reflect relative differences in the backbone fluctuations on the nanosecond time scale.

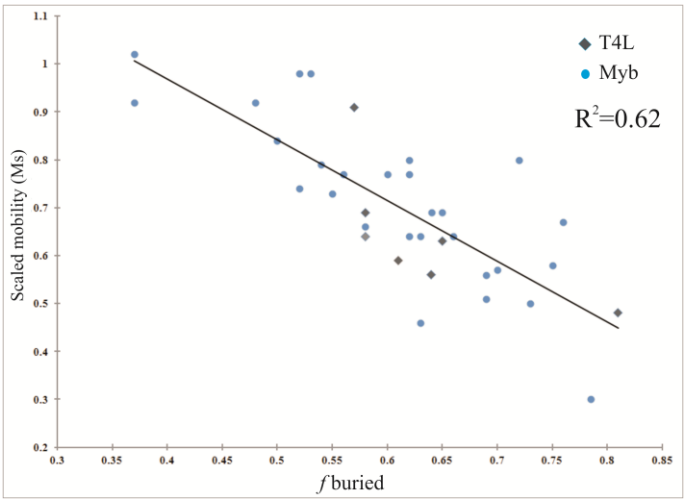


Figure 4.8: Correlation between scaled mobility of R1 at surface sites with segmental fraction of surface buried in myoglobin and T4 lysozyme (shown in gray diamonds).

4.4.2 Regions in slow conformational exchange. The structural and conformational properties of holomyoglobin *in solution* have been previously investigated *via* NMR Cα chemical shifts, ^{1}N-^{15}H NOE measurements, and hydrogen/deuterium exchange [22,24-25]. These studies revealed that the protein in solution adopts a structure very similar to those

obtained by x-ray crystallography [22] and exhibit restricted backbone flexibility [25]. Hydrogen/deuterium exchange measurements [24] revealed relatively high protection factors for residues in helices A, B, C, E, G, and G, suggesting that those regions correspond to the most ordered-sequence of the protein. Conversely, amides in helices D and F exhibit relatively low protection factors and are thus the least stable helices of the holo protein.

The data shown in Figure 4.7 revealed that, save for a few sites, most of the sequence of the holomyoglobin molecule exhibit little sensitivity to osmotic-perturbations supporting the notion that the protein is well ordered on the microsecond to millisecond time scale as suggested by the aforementioned studies. Interestingly, the EPR spectra of sites in the regions corresponding to the first turn of the E-helix (residues 62R1 and 63R1), the central region of the F-helix (residues 87-91R1) and in residues 31 and 106 were multicomponent and exhibit sensitivity to osmotic perturbation indicating conformational exchange between substates with lifetimes of > 100 ns. In each case, the more mobile component observed reflects anisotropic motion characteristic of R1 at helical surface sites, while the more immobilized state indicates tertiary interaction of the nitroxide side chain with the protein environment. Thus it is possible that the conformational equilibria observed particularly in the first turn of the E-helix and the central region of the F-helix involve subtle rotations of the local region whereby the R1 side chain moves from a solvent-exposed environment to one of tertiary contact with nearby residues as depicted in Figure 4.9. It is pertinent to mention that the sequence near residues 87 and 91 contains a “helix-breaking” proline at position 88, which distorts the helix and may be responsible for the conformational flexibility observed in the central region of the F-helix.

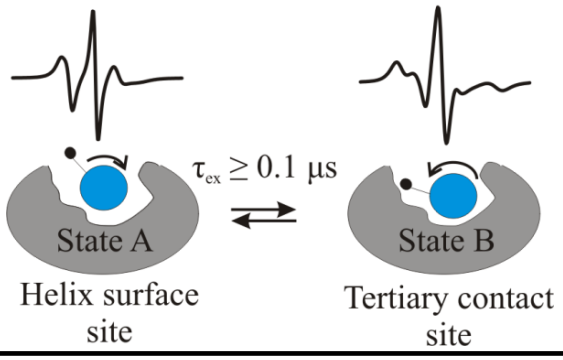


Figure 4.9: Proposed conformational exchange occurring on the first turn of E-helix and the central region of the F-helix. Conformational exchange involving local rotation of the helix whereby the nitroxide "senses" two distinct environments; solvent exposed (state A) and in tertiary contact with the protein (state B).

Although the conformational exchange in the N-terminal region of the E-helix, in the F-helix and in residues 31R1 and 106R1 were not reported in the NMR studies [25], the data from this study is not inconsistent with that of the NMR. The Cα resonances for residues 86-87 in the F-helix were not assigned in the NMR studies [61] and the average Cα chemical shift value for the residues in the center of the F-helix (86-90) is smaller than that of other helices. In addition, H/D exchange experiments show low protection factors for several amides within the F-helix. Residues 62 and 63 on the other hand have Cα chemical shift values and protection factors characteristic of residues in well-ordered helices. However, exchange of the helical segment containing residues 62-63 between two positions may not be detected by the NMR active-nuclei (Cα) if the environments around the (Cα) are very similar in both states. On the other hand, the R1 side chain can be very sensitive to even small structural changes (*i.e.* subtle helical rotation) given that such changes can bring the nitroxide side chain into tertiary contact with nearby residues of the protein as depicted in Figure 4.9. It should be noted that the second conformation for most sites exhibiting sensitivity to osmotic perturbation is relatively low (ca. 10 %) as judged by the spectral intensities corresponding to the more immobilized state and by spectral simulation (*see* Figure 4.10). The low population of the state giving rise to the more immobile

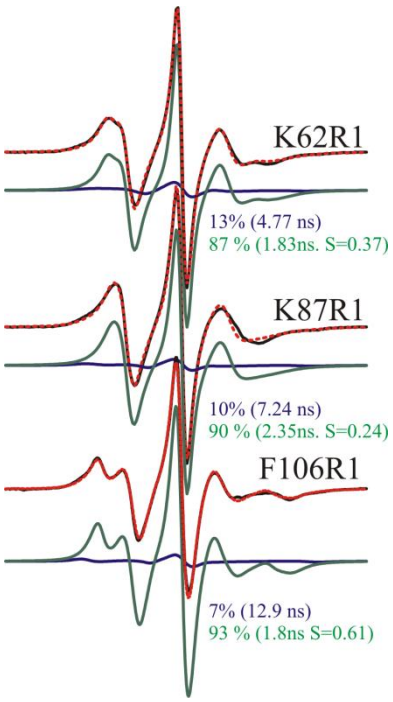


Figure 4.10: MOMD fit of the multicomponent spectra of residues 62R1, 87R1, and 106R1. The experimental data is shown in black trace and the fit is shown as dashed red lines. Deconvolution of the multicomponent spectra is shown for each site.

component is reduced when the data is recorded at 308 K (see Appendix), which is the temperature at which the NMR experiments were done, thus it is possible that the methods used in previous studies were not sensitive enough to detect such exchange between a highly populated ground state and a transiently populated state. In regard to the conformational exchange observed in residue 31R1, it should be noted that the neighboring spin-labeled sites in the sequence corresponding to helix B (*e.g.* 22R1, 23R1, 27R1, and 35R1) reflect single component spectra and/or little sensitivity to osmolyte perturbation (Figures 4.3 and 4.7), thus it is unlikely that the exchange observed at site 31R1 arises from conformational flexibility of the B-helix.

The functional relevance of the conformational flexibility in the N-terminal region of the E-helix observed here is unclear, however it is interesting that previous time-resolved crystallography studies done at room temperature showed concerted structural changes at the N-terminus of the E-helix (~ 300 ns) following laser photolysis of CO-bound protein, which are apparently relevant for ligand migration [58]. In regard to the conformational flexibility observed in the F-helix, residues 87 and 91 are close to the proximal site (His 93), which is directly bound to the heme group. Hence, it is possible that the localized conformational exchange identified on the F-helix may be functionally significant given that this region has shown plasticity upon binding of different ligands to the heme iron [48,59,60].

One interesting outcome of this study is that the sequences that exhibit high flexibility on the nanosecond time scale (*e.g.*, helices C and D, and C-terminal region of H-helix) do not correspond to the regions in conformational exchange on the microsecond to millisecond time scale suggesting that the molecular motions on the two time-domains are not necessarily coupled.

4.4.3 EPR spectrum of residue A71R1. Although the EPR resonance of the heme iron is not observed at room temperature due to its short T_1, in a few cases it is possible to observe its effect on the spectra of nitroxide-labeled proteins. The spectrum of residue 71R1 appears to be an example. As discussed in section 2.8.3 of Chapter 2 of this dissertation, when there is a second paramagnetic species near the nitroxide side chain, there is a distance-dependent ($r \leq 8$Å) exchange interaction between the species arising from overlap of the orbitals of the unpaired electrons, resulting in broadening of the spectrum. As depicted in Figure 4.11, residue 71 projects towards the porphyrin ring, while residue 70R1 projects towards the solvent and away from the porphyrin ring. Interestingly, the lineshape of 71R1 is broader than a typical spectrum of at a helical site, while the spectrum of residue 70R1, which is *i*-1, does not exhibit any broadening, thus it is possible that the spectrum of residue 71 has contributions from weak interaction between the nitroxide side chain and the iron due to its closer distance to the porphyrin ring. To investigate if this is the case, myoglobin in the apo form was reconstituted with zinc-porphyrin, which is diamagnetic, and

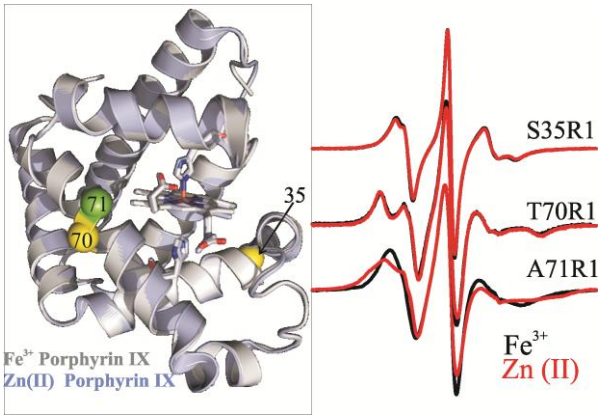


Figure 4.11: Exchange interaction between the R1 side chain at site 71 and iron atom of the heme porphyrin. *Left panel*: Overlay of the x-ray structures of iron (gray) and zinc-substituted (lightblue) myoglobin. *Right panel*: EPR spectra of the indicated sites in the iron and zinc-susbtituted myoglobin variants.

the EPR spectrum of residue 71R1 was recorded. As a control, the spectra of residues 35R1 and 70R1, both of which are not in close proximity to the porphyrin ring, were also recorded in the zinc-substituted protein. As shown in Figure 4.11, there are significant changes in the spectrum of residue 71R1 in the zinc-bound form, while the spectra of residues 35R1 and 70R1 are indistinguishable in both forms[7]. In addition, saturation recovery experiment on site 71R1 revealed faster T_1 of the nitroxide in the ferric form compared to the zinc-bound form (M. Bridges and W.L. Hubbell, unpublished). The above results combined support the contention that the lineshape of residue 71R1 has contributions from exchange interaction with the ferric ion.

[7] In a recent study, the high-resolution structure of the zinc-substituted form of myoglobin was solved (Koshiyama et al. [62]); as shown in Figure 4.11A, there are little structural differences between the iron and zinc-substituted forms, thus any spectral changes are unlikely to arise due to structural differences of the protein

4.5 Appendix

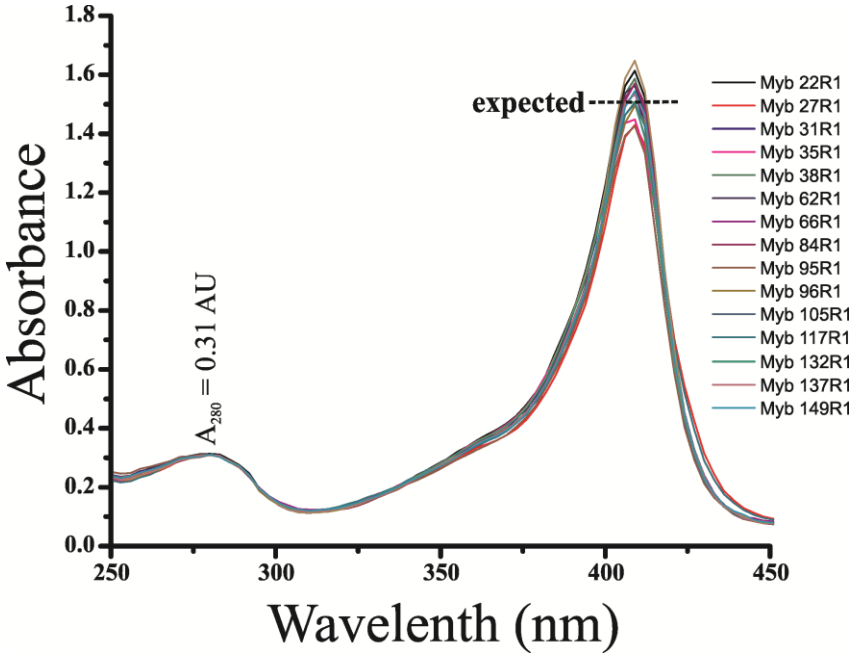


Figure 4.A.1: UV-Vis scan of a set of R1 mutants of myoglobin reconstituted with bovine heme. The spectra were normalized to the same protein concentration (10 μM) as judged by the absorption at 280 nm. The values of ε_{409} and ε_{280} are 157,000 and 31,000 M^{-1} cm^{-1}, respectively. The expected A_{409}/A_{280} is ~ 5.

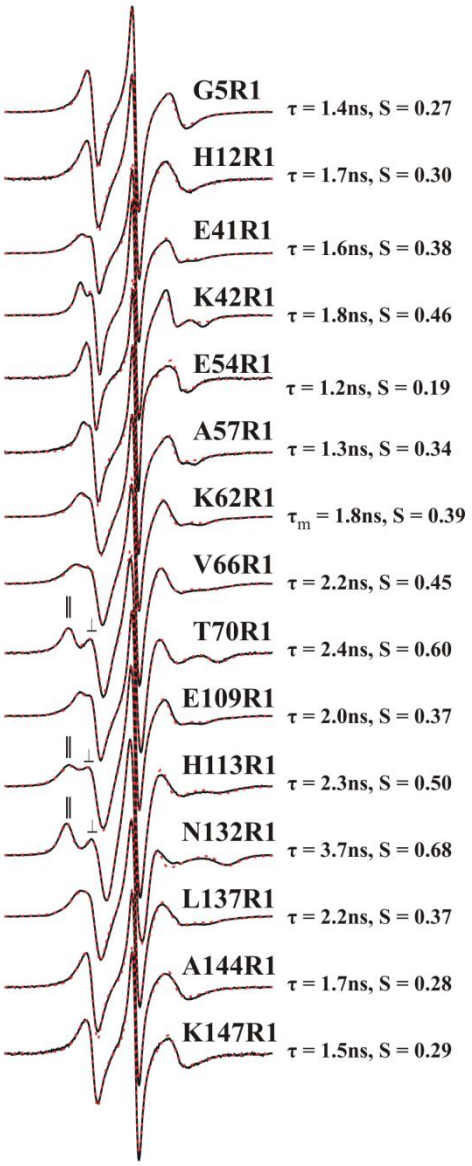


Figure 4.A.2: MOMD fits of a set EPR spectra of R1 at solvent-exposed sites of metmyoglobin with parameters obtained from the MOMD fits. The parameters reported for residue 62R1 are those of the more mobile state.

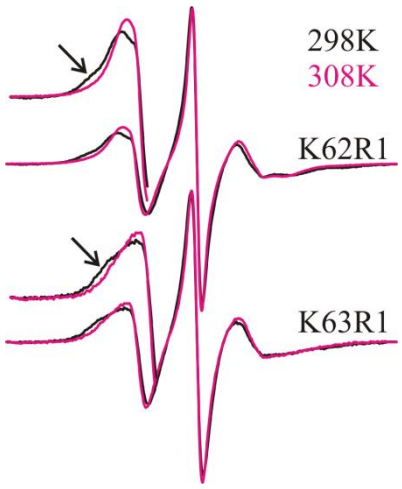


Figure 4.A.3: EPR spectra of residues 62R1 and 63R1 recorded at 298K (black) and 308K (magenta). The reduction in the relative population of the immobile component is highlighted. The low field region of the spectra was amplified for clarity.

4.6 References

1. Mittermaier, A, and Kay. L.E. 2006. New tools provide new insights in NMR studies of protein dynamics. *Science* 312:224-228.
2. Wales, T.E., and Engen, J.R. 2006. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spec. Rev*. 25:158-170.
3. Weiss, S. 2000. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat. Struc. Mol. Biol*. 7:724-729.
4. Hubbell, W.L., Mchaourab, H.S., Altenbach, C., and Lietzow, M.A. 1996. Watching proteins move using site-directed spin labeling. *Structure* 4:779-783.
5. Hubbell, W.L., Cafiso, D.S., and Altenbach, C. 2000. Identifying conformational changes with site-directed spin labeling. *Nat. Struc. Biol*. 7:735-739.
6. Altenbach, C., Yang, K., Farrens, D.L., Farahbakhsh, Z.T., Khorana, H.G., and Hubbell, W.L. 1996. Structural features and light-dependent changes in the cytoplasmic

interhelical E-F loop region of rhodopsin: A site-directed spin-labeling study. *Biochemistry* 35:12470-12478.

7. Farahbakhsh, Z.T., Ridge, K.D., Khorana, H.G., and Hubbell, W.L. 1995. Mapping light-dependent structural changes in the cytoplasmic loop connecting helices C and D in rhodopsin: A site-directed spin labeling study. *Biochemistry* 34:8812-8819.
8. Farrens, D., Altenbach, C., Yang, K., Hubbell, W.L. & Khorana, H.G. 1996. Requirement of rigid-body motions of transmembrane helices for light activation of rhodopsin. *Science* 274:768–770.
9. Altenbach, C., Kusnetzow, A.K., Ernst, O.P., Hofmann, K.P., and Hubbell, W.L. 2008. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Nat. Acad. Sci.* 105:7439-7444.
10. Bridges, M.D., Hideg, K., and Hubbell, W.L. 2010. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson*. 37:363-390.
11. McCoy, J. and Hubbell, W.L. 2011. High-pressure EPR reveals conformational equilibria and volumetric properties of spin-labeled proteins. *Proc. Nat. Acad. Sci.* 108:1331-1336.
12. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011 Structure and dynamics of a conformationaly constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246.
13. Columbus, L. and Hubbell, W.L. 2002. A new spin on protein dynamics. *Trends Bioch. Sci.* 27:288-295.

14. Columbus, L. and Hubbell. W.L. 2004. Mapping backbone dynamics in solution with site-directed spin labeling:GCN4-58 bZip free and bound to DNA. *Biochemistry*, 43:7273:7287.
15. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35:7692-7704.
16. Langen, R., Oh, K.J., Cascio, D., and Hubbell, W.L. 2000. Crystal structures of spin labeled T4 lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry*, 39:8396-8405.
17. Fleissner, M.R., Cascio, D., and Hubbell, W.L. 2009. Structural origins of weakly ordered motion in spin-labeled proteins. *Protein Sci.* 18:893-908.
18. Kroncke, B.M., Horanyi, P.S., and Columbus, L. 2010. Structural origins of nitroxide side chain dynamics on membrane proteins α-helical sites. Biochemistry 49:10045-10060.
19. Columbus, L., Kálai, T., Jeko, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in alpha-helices:analysis by variation of side chain structure. *Biochemistry* 40:3828-3846.
20. Mchaourab, H.S., Kálai, T., Hideg, K., and Hubbell, W.L. 1999. Motion of spin-labeled side chains in T4 lysozyme. Effect of side chain structure. *Biochemistry* 38:2947-2955.
21. Warshaviak, D.T., Serbulea, L., Houk, K.N., Hubbell, W.L. 2011. Conformational analysis of a nitroxide side chain in an α-helix with density functional theory. *J Phys Chem B* 115:397-405.

22. Ösapay, K., Theriault, Y., Wright, P.E., and Case, D.E. 1994. Solution structure of carbomonoxy myoglobin determined from nuclear magnetic resonance distance and chemical shifts constraints. *J. Mol. Biol*. 244:183-197.

23. Bismuto, E., Irace, G., and Gratton, E. 1989. Multiple conformational state of myoglobin revealed by frequency domain fluorometry. *Biochemistry* 28:1508-1512.

24. Cavagnero, S., Theriault, Y., Narula, S.S., Dyson, H.J., and Wright, P.E. 2000. Amide proton exchange rates for sperm whale myoglobin obtained from ^{15}N-^{1}H NMR spectra. *Protein Sci*. 9:186-193.

25. Eliezer, D., Yao, J., Dyson, H.J., and Wright, P.E. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struc. Biol*. 5:148-155.

26. Halle, B. 2002. Flexibility and packing in proteins. *Proc. Nat. Acad. Sci.* 99:1274-1279.

27. Zhang, F. and Brüschweiler, R. 2002. Contact model for the prediction of NMR N-H order parameters in globular proteins. *J. Am. Chem. Soc.*, 124:12654-12655.

28. Kendrew, J.C., Dickerson, R.E., Strandberg, D.E., Hart, R.G., Davies, D.R., Phillips, D.C., and Shore, V.C. 1960. Structure of myoglobin. Nature 185:422-427.

29. George, P. and Stratmann, C.J. 1954. The oxidation of myoglobin to metmyoglobin by oxygen. *Biochem. J*. 57:568-573.

30. Springer, B.A., Egeberg, K.D., Rholfs, R.J., Matthews, A.J., Olson, J.S., and Sligar, S.G. 1989. Discrimination between oxygen and carbon monoxide and inhibition of autooxidation by myoglobin. *J. Biol. Chem*. 264:3057-3060.

31. Austin, R.H., Beeson, K.W., Einstein, L., Frauenfelder, H., and Gunsalus, I.C. 1975. Dynamics of ligand binding to myoglobin. *Biochemistry* 14:5355-5373.

32. Frauenfelder, H., Sligar, S.G., and Wolynes. P.G. 1991. The energy landscapes and motions on proteins. *Science* 254:1598-1603.

33. Šrajer, V., Ren, Z., Teng, T.Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M., and Moffat K. 2001. Protein conformational relaxation and ligand migration in myoglobin: A nanosecond to millisecond molecular movie from time-resolved laue x-ray diffraction. *Biochemistry* 40:13802-13815.

34. Tomita, A., Kreutzer, U., Adachi, S., Koshihara, S., Jue, T. 2010. 'It's hollow': the function of pores within myoglobin. *J. Exp. Biol*. 213:2748-2754.

35. Guo, Z.F., Cascio, D., Hideg, K., Kalai, T., and Hubbell, W.L. 2007 Structural determinants of nitroxide motion in spin-labeled proteins: tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. *Protein Sci*. 16:1069–1086.

36. Guo, Z.F., Cascio, D., Hideg, K., and Hubbell, W,L. 2008. Structural determinants of nitroxide motion in spin-labeled proteins: solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci*. 17:228–239.

37. Hargrove, M.S., Barrick, D., and Olson, J.S. 1996. The association rate constant for heme binding to globin is independent of protein structure. *Biochemistry* 35:11293-11299.

38. Liong, E.C., Dou, Y., Scott, E.E., Olson, J.S., and Phillips, G.N. 2001. Waterproofing the heme pocket. *J. Biol. Chem*. 276:9093-9100.

39. Phillips, G.N. 1990. Comparison of the dynamics of myoglobin in different crystal forms. *Biophys. J*. 57:381-383.

40. Frauenfelder, H., Hartmann, H., Karplus, M., Kuntz, I. D., J. Kuriyan, Parak, F., Petsko, G.A., Ringe, D., Tilton, R. F., Connolly, M. L., and Max, N. 1987. Thermal expansion of a protein. *Biochemistry* 26:254-261.

41. Barrick, D., Hughson, F. M., and Baldwin, R. L. 1994. Molecular mechanism of acid denaturation: The role of histidine residues in the partial unfolding of apomyoglobin. *J. Mol. Biol. 237*:588-601.

42. Shu, F., Ramakrishnan, V., and Schoenborn. 2000. Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proc. Nat. Acad. Sci.* 97:3872-3877.

43. Fraczkiewicz, R. and Braun, W. 1998. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem.* 19:319-333.

44. Fleissner, M.R. 2007. X-ray structures of nitroxide side chains in proteins: a basis for interpreting distance measurements and dynamic studies by electron paramagnetic resonance. PhD thesis, University of California, Los Angeles.

45. Columbus, L. 2001. Investigating backbone and side chain dynamics of α-helices in the nanosecond regime by site-directed spin labeling. PhD thesis. University of California, Los Angeles.

46. Lietzow, M.A., and Hubbell, W.L. 2004. Motion of spin label side chains in cellular retinol-binding protein: Correlation with structure and nearest-neighbor interactions in an antiparallel β–sheet. *Biochemistry* 43:3137-3151.

47. Brucker, E.A., Olson, J.S., Phillips Jr., G.N., Dou, Y., and Ikeda-Saito, M. 1996 High resolution crystal structures of the deoxy, oxy, and aquomet forms of cobalt myoglobin. *J.Biol.Chem*. 271: 25419-25422

48. Hartmann, H., Parak, F., Steigemann, W., Petsko, G.A., Ponzi, D.R., and Frauenfelder, H. 1982. Conformational substates in a protein:Structure and dynamics of metmyoglobin at 80K. *Proc. Nat. Acad. Sci.* 79:4967-4971.

49. Arakawa, T. and Timasheff, S.N. 1985. The stabilization of protein by osmolytes. Biophys J. 47:411-414.

50. Liu, Y. and Bolen, W.D. 1995. The peptide backbone plays a dominant role in protein stabilization in naturally occurring osmolytes. *Biochemistry* 34:12884-12891.

51. Garcia-Moreno, B.E., Chen, L.X., March, K.L., Gurd, R.S., and Gurd, F.R.N. 1985. Electrostatic interaction in sperm whale myoglobin. *J. Biol. Chem*. 260:14070-14082.

52. Holladay. L.A. 1986. Thermodynamics of carp metmyoglobin folding. *Comp. Biochem. Physiol*. 83B:365-370.

53. Mittermaier, A., Kay, L.E., and Forman-Kay, J.D. 1999. Analysis of deuterium relaxation-derived methyl axis order parameter and correlation with local structure. *J. Biomol. NMR* 13:181-185.

54. Goodman, J.L., Pagel, M.D., and Stone, M.J. 2000. Relationship between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J. Mol. Biol*. 295:963-978.

55. Huang, S.H., Shih, C.H., Lin, C.P., and Hwang, J.K. 2008. Prediction of NMR order parameters in proteins using weighted protein contact number. *Theor. Chem. Acc*. 121:197-200.

56. Sezer, D., Freed, J.H., and Roux, B. 2008. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine α–helix. *J. Phys. Chem*. 112:5755-5767.

57. Seezer, D., Freed, J.H., and Roux, B. 2009. Multifrequency electron spin resonance of a spin-labeled protein calculated from molecular dynamics simulations. *J. Am. Chem. Soc.* 131:2597-2605.

58. Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A.E., Sciara, G., Wulff, M., Anfinrud, P., and Brunori, M. 2003. Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography. *Proc. Nat. Acad. Sci.* 100:8704-8709.

59. Phillips, S.E.V. 1980. Structure and refinement of oxymyoglobin at 1.5Å resolution. *J. Mol. Biol*. 142:531-554.

60. Kachalova, G.A., Popov, A.N., and Bartunik, H.D. 1999. A steric mechanism for inhibition of CO binding to heme proteins. *Science* 284-473-476.

61. Jennings, P.A., Stone, M.J., and Wright, P.E. 1995. Overexpression of myoglobin and assignments of its amide, C^{α} and C^{β} resonances. *J. Biomol. NMR* 6:271-276.

62. Koshiyama, T., Shirai, M., Hikage, T., Tabe, H., Tanaka, K., Kitagawa, S., Ueno, T. 2011. Post-crystal engineering of zinc-substituted myoglobin to construct a long-lived photoinduced charge-separation system *Angew.Chem.Int.Ed.Engl.* 50:4849-4852.

63. Bracken, C., Carr, P.A. Cavanagh, J., and Palmer, A.G. 1999. Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA. *J. Mol. Biol*. 285:2133-2146.

64. Morin, B., Bourhis, J.M., Longhi, S. 2006. Assessing induced folding of an *IDP* by SDSL EPR spectroscopy. *J. Phys. Chem. B.* 110:20596–20608.

65. Belle, V., Rouger, S., Costanzo, S., Liquiere, E., Strancar, J., Guigliarelli, B., Fournel, A., Longhi. S 2008 Mapping alpha-helical induced folding within the intrinsically

disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins J.* **73***:*973–988.

Chapter 5: Mapping molecular flexibility of partially folded states of proteins with site-directed spin labeling: a case study of myoglobin

5.1 Summary

It has been recently recognized that a substantial fraction of the proteome consists of functional sequences that lack stable secondary and tertiary folds. Thus, a detailed structural and dynamic characterization of natively unfolded and partially folded proteins has become increasingly important for understanding the molecular mechanisms of protein function. CW EPR in combination with the osmotic perturbation (OP) method introduced in chapter 3 of this dissertation has the potential of providing functionally relevant site-specific information of fast backbone motions and conformational equilibria of natively unfolded and partially folded proteins. Sperm whale apomyoglobin provides an excellent model system to demonstrate the utility of CW lineshape analysis and OP to characterize such systems as the protein populates multiple unfolded and partially folded states at equilibrium conditions, each of which has been extensively characterized with respect to solution dynamics by NMR. The data shown here illustrate how slow conformational exchange (μs to ms time scale) reveals itself in the CW EPR spectra in partially folded states and provides a sequence-specific map of the changes in the conformational properties that is in excellent agreement with that from NMR. As a result of the intrinsic EPR time domain, the EPR data provide additional insight into the location and conformational properties of sequences of the molecule in the N_{Apo} state for which direct information was not accessible by other methods.

5.2 Introduction

For many decades the prevalent view in structural biology was the requirement of a stable three-dimensional folded structure of a protein for function. This long-standing paradigm has been challenged by a myriad of studies done in the last 20 years which have revealed that a large portion of protein sequences do not fold into a stable 3D structure at physiological conditions (*i.e.,* 25 °C and pH ~ 7) [1]. In fact, sequence analysis of complete genomes of prokaryotes, archaea, and unicellular and multicellular eukaryotes suggests that unfolded and partially folded proteins are very common with the more complex organisms showing the higher proportion of intrinsic protein disorder [2-3]. Protein disorder can extend from local sequences involving a few amino acids to entire domains or involve global disorder [4]. Most importantly, natively unfolded and partially folded protein sequences have been implicated in fundamental biological processes such as signal transduction, cell cycle regulation, transcription, and catalysis [4-7]. There are several mechanistic interpretations for the role of dynamically disordered protein sequences in protein function that include: (1) facilitating interaction with multiple targets (molecular promiscuity) [8], (2) reducing sensitivity to environmental conditions, (3) increasing the rate of binding to other macromolecules *via* "fly-casting" mechanism [9], and (4) maintaining a large intermolecular interface to volume ratio [10]. Thus, to understand the mechanisms of protein function involving unfolded and partially folded sequences, it is necessary to develop experimental techniques capable of characterizing the structural and conformational properties of such states.

Current biophysical methods that have been employed to identify and characterize unfolded and partially folded protein sequences include X-ray crystallography, CD spectroscopy

[11], fluorescence [12], and solution NMR spectroscopy [13]. In X-ray crystallography, only **local** protein disorder can be inferred *via* the lack of electron density of protein segments, however the intrinsic structural and dynamic heterogeneity of globally unfolded and partially folded proteins precludes their structural characterization by X-ray crystallography. On the other hand, solution NMR methods have the advantage of providing site-specific information on the structure and conformational properties of dynamically disordered protein sequences. In particular, the development of heteronuclear NMR techniques has been fundamental in overcoming the difficulty of poor resonance dispersion of protons in natively unfolded proteins. Particularly useful metrics to study natively unfolded protein sequences *via* NMR are the chemical shift values, which can provide site-specific information on the secondary structure and ^{15}N-^{1}H relaxation measurements, which gives insights into the local dynamics of the backbone [13]. Interestingly, the number of reports describing the structural and dynamic characterization of partially folded states such molten globule (MG) states using NMR spectroscopy is limited likely as a result of intrinsic technical challenges. For example, MG states often exhibit poor chemical shift dispersion and are highly heterogeneous with structural interconversion between substates in the ms time scale, which often leads to extreme broadening of the NMR resonance lines [13]. In addition, MG states often aggregate at the concentrations required for NMR. In favorable cases, the broad resonances and aggregation problems can be solved by the use of high temperatures and ethanol [14], however these conditions can, in principle, affect the structural and conformational properties of the protein. SDSL-EPR offers an alternative strategy to characterize the molecular flexibility of unfolded and partially folded states in solution at physiological temperatures and solvent conditions. Indeed, SDSL-EPR has been employed to identify and characterize unfolded protein sequences in several proteins, including membrane

proteins [15-17]. The study presented here aims at showing the utility of osmotic-perturbation EPR in combination with CW lineshape analysis and interspin distance measurements to provide a sequence-specific map of fast backbone dynamics and regions in slow conformational exchange of a protein in multiple partially folded states populated at equilibrium conditions.

Myoglobin is an ideal and unique model system for this purpose because conditions have been established wherein the molecule populates folded and partially folded states under equilibrium conditions suitable for biophysical characterization. The states are: (1) the native holo (N_{Holo}), which is the functional form containing the heme ligand described in the previous chapter; (2) native apo protein (N_{Apo}); (3) the molten globule intermediate (I_{MG}), and (4) the acid unfolded state (U_{Acid}). High resolution crystal structures are available for the holo form [18-19], while solution NMR [20-28] and CD spectroscopy [29], along with small angle X-ray scattering [30], and mass spectrometry [31] have provided a wealth of information on the structure and dynamics of the protein in solution for each of the states.

In chapter 3 of this dissertation, the F-helix sequence of myoglobin in the N_{Apo} state was studied to evaluate the potential of osmotic perturbation EPR as a tool to identify conformational equilibria in spin-labeled proteins. In the preceding chapter, CW lineshape analysis and osmotic perturbation EPR was employed to map the flexibility of myoglobin in the well-ordered N_{Holo} state. The data presented in this chapter serves as an extension of both studies. The molecular flexibility of apomyoglobin, as seen through the EPR time window, was monitored in the N_{Apo}, I_{MG}, U_{acid} states. Spectral changes for two of the transitions (*i.e.*, $N_{Holo} \rightarrow N_{Apo}$, $N_{Apo} \rightarrow I_{MG}$) were analyzed to get insights into the site-specific progressive increase in backbone (ns time scale) and conformational flexibility (μs-ms time scale). In addition, distance measurements

between a set of R1 pairs were done using double electron-electron resonance (DEER) spectroscopy to provide additional insights into the structural changes that follow the $N_{Holo} \rightarrow N_{Apo} \rightarrow I_{MG} \rightarrow U_{acid}$ transitions.

5.2.1 Background of structural and dynamic characterization of partially folded states of apomyoglobin by other methods. Myoglobin has been the subject of many structural and kinetics studies and has served as a paradigm to study protein folding [32-33]. As discussed in the previous chapter, in its functional form, myoglobin contains a heme group that enables oxygen binding. Removal of the heme group to generate the apomyoglobin form leads to a more unstable protein with lower helical content [29,34]. In early studies it was suggested that apomyoglobin at neutral pH exhibits characteristics of a molten globule state (*i.e.,* compact with native-like secondary structure, but lacking a stable tertiary fold) [35]. However, in subsequent NMR studies done by Eliezer and Wright et al. [31] it was shown that apomyoglobin exhibits holo-like secondary and tertiary structures with dynamic disorder in local sequences of the backbone as judged by broadening of the NMR resonance lines. Specifically, the regions that become dynamically disordered in the absence of heme are the sequences corresponding to the E-F turns, F-helix, F-G turns, first turn of G-helix, and the C-terminal region of the H helix as show in Figure 5.1. Interestingly, as the pH of the solution is lowered to 4 in the absence of salt, apomyoglobin populates a partially folded equilibrium molten globule state (I_{MG}) containing a helix content of ~ 35 % [34] and a higher hydrodynamic ratio compared to the native state [30]. Hydrogen/deuterium exchange experiments [34], $^{13}C\alpha$ chemical shifts values, and ^{15}N relaxation data on the I_{MG} state revealed that the sequences corresponding to helices A, B, G, and H are folded and form the hydrophobic core as depicted in Figure 5.1. In addition, ^{15}N relaxation data revealed that the core helices exhibit restricted backbone flexibility, while the rest of the

sequence exhibits lower helical propensity with high backbone flexibility [22,25]. Remarkably, the aforementioned equilibrium intermediate closely resembles an obligatory kinetic intermediate formed within the first 6 ms during apomyoglobin folding [32-33,36,38], which makes the I_{MG} state of apomyoglobin an excellent model to study the structure and dynamics of the polypeptide during the intermediate states of protein folding.

Upon further reduction of the pH of the solution to 2 at low salt conditions, apomyoglobin populates an unfolded state (U_{Acid}) with native-like helical propensity in the regions corresponding to helices A and H, and a non-native helix corresponding to the end of the D and the beginning of the E helices in the native protein [22,26]. Remarkably, paramagnetic relaxation enhancement (PRE) and ^{15}N relaxation data revealed the formation of local hydrophobic clusters and long-range native-like tertiary interactions in the U_{Acid} state, which may play an important role in the initial folding process of myoglobin [26,37].

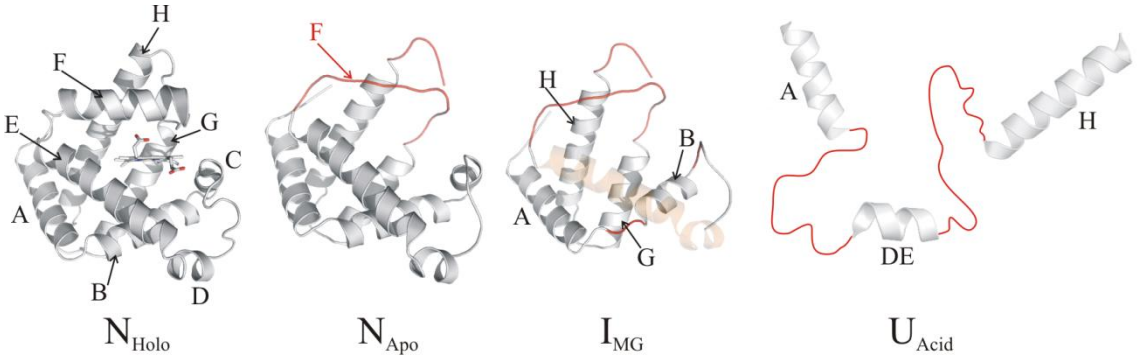


Figure 5.1: Diagrammatic representation of the structure of myoglobin in each of the folded and partially folded states populated at equilibrium conditions based on NMR data (Eliezer et al. [22]). The dynamically disordered sequences are shown in red. The location of each of the helices is indicated in the X-ray structure of the holo protein shown in the left panel (PDB ID: 2mbw).

5.3 Results:

5.3.1 *Characterization of the R1 mutants.* For this study, 41 cysteine substitutions at solvent exposed sites were engineered throughout α-helices A to H and the A/B, E/F and G/H turns, 38 of which are the same as those discussed in the previous chapter. Substitution of R1 at surface sites of myoglobin caused little change to the structural properties of myoglobin in the states studied here as judged by Far UV CD spectroscopy done in this work (*see* Appendix) and in previous studies [39]. As shown in the previous chapter, cysteine mutations of the surface residues selected here did not impair heme binding to the apo protein.

Purification of the myoglobin mutants in the apo form (*i.e.*, no heme) involves isolation of the protein from inclusion bodies followed by *in vitro* refolding [40] (*see* Chapter 8). Since the amino acid sequence defines the folding pathway and the final three-dimensional structure of a protein [41], mutations of native residues can alter the folding mechanism favoring alternative pathways, which may result in the formation of misfolded and aggregated species. Thus, it is essential to detect and remove such aggregates after *in vitro* refolding. To identify and remove any soluble aggregates following *in vitro* refolding of the apo myoglobin mutants, the samples were subjected to size exclusion chromatography (Figure 5.2) using a Superdex 75 column (*see* Chapter 8). The WT protein was also injected in the column as a control. As shown in Figure 5.2A, the WT sample eluted from the column at an elution volume (V_e) of ~ 12.6 ml, corresponding to the monomeric species (based on calibration curves and SDS-PAGE electrophoresis). Remarkably, the profiles of size exclusion chromatography for some of the mutants clearly revealed the presence of a second, higher molecular weight, but soluble species eluting at the void volume as shown in Figure 5.2B. Denaturing SDS-PAGE electrophoresis

revealed that the molecular mass of the species eluting at the void volume corresponded to that of the monomer under denaturing conditions (Figure 5.2C). It is emphasized that refolding was done in the presence of DTT, hence eliminating disulfide formation between monomers as the source of misfolding and subsequent oligomerization. Interestingly, 3 of the 42 cysteine mutants (*e.g.*, R118C, A127C, and K133C) resulted in 100% oligomer formation (Figure 5.2D) and were not further investigated.

The misfolded oligomeric species observed for some of the mutants in this work are likely different from the soluble aggregates observed in earlier studies, which revealed that *in vitro* refolding of WT apo myoglobin from inclusion bodies results in the formation of a significant fraction of aggregates with an average M_W > 1MDa that were undetected in gel filtration purification due to adsorption on the column matrix [42]. It is noted that the *soluble* aggregates contribute significantly to the EPR spectra of the spin-labeled samples (*see* Appendix), which shows the importance of removing aggregated species prior to EPR data collection. For each case presented hereon, the EPR spectra are derived from pure monomeric species purified by size exclusion chromatography.

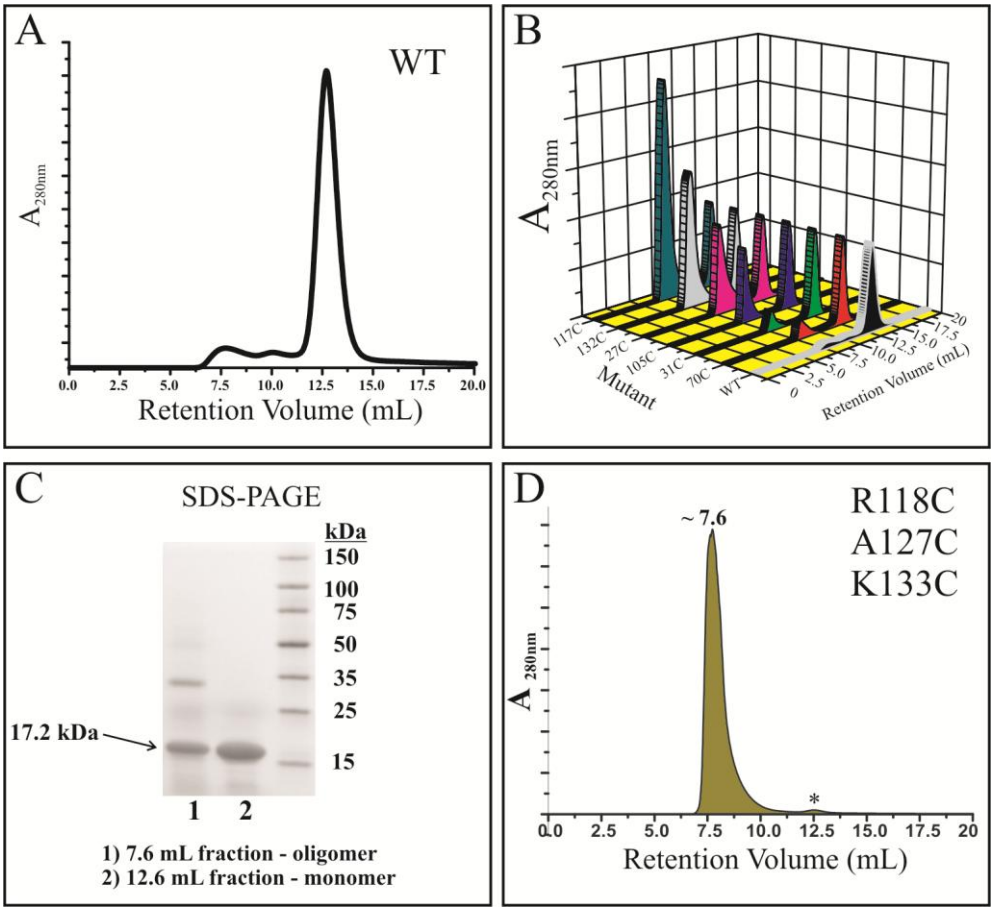


Figure 5.2: Elution profile from size exclusion chromatography of WT and cys myoglobin mutants. A) WT apomyoglobin. B) 3D-graph showing the elution profiles of WT and selected set of mutants injected in the Superdex 75 column. C) SDS-PAGE of fractions eluted from the size exclusion column. D) Elution profiles of mutants that completely misfolded. The asterisk indicates the retention volume of the monomeric protein.

5.3.2 *Mapping backbone flexibility of myoglobin in N_{Apo} state.* The structure and dynamics of myoglobin in the N_{Apo} state in solution have been previously characterized by multidimensional NMR and CD spectroscopy. These studies revealed that myoglobin in the N_{Apo} state is destabilized relative to the holo protein by about 5.5 kcal/mol [29], has a reduction in the overall helical content (~20%) [22,43], and adopts a tertiary structure that closely resembles the holo protein [21-23]. Based on sequence-specific resonance assignments, residues 78-106 plus residues 146-153 corresponding to the EF turn, F-helix, FG turn, the beginning of the G-helix,

and the C-terminal region of the H-helix, respectively are apparently in conformational exchange (microseconds to milliseconds time scale) in N_{Apo} state as inferred from NMR line broadening, however there is no direct information available on the structures of the substates involved due to complete loss of resonances in these regions [21]. A key advantage of SDSL is that slow exchange in the microseconds to milliseconds time window does not cause loss of spectra information, which makes SDSL particularly attractive for obtaining *direct* information on the substates involved.

In the previous chapter the molecular flexibility of myoglobin in the N_{Holo} state as seen through the EPR time domain was discussed. To map the changes in the conformational flexibility that arise due to heme removal ($N_{Holo} \rightarrow N_{Apo}$ transition) *via* SDSL-EPR, the EPR spectra of the same 38 sites studied in the N_{Holo} state were recorded in the N_{Apo} state in Ficoll 70 at pH 6.1. The entire set of spectra is given in the Appendix section, and a subset shown in Figure 5.3 is compared to the same sites in N_{Holo} to illustrate the nature of the changes. The EPR spectra of residues 41R1, 66R1, 84R1, 87R1, 91R1 and 96R1 in the N_{Apo} state have been reported elsewhere (refer to chapter 3) [39] and are reproduced here. Remarkably, 24 of the 38 sites in N_{Apo} show multicomponent spectra compared to only 10 for N_{Holo} state. In most cases, a new component appears corresponding to an immobilized state of the nitroxide reflecting tertiary interaction with the environment. Examples in Figure 5.3A include residues 42R1, 78R1, 109R1, 140R1, 147R1 and 149R1 located in the helices identified in the Figure; other similar sites are, 41R1, 53R1, 96R1, 106R1, 137R1, and 144R1 (*see* Appendix).

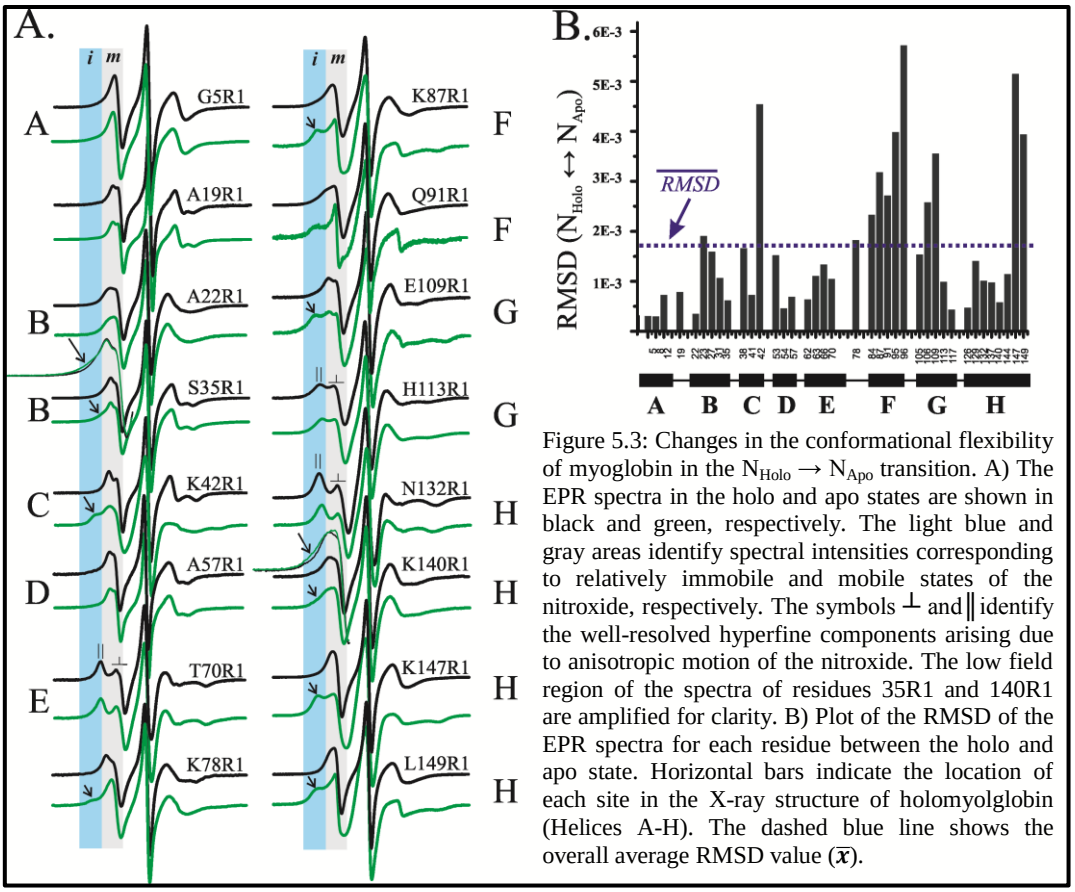


Figure 5.3: Changes in the conformational flexibility of myoglobin in the $N_{Holo} \rightarrow N_{Apo}$ transition. A) The EPR spectra in the holo and apo states are shown in black and green, respectively. The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively. The symbols ⊥ and ∥ identify the well-resolved hyperfine components arising due to anisotropic motion of the nitroxide. The low field region of the spectra of residues 35R1 and 140R1 are amplified for clarity. B) Plot of the RMSD of the EPR spectra for each residue between the holo and apo state. Horizontal bars indicate the location of each site in the X-ray structure of holomyolglobin (Helices A-H). The dashed blue line shows the overall average RMSD value ($\bar{x}$).

Interestingly, the more mobile component in the spectra of residues in helix F (87R1, 91R1, 95R1) and the C terminus of helix H (149R1) reflect relatively high mobility of the backbone (τ ~ 1 ns; refer to chapter 3), which may arise from local unfolding or from a helical sequence exhibiting high amplitude of backbone motions (*see* Discussion).

A simple measure to illustrate the site-specific differences between the N_{Holo} and the N_{Apo} is the root mean square difference (RMSD) of the normalized EPR spectral lineshapes. A plot showing the RMSD per residue is shown in Figure 5.3B; the average value for all sites is shown by the dotted blue line. Above average values of RMSD are observed in four regions of the molecule, namely, residue 23 on helix B, residue 42R1 on helix C, residues 78-109, which

correspond to the EF turn, F-helix, the FG turn, and the first half of G-helix, and residues 147-149 corresponding to the C-terminal region of the H-helix.

To identify which of the multicomponent spectra observed in the N_{Apo} state arise from conformational exchange, the EPR spectra of all such sites recorded in Ficoll were compared to those in sucrose solutions. The EPR spectra of all the multicomponent sites showing sensitivity to osmotic-perturbation are shown in Figure 5.4A. The complete data set is provided in the Appendix section. Fourteen of the 24 multicomponent spectra showed a clear shift in population from a more mobile to immobile state, identifying those sites as residing in regions of conformational exchange. Interestingly, most of these sites are in the same sequences that exhibit significant spectral changes between the holo and apo states as shown in Figure 5.4B, indicating that the spectral changes observed in the holo→apo transition arise from an increase in the conformational flexibility of myoglobin in the apo form. As expected, osmolyte perturbation has little effect on sites that have a single component (*see* Appendix) although there are subtle changes in mobility of R1 in a few cases that could reflect the effect of osmolyte on the fast (ns) motions of the backbone.

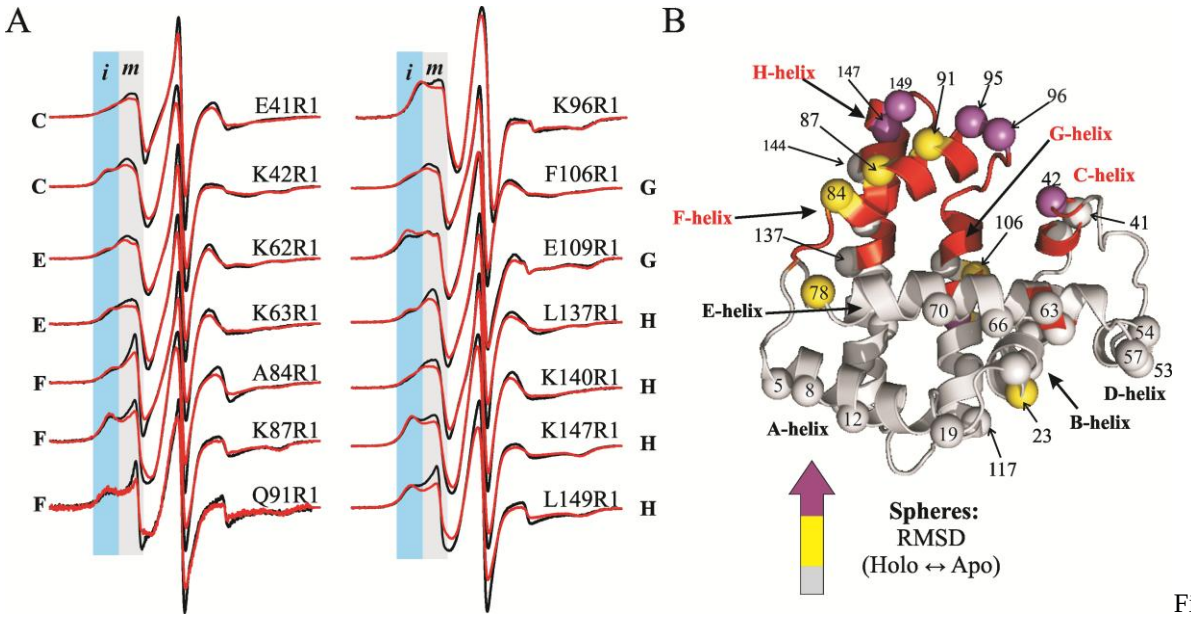


Figure 5.4: Conformational equilibria in myoglobin in the N_{Apo} state. A) EPR spectra of R1 at sites exhibiting sensitivity to osmotic-perturbation. The spectra recorded in Ficoll 70 and sucrose are shown in black and red, respectively. The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively. B) Ribbon diagram of myoglobin showing the sites with above average RMSD values. The spheres at the Cα have been color coded based on the magnitudes of RMSD values shown in Figure 5B. $x < \overline{RMSD}$ (gray), $\overline{RMSD} \leq x \leq \overline{RMSD} + \boldsymbol{\sigma}$ (yellow), and x > $\overline{RMSD} + \boldsymbol{\sigma}$ (magenta). Red backbone indicates regions in conformational exchange as judged by osmotic perturbation.

5.3.3 *Conformational flexibility of myoglobin in the I_{MG} state***.** As mentioned above, earlier studies have shown that apomyoglobin at pH 4 in the absence of salt populates a compact equilibrium molten globule state (I_{MG}) that strongly resembles the obligatory kinetic intermediate formed within 6 ms of the folding pathway of myoglobin [25,32]. Multidimensional NMR studies have revealed that helices A, G, H, and the C-terminal region of the B helix are folded and form a compact hydrophobic core in the I_{MG} state, whereas the rest of the protein is conformationally heterogeneous (save for the sequence corresponding to the F-helix in the N_{Holo} state, which behaves as a free flight random coil) [22,25]. NMR relaxation experiments also showed that the backbone flexibility (subnanosecond) for the ABGH core region is restricted, but possibly undergoing slower conformational exchange processes [22,25]. Due to poor spectral resolution, the NMR studies mentioned above were performed at high temperature (323 K), which may affect the conformational properties of the protein. SDSL is well-suited to provide

complementary information on the conformational flexibility of myoglobin in the I_{MG} sate with site-specific resolution at room temperature (298 K) and at 323 K.

To map the changes in conformational flexibility in the transition from N_{Apo} to I_{MG} with SDSL, the pH of the protein solution for 36 of the spin-labeled mutants was lowered to 4.1 to populate the molten globule state and the EPR spectra were recorded at 298 K. Remarkably, the EPR spectra of R1 at every site studied in the I_{MG} state revealed multiple dynamic states of R1, some of which reflect a single component in the N_{Holo} and N_{Apo} states. A subset of the spectra is shown in Figure 5.5A along with a comparison of the corresponding spectra in the N_{apo} form to illustrate the nature of the spectral changes observed in the $N_{Apo} \rightarrow I_{MG}$ transition; the complete set of spectra for I_{MG} is given in the Appendix. In addition to the appearance of new immobilized states at sites in helices A, B, G, and H, a sharp component reflecting a high degree of mobility on the nanosecond time scale ($\tau \leq 1$ ns; Figure 5.6) appears or is increased in population in the spectra for sites within the C, D, E, and F helices (*e.g.,* 41, 57R1, 66R1, 70R1, 84R1). This component has a lineshape and mobility similar to that of R1 in unstructured protein sequences [44-45], suggesting that these helices sample unfolded conformations. On the other hand, the more mobile component observed for residues in the rest of the sequence (helices A, B, G, and H) reflect a more restricted motion of the backbone in the nanosecond time scale.

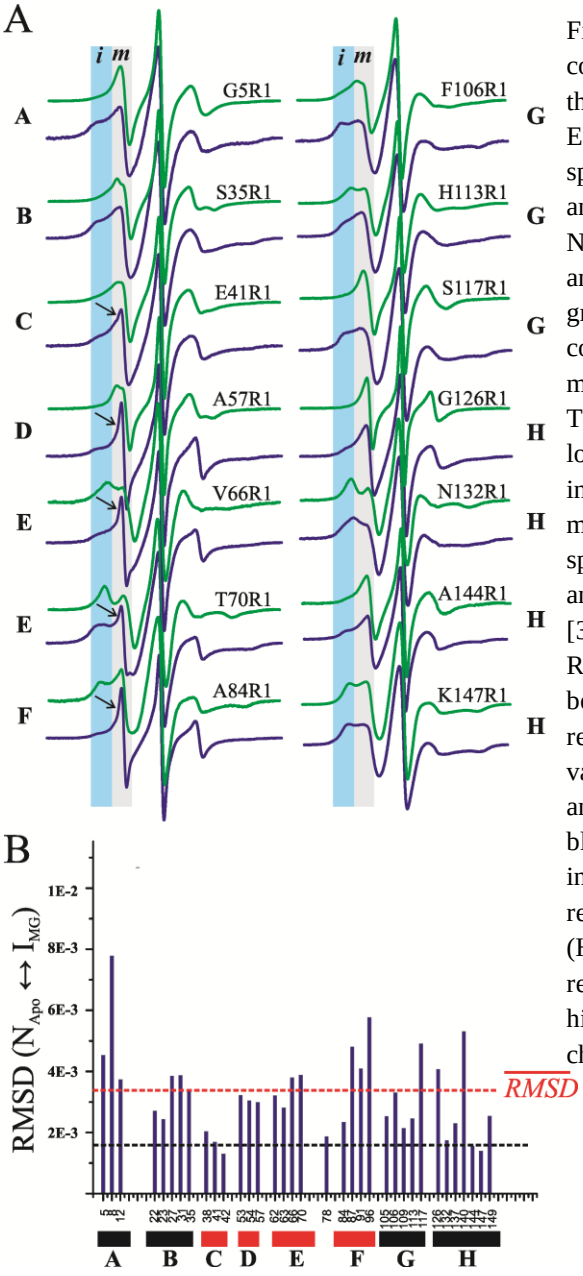
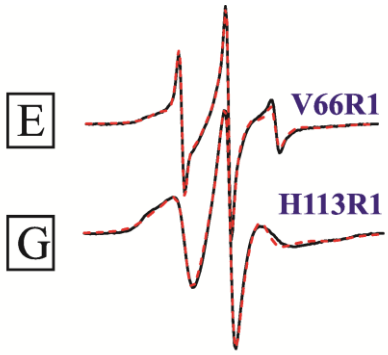


Figure 5.5: Mapping changes in the conformational flexibility of myoglobin in the $N_{Apo} \rightarrow I_{MG}$ transition. A) Representative EPR data of a set of sites reflecting the spectral changes observed between the N_{Apo} and the I_{MG} states. The EPR spectra in the N_{Apo} and the I_{MG} states are shown in green and blue, respectively. The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively. The region in which each R1 side chain is located is indicated. The arrows are use to indicate sharp component reflecting high ns mobility of the backbone (see text). The EPR spectra of residues 41R1 and 66R1 in N_{Apo} and I_{MG} states in Ficoll have been reported [39] and are reproduced here. B) Plot of the RMSD of the EPR spectra for each residue between the N_{Apo} and I_{MG} state. The dashed red line shows the overall average RMSD value. The average RMSD between the N_{Holo} and N_{Apo} states is shown for comparison (see black dashed line). The horizontal bars indicate the location of each site in the high-resolution structure of holo myolglobin (Helices A-H). The red bars identify the regions in which the EPR spectra show a highly mobile state of the nitroxide side chain.

As an example the more mobile component in the spectrum of residue 113R1 in helix G can be reasonably fit with the MOMD model using τ = 2.3 ns, S = 0.26 (Figure 5.6), which is similar to the values observed for R1 at helix surface sites. To elucidate whether there are temperature-dependent changes in the conformation and flexibility of the protein, spectra were also recorded at 323 K, the temperature at which NMR data has been reported for the I_{MG} state.



Mutant	τ (ns)	S	%
I_{MG} – 66R1	0.7	0	17
	1.9	0.28	21
	3.2	0.49	62
I_{MG} – 113R1	2.3	0.26	42
	5.9	0	58

Figure 5.7: A) Representative EPR spectra in the I_{MG} state in solution (blue) and in CNBr-activated Sepharose (gray). B) EPR spectrum of 19R1 in solution and on the solid support prior to acid unfolding (black) and after refolding (magenta).

The immobile component observed in the spectra of many sites in helices A, B, G and H in I_{MG} state 298 K are also observed in the spectra recorded at 323 K (*see* Appendix) suggesting that the higher temperature does not fundamentally alter the conclusions derived from the 298K data. On the other hand, the immobile component observed for R1 residues in the F helix at 298 K disappears at 323 K, suggesting changes in the conformational properties of this sequence at the higher temperature.

For a simple comparison of the changes in the N_{Apo} to I_{MG} transition, the RMSD for each site was computed as done for the $N_{Holo} \rightarrow N_{Apo}$ transition. As shown in Figure 5.5B, there are significant changes at most sites between the native apo and the MG state. In fact, the average root mean square difference value observed in the N_{Apo} to MG transition (red line on Figure 5.5B) is two times larger of that observed for the holo to apo transition (*see* dashed rectangle in Figure 5.5B) suggesting significant changes in the mobility of the nitroxide throughout the protein. Above average changes in the EPR spectra were observed at sites in helices A, F, G, and H.

Although soluble aggregates were removed in the purification of the spin labeled mutants, one might argue that the existence of relatively immobile states in the spectra of R1 at every site in the I_{MG} state is due the reappearance of aggregates at pH 4. To evaluate this

possibility, a series of spin-labeled mutants of myoglobin in the N_{Apo} state (100% monomer) were attached to CNBr modified Sepharose solid support and the pH of the solution was lowered from 6.1 to 4.1 in order to populate the molten globule state. The rationale behind this strategy is that protein immobilization on the solid support should, in principle, eliminate the possibility of aggregation. Save for residue A19R1, the EPR spectra of the immobilized proteins in the molten globule state are similar to those in solution (Figure 5.7A), suggesting that the immobile component is not from protein aggregation. Furthermore, as discussed above, the immobile component for some of the mutants disappears at high temperature and reappears at 298K, thus showing complete reversibility, which would not be expected for a sample containing soluble aggregates. Additional evidence that eliminates the possibility of protein aggregation as the origin for the more immobile component in the EPR spectra of myoglobin mutants in the I_{MG} state will be presented in below. The differences in the spectra of residue 19R1 on CNBr and in solution deserve comment. As shown in Figure 5.7B, the spectra of 19R1 in the N_{Apo} state before and after refolding from the I_{MG} state is reversible on the solid support, but not in solution, which supports the contention that the differences observed in the spectra of residue 19R1 on the solid support and solution in the I_{MG} state arise due to protein aggregation in the latter.

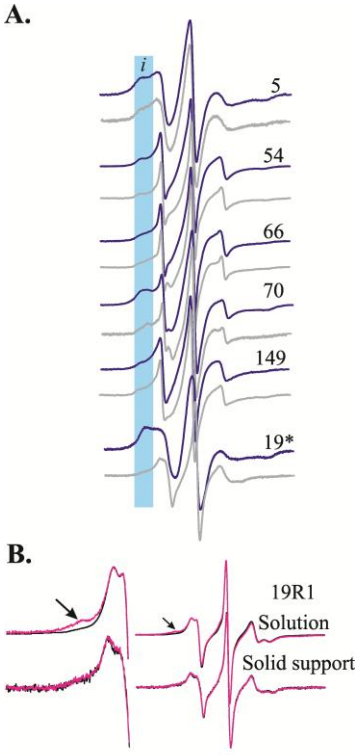


Figure 5.7: A) Representative EPR spectra in the I_{MG} state in solution (blue) and in CNBr-activated Sepharose (gray). B) EPR spectrum of 19R1 in solution and on the solid support prior to acid unfolding (black) and after refolding (magenta).

To examine whether the multicomponent spectra observed in the I_{MG} state arise from multiple conformations of the protein as opposed to R1 rotameric equilibria, the effect of

osmotic-perturbation was evaluated for every mutant. The results for a representative set of sites are shown in Figure 5.8A; the complete data set is shown in the Appendix section. With few exceptions, the addition of sucrose increased the relative population of the immobile component relative to the more mobile one, reflecting a shift in the equilibrium of the states that give rise to these components. The above results collectively suggest that the I_{MG} state at 298K is conformationally heterogeneous, with helices C, D, E, and F sampling highly dynamic unfolded conformations. These conclusions are summarized in Figure 5.8B, where the spheres representing R1 sites are color-coded according to the magnitude of the RMSD in the transition from N_{Apo} to I_{MG}, and the backbone is coded by color and shape to identify regions involved in conformational exchange or regions sampling unfolded states.

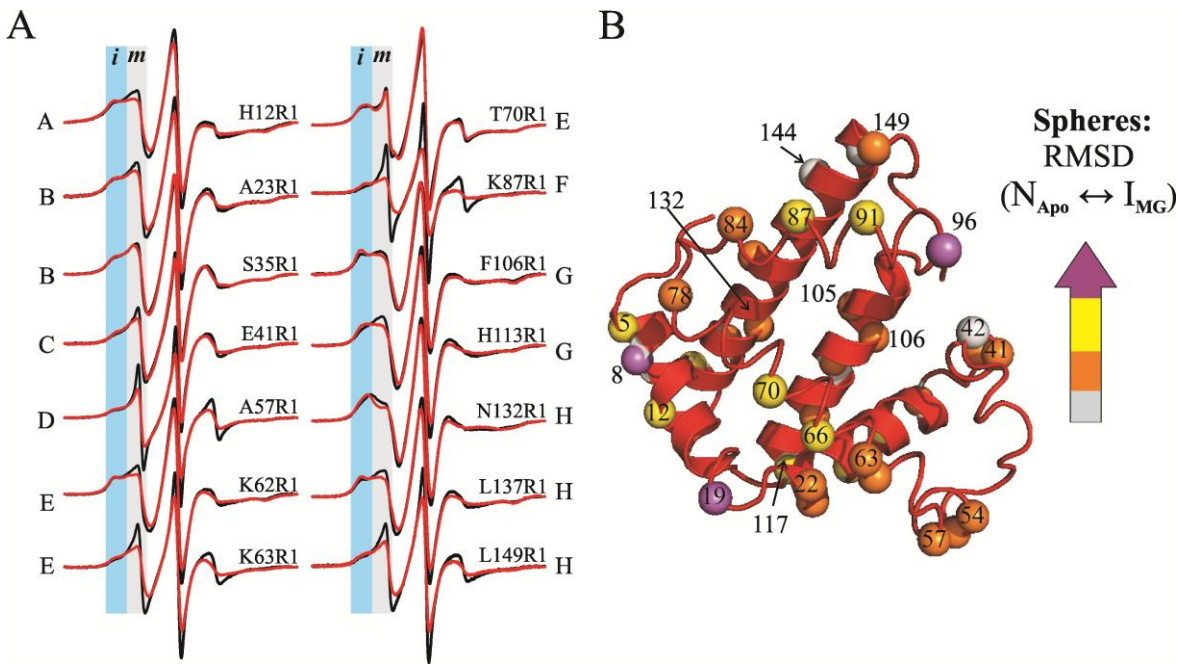


Figure 5.8: Osmotic-perturbation effects on the spectra of R1 of myoglobin in the I_{MG} state. A) Representative spectra recorded in Ficoll 70 (black) and sucrose (red) are superimposed. . The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively. B) Ribbon diagram of myoglobin showing the magnitude of RMSD between the N_{Apo} and I_{MG} states. The Cα spheres are color coded according to the magnitude of RMSD between the N_{Apo} and I_{MG} states (data from Figure 7B): **gray**, below $\overline{\boldsymbol{RMSD}}$ holo to apo (below dashed black line); **yellow**, $\overline{\boldsymbol{RMSD}} \leq$ RMSD $\leq \overline{\boldsymbol{RMSD}} + \sigma$; **magenta**, RMSD > $\overline{\boldsymbol{RMSD}} + \sigma$, where σ is the standard deviation from the mean.

5.3.4 *The acid unfolded state (U_{Acid}) of myoglobin*. The U_{Acid} state of myoglobin has been used as a paradigm for investigating the early events during protein folding [26-27,37,46]. Previous NMR studies have shown that apo myoglobin at pH 2.3 populates an expanded unfolded state that retains some residual helical structure in the regions corresponding to the helices A and H in the native state [22,26,46]. A non-native helix is also observed along residues 52-61 corresponding to the D-E helical region of the native protein. In addition, (PRE) studies using the R1 side chain at selected sites have identified transient long-range interactions between the N- and C-terminal regions of the protein, which may play an important role in initiating chain collapse during folding [27,37,47]. In the PRE studies, the EPR spectra of the R1-labeled mutants were not reported. To provide additional information on the local environment around the R1 side chain along the regions involved in long-range interactions, the EPR data for some of the spin-labeled mutants used in the PRE study were recorded including residue L11R1, which was used exclusively in this report to study the U_{Acid} state because it is a tertiary contact site in the native state. Additional mutants not included in the PRE study were also studied here. Unfortunately, Ficoll 70 is not stable at low pH (ca. < 3), thus sucrose was used to reduce the contributions of overall rotational diffusion of the protein to the EPR spectrum. Although sucrose is expected to shift conformational equilibria, the substates observed in the presence of sucrose should also be present in the absence of it, but at a different ratio. Interestingly, the EPR spectra of the majority of the sites investigated reflect multiple dynamic states of R1, with a relatively immobile component reflecting tertiary interaction of the nitroxide side chain with the environment (Figure 5.9A).

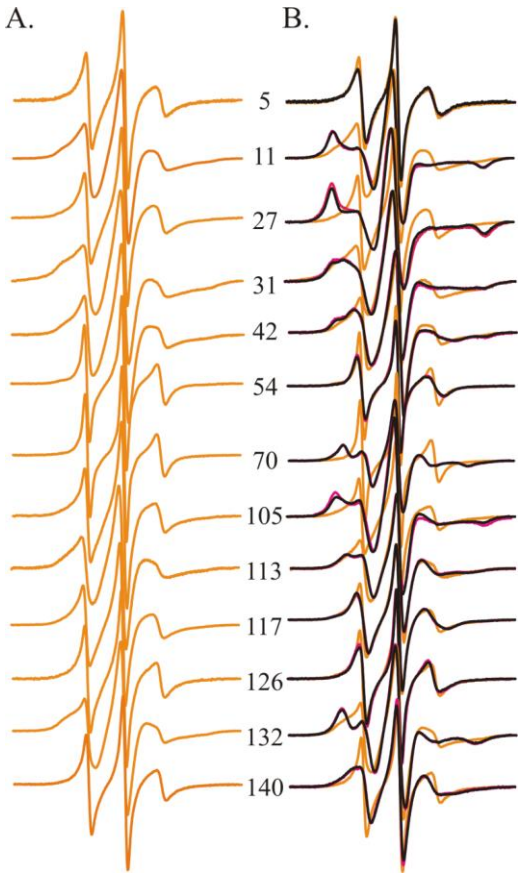


Figure 5.9: Characterization of the acid unfolded state (U_{Acid}) of myoglobin with SDSL. A) EPR spectra of the indicated myoglobin mutants in the U_{Acid} state. B) Refolding of each myoglobin mutant from the U_{Acid} state is completely reversible. The EPR spectra of each mutant in the N_{Apo} state recorded prior acid unfolding (black) and after refolding (magenta) from the U_{Acid} state are superimposed.

The presence of multiple spectral components likely arises from slow exchange between conformational substates, but this possibility could not be evaluated using the osmotic perturbation method due to the instability of Ficoll 70 at low pH as mentioned above. On the other hand, the EPR spectra of residues 54R1 and 70R1, corresponding to helices D and E in the native sequence show a single component reflecting high mobility of the nitroxide on the nanosecond time scale ($\tau \leq 1$ ns; Figure 5.10).

The EPR spectra of R1 in sites 31R1, 42R1, and 113R1, corresponding to helices B, C, and G, respectively in the native protein, reflect a restriction of the nitroxide motion on the nanosecond time scale. For example, the EPR spectrum of residue 31R1 located in helix B can be reasonably fit with a mobile component (*m*) and a relatively immobile component (*i*) of restricted mobility (Figure 5.10). Interestingly, these sequences were previously identified as being completely unfolded in NMR studies as judged by chemical shift values [22,26]. However, NMR relaxation [26] and PRE data [37] showed restricted backbone motions and long-range interaction of the aforementioned sequences due to local hydrophobic collapse in the U_{Acid} state. Thus, the restricted mobility for R1 in site 31, 42, and 113 in the U_{Acid} state likely arises from the transient tertiary contact within these locally unfolded segments.

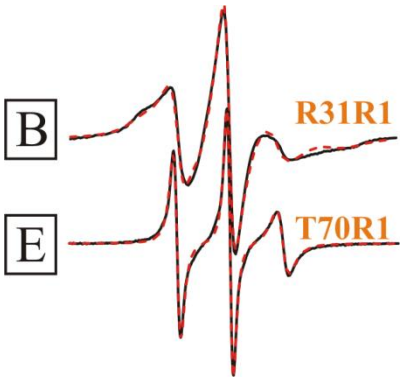


Figure 5.10: Experimental (black) and MOMD fit (dashed red line) of spectra of residue 31R1 and 70R1 in the U_{Acid} state. The spectrum of 31R1 was reasonably simulated with 78% *m* (τ_m = 1.4 ns, S_m = 0.21) and 22% *i* (τ_i = 5.3 ns). The spectrum of residue 70R1 was fit with a single isotropic motion with τ = 1 ns.

To examine whether the immobile component observed at many sites in the U_{Acid} state arises from to *soluble* aggregates, the unfolded protein was refolded back to the native state (N_{Apo}) by changing the buffer solution to 10mM sodium acetate pH 6.1, and the spectra were recorded and compared to the spectra obtained prior to populating the U_{Acid} state. As shown in Figure 5.9B, the EPR spectra of all sites in the N_{Apo} state before and after unfolding show complete reversibility, eliminating aggregation as the origin of the immobile component observed in both the I_{MG} (see above) and U_{Acid} states. The very small difference in the spectrum of residue 105R1 after refolding from the U_{Acid} state likely arises from alternative local re-packing due to the presence of the nitroxide side chain during refolding.

5.3.5 *Structural changes between N_{Holo}→ N_{Apo}→ I_{MG} transitions monitored via DEER spectroscopy*. The previous sections focus on the changes in flexibility of myoglobin as probed by EPR spectroscopy of single-labeled myoglobin mutants upon populating multiple folded states, however mapping structural changes with SDSL requires a fundamentally different strategy. As discussed in chapter 2 (section 2.8) of this dissertation, double electron-electron resonance (DEER) spectroscopy is a pulse EPR technique [48] that measures the probability distribution of distances between unpaired electrons in the 17 – 80 Å range [49] for frozen samples at cryogenic temperatures (< 80K). The most probable distance and width of the distribution provide direct information on the structure and structural heterogeneity, respectively. The structural heterogeneity is presumably related to the amplitude of molecular motion at physiological temperatures, which has contributions from both R1 motion and the intrinsic protein flexibility [50-51].

To monitor structural changes that take place in myoglobin in the $N_{Holo} \rightarrow N_{Apo} \rightarrow I_{MG}$ transitions, four pairs of R1 mutants were engineered and the interspin distance distribution was monitored for each of them in the N_{Holo}, N_{Apo}, and I_{MG} states using DEER spectroscopy. The X-ray structure of the holo protein (PDB ID: 2MBW) was used to model the R1 side chain based on the rotamer library determined from crystal structures, so predicted inter-nitroxide distances could be directly compared to the experimental values obtained for the well-ordered N_{Holo} state.

Figure 5.11A shows the location of the R1 pairs in myoglobin selected to monitor the distances between helices A and H (12R1/132R1), B and E (31R1/70R1), B and F (31R1/87R1) and D and H (57R1/132R1). Figure 5.11B gives the dipolar evolution functions (left panel) and derived distance distributions (right panel). The most probable distances for the four pairs in the

N_{Holo} state are in close agreement (± 1.5Å) with the predicted interspin distances from modeling the R1 side chain in the high resolution crystal structure (Table 5.1). Interestingly, the distance distribution for the 57R1/132R1 shows a second maximum with interspin-distance shifted ≈5 Å in comparison to the most probable distance, which may reflect a second position for the short and flexible D helix, but could also arise from multiple rotamers of R1 whereby the environment around the side chain remains the same for both rotamers (*see* Discussion). The relatively narrow distributions in N_{Holo} for 12R1/132R1, 31R1/70R1 are a reflection of the low structural heterogeneity in the frozen state and presumably a correspondingly low flexibility of the tertiary fold sampled by these R1 pairs at physiological temperature. On the other hand, the distribution of distance for 31R1/87R1 is broader and asymmetric compared to the other pairs. A greater structural heterogeneity here is consistent with the conformational exchange involving helix F identified by the osmolyte perturbation of the multicomponent spectrum for 87R1.

The distribution of distances sampled by 12/32, 31/70 and 57/132 in the N_{Apo} state are very similar to those in the N_{Holo} form (Figure 5.11; Table 5.1), indicating little change in the overall tertiary fold sampled by these pairs. Although there is no change in the most probable distance for the 31R1/70R1 pair, the width of the distance distribution observed in the N_{Apo} state is greater by ≈ 5 Å (Table 5.1), likely reflecting higher backbone flexibility in the N_{Apo} state. The interspin distance distribution for the 31R1/87R1 pair showed dramatic changes in the $N_{Holo} \rightarrow N_{Apo}$ transition, specifically the distance distribution between the B and F helices becomes bimodal with a long distance similar to that observed in the holo form (36 Å) and a new interspin distance with maximum at 20 Å. Because there is no change in the most probable distance and only a slight change in distribution width for the 31R1/70R1 pair, the origin of the bimodal distribution for the 31R1/87R1 pair must reside in movement of 87R1 in the F helix. A model to

account for the striking change in distance distribution in the transition from N_{Holo} to N_{Apo} will be considered in the Discussion.

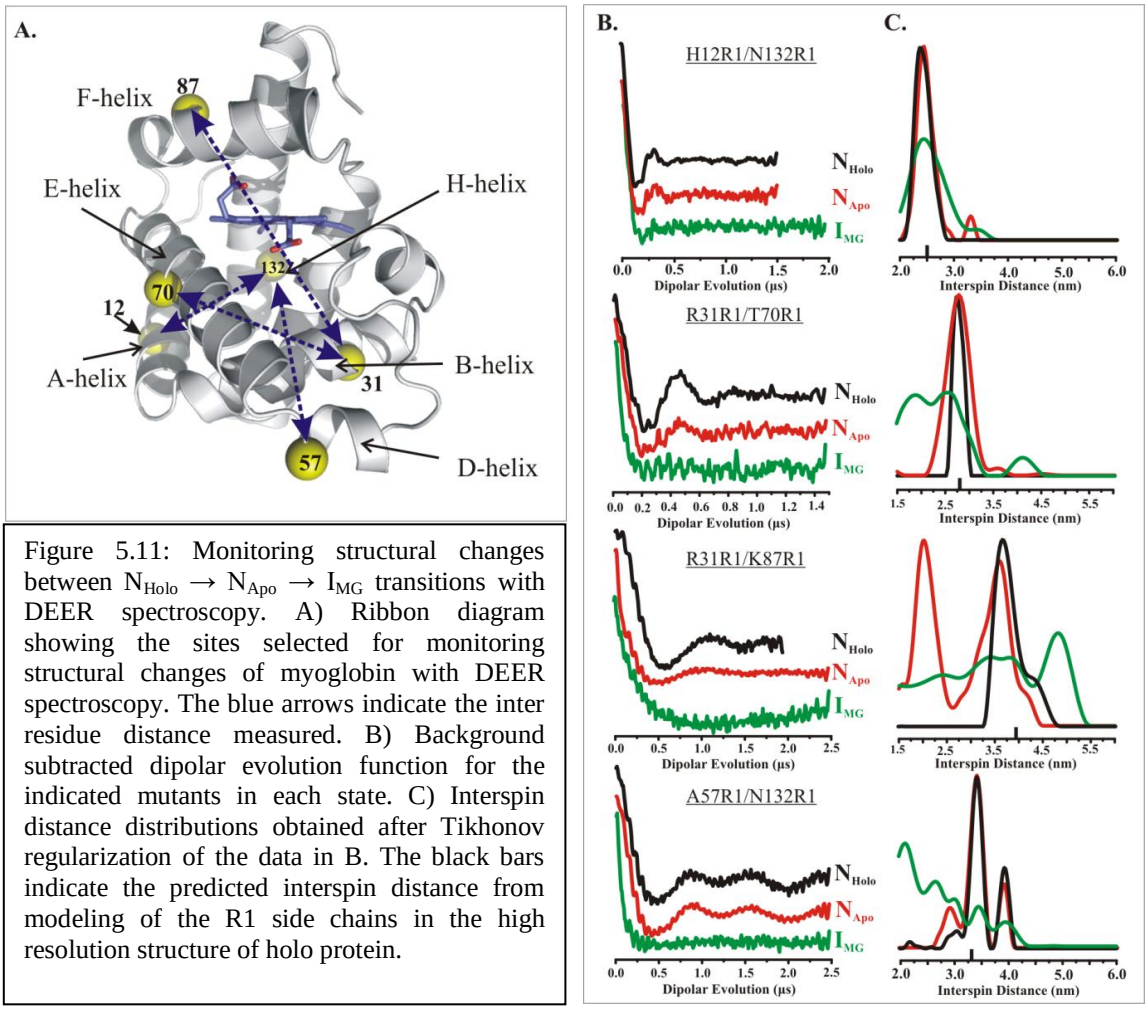


Figure 5.11: Monitoring structural changes between $N_{Holo} \rightarrow N_{Apo} \rightarrow I_{MG}$ transitions with DEER spectroscopy. A) Ribbon diagram showing the sites selected for monitoring structural changes of myoglobin with DEER spectroscopy. The blue arrows indicate the inter residue distance measured. B) Background subtracted dipolar evolution function for the indicated mutants in each state. C) Interspin distance distributions obtained after Tikhonov regularization of the data in B. The black bars indicate the predicted interspin distance from modeling of the R1 side chains in the high resolution structure of holo protein.

Upon formation of the I_{MG} all R1/R1 distance distributions are dramatically broader than in the N_{Holo} and N_{Apo} state, reflecting a structural heterogeneity consistent with the conformational exchange at physiological temperatures revealed by the osmotic perturbation response presented in Figure 5.8.

Table 5.1 Most probable interspin distance and distance distributions for R1 doubles of myoglobin in the N_{Holo}, N_{Apo}, and I_{MG} states.

Mutant	Helices	R1-R1 distance from modeling[†] (Å)	Most probable R1-R1 distance from DEER (Å)	Width of distance distribution from DEER[‡] (Å)
12R1/132R1 N_{Holo}	A/H	25	23.7	5.8
12R1/132R1 N_{Apo}		-	24.3	5.8
12R1/132R1 I_{MG}		-	24.3	~14.0
31R1/70R1 N_{Holo}	B/E	28	27.3	4.3
31R1/70R1 N_{Apo}		-	27.6	9.2
31R1/87R1 N_{Holo}	B/F	39	36.7	11.0
31R1/87R1 N_{Apo}		-	$20.2^{53\%}$ & $36.0^{47\%}$	10.0 & 14.5
57R1/132R1 N_{Holo}	D/H	33	34.1	3.9
57R1/132R1 N_{Apo}		-	34.1	4.4

[†] Estimates of inter-nitroxide distances from modeling the R1 side chain in the crystal structure of the holo protein (PDB ID: 2MBW)
[‡] Width of the distribution taken as R1-R1 distances between 16 to 84% probability

5.4 Discussion

An initial goal of the study presented here was to evaluate how molecular flexibility of partially folded states is revealed as viewed through the CW EPR time window. To this end the changes in backbone flexibility and conformational properties of myoglobin were monitored *via* lineshape analysis and the osmolyte perturbation criterion introduced in chapter 3 of this dissertation. As defined in a site-specific fashion by NMR [22], there are four states of myoglobin, which can be populated at equilibrium conditions that exhibit conformational flexibility increasing in the order $N_{holo} \rightarrow N_{apo} \rightarrow I_{MG} \rightarrow U_{acid}$. Thus, in the present study the conformational flexibility of myoglobin was systematically "tuned" and monitored by SDSL. The results presented in this chapter combined with the results discussed in Chapter 4 of this dissertation establish an important principle, namely that *the onset of conformational flexibility of folded and partially folded states on a time scale >100 ns may be revealed by the appearance of components in the CW EPR spectrum corresponding to immobilized states of R1.* The fact that conformational flexibility shows up as immobilization of R1 is simply due to the fast time scale of the EPR experiment. The lifetime of the individual conformational substates is long on the EPR time scale, so R1 will sense each substate frozen in time. In some of the substates, the local conformation can be such that R1 experiences immobilizing contacts with other parts of the structure as depicted in Figure 5.12A. On the other hand, the presence of *a conformational substate that involves a locally unfolded sequence exhibiting high backbone flexibility will give rise to a sharp spectral component in the CW EPR spectrum corresponding to a highly mobile state of R1* (*see* Figure 5.12C).

Details of the results leading to the above conclusions are discussed below with respect to the EPR spectral changes observed in the $N_{Holo} \rightarrow N_{Apo} \rightarrow I_{MG} \rightarrow U_{Acid}$ transitions and comparison with relevant NMR data.

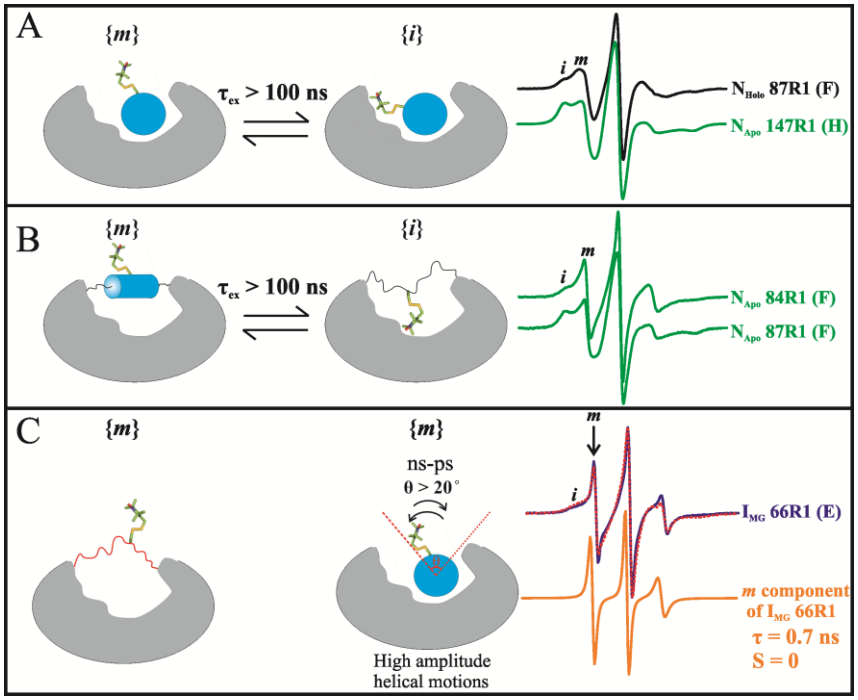


Figure 5.12: Cartoon representation of the proposed substates and exchange mechanism that give rise to the EPR spectra observed in the N_{Holo}, N_{Apo}, and I_{MG} states of myoglobin. For each scenario, examples of sites undergoing the proposed exchange mechanism are shown in the right panel (see text). A) Helix rocking motion occurring in the μs-ms time scale brings the nitroxide into tertiary contact with the protein environment giving rise to a relatively immobile state of R1. B) Exchange between an unfolded, but "collapse" state (*i*) and a helical state of exhibiting high backbone mobility. C) Highly mobile state of R1 arising either from unfolded polypeptide chain with high backbone flexibility or a flexible helix undergoing high angular amplitude of backbone fluctuation ($\theta > 20^{\circ}$) in the ns-ps time scale. Spectral simulation (red dashed lines) and the deconvoluted mobile component obtained from the fit (orange) of the spectrum of residue 66R1 in the I_{MG} state are shown.

5.4.1 N_{Apo} State. The backbone flexibility and conformational properties of myoglobin in the N_{Holo} state have been discussed in detail in the previous chapter. The most common change in the EPR spectra between N_{Holo} and N_{Apo} is the appearance of a new spectral component for R1 in the latter state. Specifically, the spectra of residues in the sequence 38-42 (helix C), residues 84-109 (helix F, FG loop, N terminal portion of helix G), and at residues 137-149 (C terminal

portion of helix H) show a new relatively immobilized state of R1 arising from new tertiary interactions of R1 with the protein environment (Figure 5.3; Appendix). In each case, an osmotic shift confirms conformational exchange. Of particular interest are the changes in the flexibility of the backbone as reported by R1 for the F-helix in the N_{Holo} to N_{Apo} transition. Although the NMR resonance lines for the F helix sequence in the N_{Apo} state are broadened beyond detection, the states contributing to the exchange are resolved, at least partially, in the EPR spectra. In contrast to the holo state, the spectra of all the sites on the helix F sequence reflect a dominant immobilized state in addition to a second population (~15-25 %) exhibiting narrow resonance lines reflecting a flexible backbone, but not necessarily arising from a free-flight random coil state. The highly mobile state could, in principle, arise from a helical segment fluctuating with relatively high angular amplitude and correlation time in the order of 1 ns. As an example, reasonably good fit was obtained for the spectrum of residue 87R1 using a correlation time and order parameter for the more mobile component of 1.3 ns and 0.08, respectively (refer to Chapter 3), the latter of which corresponds to an angular amplitude of backbone motions of about 15°.

Insight into the structural origins of the two-component spectra observed for sites in the F helix in the apo state was obtained by analysis of interspin distance distribution between 31R1/87R1 pair, which reflects a bimodal distance distribution between residue 87R1 in helix F and a reference site at 31R1 in helix B (Figure 5.11). The interspin distance of 40 Å is similar to that observed in the N_{Holo} state, while the shorter distance near 20Å can only be satisfied by a second population whereby the 87R1 side chain is buried in the now empty heme cavity. It is difficult to imagine the entire F helix rotating to position 87R1 (and the other sites in F) in the cavity, so it is more likely that the state(s) giving rise to the immobilized state of the nitroxide and the short distance measured arise from a disordered sequence partially filling the empty

cavity. In accord to the above interpretation of the EPR data, radical footprinting studies [52] and recent relaxation dispersion NMR data [53] both suggest that the heme binding site is partially occupied in the apo state, likely by the flexible F-helix sequence.

As anticipated, the sequences reflecting new spectra components in the N_{Apo} state overlap to a large extent with those identified by NMR to be in conformational exchange. For example, resonances for $^{13}C\alpha$, ^{1}HN and ^{15}N were not observed in residues 81-102 and 144-148, a result attributed to conformational exchange on the ms time scale [21-22]. Although the sequences identified to be in exchange by SDSL-EPR coincide with those exhibiting broad NMR resonances lines, the EPR data reveal exchange beyond those sequences to include residues 38-42, 103-109 and 137-144. Recent NMR relaxation dispersion measurements of myoglobin in the apo state indeed revealed exchange processes for residues in the sequences 103-108 and 136-152 with exchange lifetimes less than about 500 ms [53]. Moreover, similar exchange rates were found for residues in the sequence 38-42 corresponding to helix C. Thus, to the extent that they can be compared, there is essentially complete agreement between the NMR chemical shift and relaxation dispersion data with the SDSL-EPR data.

As pointed out from the NMR relaxation dispersion studies [53], the sequences involved in conformational exchange are in helical segments that are in the vicinity of helix F (C, E, N-terminus of G and C terminus of H) and it was suggested that the relaxation dispersion arises from ms fluctuations in structure and/or packing of the F helix region. Although it seems very plausible that the conformational fluctuations in these other helices are indeed coupled to fluctuations in F against which they are packed (*e.g.*,137R1), the conformational exchange observed in these segments extends to sites beyond those in direct contact with helix F. For

example, osmolyte shifts are detected in two-component spectra for R1 residues along an extended solvent-exposed surface of helix H directly opposite that in contact with F (140R1 to 149R1); the most distal of these residues is a full turn from the nearest contact with helix F. Similar arguments can be made for helix G, where R1 residues along the solvent exposed surface that extend 2 turns away from F sense local packing changes (106-109). These results argue for fluctuations in the N_{apo} structure delocalized from sites in direct contact with helix F. The fluctuations sensed by R1 in helices in these cases may involve states such as those shown in Figure 5.13A. Further evidence for such fluctuations may be found in the reduced values of the $^{13}C\alpha$ shifts for 102-109, 137-144 and 38-42 in N_{Apo} relative to N_{Holo} [22], suggesting the presence of states with distorted backbone geometry in exchange on the ms or faster time scale with the native helical state.

Collectively, the SDSL-EPR data are in remarkably good agreement with the NMR data and most importantly, the EPR data provides direct information on both the backbone flexibility and locations of the F helix in N_{Apo} state.

5.4.2 *I_{MG} state*. The transition from N_{Apo} to I_{MG} is remarkable in that the spectra for R1 at every site analyzed reflect two resolved components, many of which have a single-component in the N_{Apo} state. Osmotic shifts suggest that for most sites the multicomponent spectra arise from conformational exchange between, at least, two states. The mobility of the nitroxide reflected in each of the resolved spectral components is of importance for obtaining insights into the backbone flexibility and structure of each of the conformational states in equilibrium. For example, the more mobile component of the spectra of residues in the A, B, G, and H sequence reflects constrained motions of the backbone (*i.e.*, $\tau > 1$ ns and nonzero order parameters), and

thus it is likely that the substate giving rise to the mobile component involves a folded structure. As an example, the more mobile component of residue 113R1, located in helix G, can be reasonably fit with τ = 2.3 and S = 0.26, which are comparable to those values observed for R1 in folded helical structures [54-55] including those in the N_{Holo} state (*see* Chapter 4). However, it is realized that there may be collapsed unfolded substates with constrained backbone motions, which would give rise to a mobile component similar to those observed for sites in the A-B-G-H sequence. Specific examples of this are discussed below.

In contrast to residues in the A-B-G-H sequence, the EPR data depict a very different situation for residues within helices C, D, E and F. On each of these helices, most of the spectra of R1 have a resolved component that reflects a fast, essentially isotropic motion likely arising from a flexible random coil state. For example, the more mobile component on the spectrum of 66R1 located on helix E, can be reasonably fit with isotropic motion (*i.e.*, S = 0) and τ = 0.9 ns (Figure 5.6), which is comparable to τ values observed in unfolded protein sequences [44-45].

Differences on the conformational flexibility of core (A-B-G-H) and non-core (C-D-E-F) helices in the I_{MG} state can also be inferred from the DEER data. Although the DEER data set is limited, qualitative differences were observed in the distance distribution between pairs involving sites within helices A and H, and those involving sites located in helices D, E, and F. For example, the width of the distance distribution for the 12R1/132R1 pair, located in helix A and H, respectively, is narrower than the other three pairs, which include sites 57R1, 70R1, and 87R1, located in helices D, E, and F, respectively. In addition to the narrower distance distribution, the 12R1/132R1 pair have a monomodal distribution, while the distribution on the

other three pairs reflect multiple peaks as depicted in Figure 5.11, presumably corresponding to different positions of the backbone.

Thus, the CW lineshape and DEER data combined suggest that the myoglobin structure becomes conformationally heterogeneous on the EPR times scale with exchange lifetimes longer than about 100 ns. Most importantly, the data revealed differences in the extent of backbone motions between helices A-B-G-H and C-D-E-F. These findings are entirely consistent with the NMR data on the I_{MG} state [22,25]. In particular, backbone chemical shifts, low temperature coefficients for H^N chemical shifts, and relatively high heteronuclear NOEs all point to a high content of helical structure with a low amplitude of fast backbone dynamics in A, the C terminal end of B, G and H. Analysis of NMR relaxation parameters revealed restricted ns mobility of the backbone in the A-B-G-H sequence, but with higher amplitudes of motion on a slower ($\leq$ ns) time scale, perhaps reflecting rigid body rotational diffusion of a compact unit, and possibly conformational exchange [22,25]. In regard to the C, D, E, and F sequence, the data from $^{13}C\alpha$ backbone chemical shifts, temperature coefficients for H^N chemical shifts, and reduced heteronuclear NOEs revealed that the C-D-E sequence has lower helical content and higher amplitude of fast backbone dynamics compared to the A-B-G-H core helices [22,25].

As mentioned above, due to poor spectral resolution, the aforementioned NMR studies on the I_{MG} state were performed at high temperatures (323 K). Because the conformation and dynamics of the protein can be influenced by temperature, the SDSL-EPR data was also acquired for selected sites at 323 K (see Appendix). It is realized that such an increase in temperature excites internal motions of the R1 side chain, sharpening the EPR spectral lines and effectively reducing resolution. Hence without further studies, it would be difficult to differentiate changes

in R1 internal motion from those due to excitation of protein motions. Nevertheless, a qualitatively similar pattern of two-component spectra is observed for R1 in the regions corresponding to helices A, B, G and H in the native state at both 298K and 323K. On the other hand, a single highly mobile component is observed at sites in the F-helix at the higher temperature (see Appendix) consistent with the high degree of backbone flexibility observed by NMR [25]. Since sites in the F-helix reflect multiple components at 298 K (Figures 8, S12, and S14), the EPR data suggests temperature-dependent differences in the conformational properties of the F-helix sequence.

Although the helices A, B, G and H maintain much of the secondary structure of N_{Holo} and N_{Apo} states, they are nevertheless involved in conformational exchange, as suggested by both the NMR and EPR data. One model consistent with the data is a conformational ensemble of these helices in the I_{MG} state involving substates such as those shown in Figure 5.13A. The remaining sequences likely involve a similar ensemble of states, with helices fluctuating between positions, but also include a dynamically disordered state, such as that shown in Figure 5.13B.

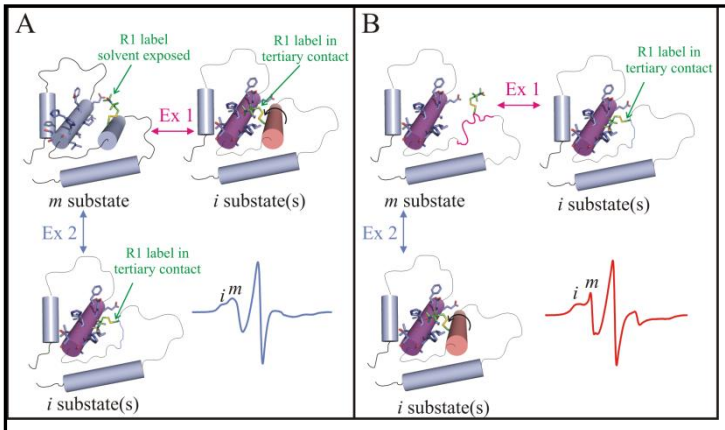


Figure 5.13: Proposed conformational exchange mechanism for apomyoglobin consistent with the EPR and NMR data in the I_{MG} state. A) Conformational substates in the A-B-G-H sequence. In the first model (Ex 1), the conformational exchange involves a helix fluctuating between different positions in the μs-ms time scale whereby the nitroxide senses multiple environments. In the second model (Ex 2), the backbone fluctuates between helical and an unfolded collapsed conformation with restricted mobility in the ps-ns time scale. For either exchange mechanism, the immobile component may arise from a heterogeneous population (*i.e.*, multiple states with similar environment around the R1 side chain). B) Proposed conformational exchange in the C-D-E-F sequence. The first model (Ex 1) involves exchange between a flexible disordered state with unfolded collapsed state(s) in which the nitroxide is in tertiary contact with the protein. The alternative scenario (Ex 2) involves exchange between a helix fluctuating between multiple positions with an unfolded state of the backbone with high flexibility in the ps-ns time scale.

5.4.3 U_{Acid}. Although the EPR data set on the acid unfolded state of myoglobin is limited, the most important finding is that the EPR spectra of many of the sites investigated reflect restricted backbone flexibility on the ns time scale and, possibly, μs-ms fluctuations. For example, two-spectral component are clearly resolved in, at least, 6 of the sites investigated (*e.g.*, 11, 31, 42, 105, 113, and 132). Among these sites, the more mobile component of residues 11R1, 31R1, 42R1, and 113R1 reflect constrained motion ($\tau > 1$ ns) of the backbone as judged by the width of the resonance line corresponding to the more mobile state and by spectral simulation. For example, the spectrum of residue 113R1 located in the G helix region of the native protein can be reasonably fit with τ and S for the mobile component of 2.3 ns and 0.26, respectively (Figure 5.10). The restricted mobility of R1 for the mobile component in the aforementioned sites could be either due to the presence of helical structures with constrained mobility or from local hydrophobic collapse of the unfolded sequence. On the other hand, the spectra of residues 54R1 and 70R1 corresponding to D and E region reflect the narrowest line width and hence the highest mobility of the backbone. For example, the spectrum of residue 70R1 can be reasonably fit with a single component reflecting isotropic motion and $\tau \approx 1$ ns (Figure 5.10). Thus, of the regions investigated, the D and E sequence appear to exhibit higher ns backbone flexibility with no evidence of μs–ms fluctuations.

Of particular interest are the differences in the mobility of R1 for sites in the same region. For example, the EPR spectrum of residue 5R1, located in the A helix, reflects a single component with fast motion, while the spectrum of residue 11R1 have 2 components of restricted mobility. The same is observed for 27R1 and 31R1. These results suggest highly localized backbone motions and site-specific tertiary-interaction.

These findings are entirely consistent with the NMR and PRE data [22,26,37]. In particular, backbone chemical shifts and NMR relaxation parameters revealed the presence of transient helical structures with restricted ns motions for the sequence corresponding to helices A and H in the native protein, while a non-native D/E helix exhibiting high backbone flexibility was observed [26]. Although the chemical shift data showed no evidence of helical structure in the B and G helix sequences, the relaxation studies suggested restricted ns motions of the backbone in these sequences, which was interpreted as arising from local hydrophobic collapse [22,26]. This latter interpretation of the NMR data was confirmed in subsequent PRE studies [37]. Thus, the sequences that exhibit constrained motions of the backbone in the ns time scale as judged by the EPR spectra of R1 (A, B, C, G, and H) coincide to a large extent with those identified in NMR studies as exhibiting helical propensities (A and H) and/or involved in long-range hydrophobic interaction (A, B, C, G, and H).

NMR relaxation studies showed evidence of μs-ms motions in the sequence corresponding to residues 8-19 and 112-116. The EPR spectra of residues 11R1 and 113R1 clearly showed two-well resolved components that likely arise from the μs-ms motions identified by NMR relaxation. The origins of the multicomponent EPR spectra for the other sites could not be evaluated due to the incompatibility of Ficoll 70 (viscogen) with acidic conditions. Future experiments involving osmotic perturbation on myoglobin immobilized on a solid support (*e.g.*, CNBr Sepharose) will be done to elucidate the origins of the multicomponent spectra for such sites in the U_{Acid} state.

5.5 Appendix

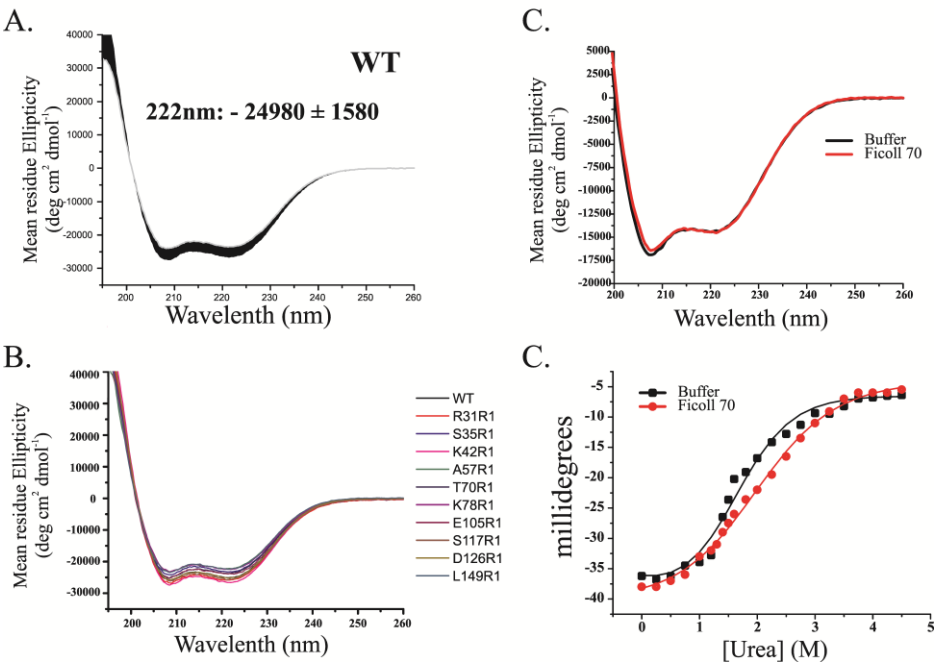


Figure 5.A.1: Effect of R1 and Ficoll 70 on the structure and stability of apomyoglobin. A) Far-UV CD spectrum of WT myoglobin showing the standard deviation of the data based on triplicate measurements. B) Far-UV CD spectra of R1 mutants of myoglobin in the $\mathbf{N_{Apo}}$ state. B) Ficoll 70 effects on the secondary structure of myoglobin in the $\mathbf{I_{MG}}$ state. C) Ficoll 70 effects on the stability of myoglobin in the $\mathbf{I_{MG}}$ state. The chemical denaturation data was fit using a nonlinear extrapolation method as described in Chapter 8. The relative stabilities obtained from the fits are shown in Table 5.A.1.

Table 5.A.1: Effect of Ficoll 70 on the stability of ApoMyb WT in the $\mathbf{I_{MG}}$ state (10mM sodium acetate pH 4)

Sample	pH	ΔG (H_2O) (kcal/mol)	*m* (kcal/mol*M)	$[Urea]_{1/2}$
Buffer	4	-1.87±0.03	1.22±0.07	1.54±0.03
Ficoll 70	4	-1.72±0.01	0.89±0.03	1.94±0.02

Change in stability of ApoMyb in Ficoll 70 solution: ΔΔG = -1.87 – (-1.72) = **0.15 kcal/mol**

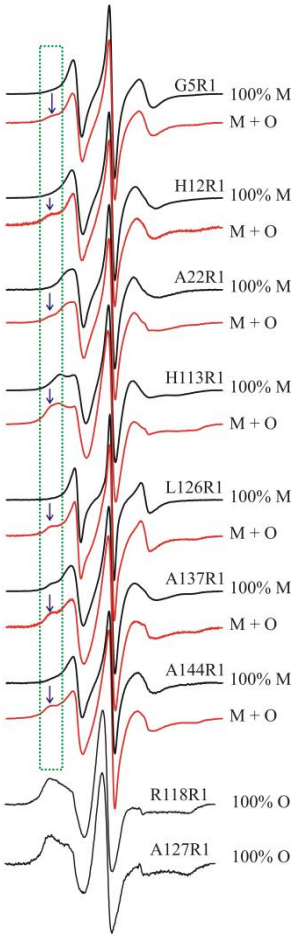


Figure 5.A.2: Contribution of oligomeric species to the EPR spectra of spin-labeled myoglobin mutants. EPR spectra of R1 mutants before and after gel filtration are shown in red and black, respectively. The arrows indicate spectral components arising due to the presence of misfolded proteins. M: monomer, O: oligomer.

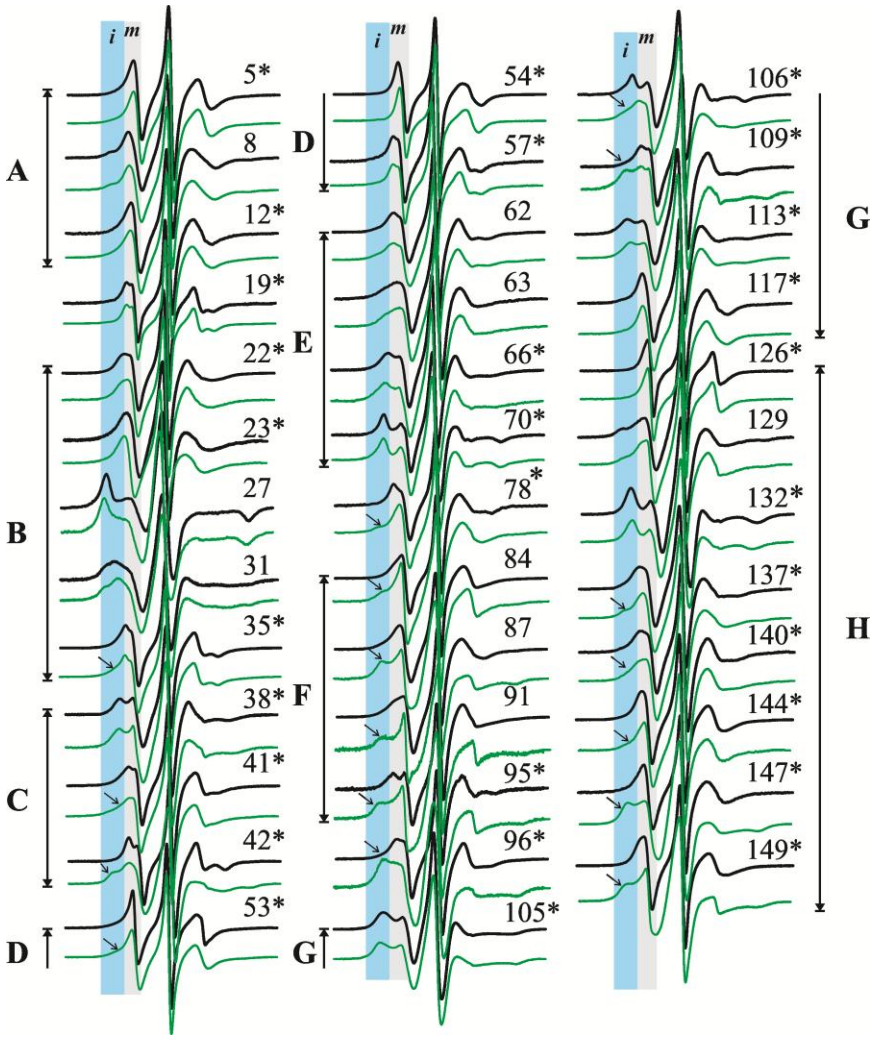


Figure 5.A.3: EPR spectra of R1 at the indicated sites in the N_{Holo} (black) and N_{Apo} (green) state. The asterisks identify the sites reflecting a single dynamic state in the N_{Holo} state.

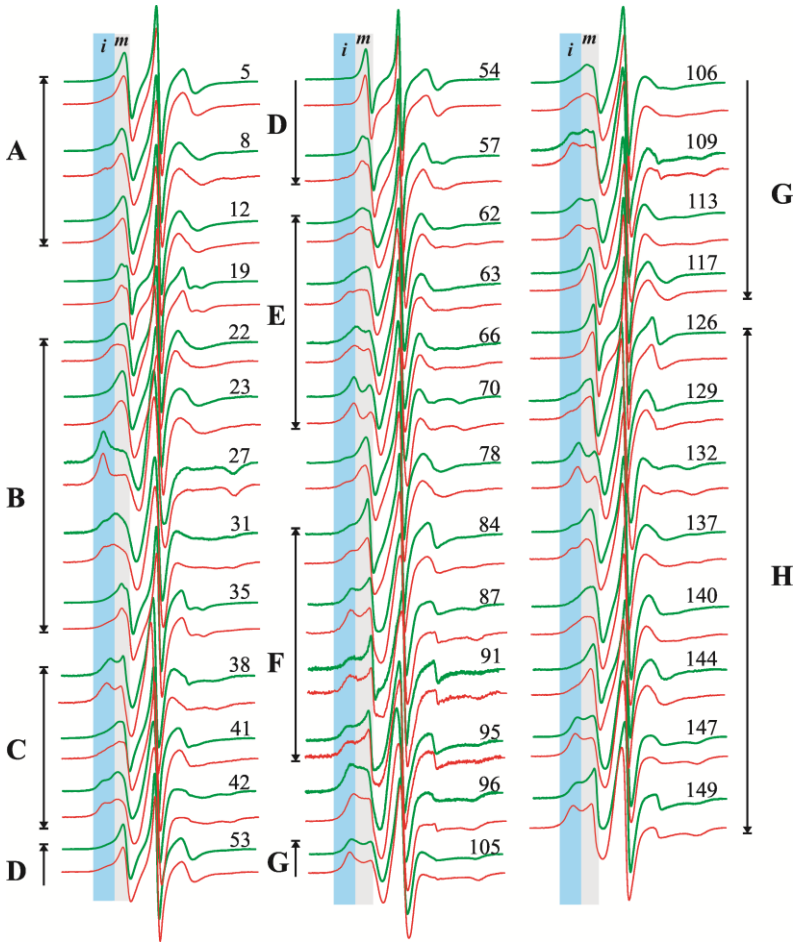


Figure 5.A.4: Mapping the conformational flexibility of myoglobin in the N_{Apo} state with osmotic-perturbation SDSL. EPR spectra of R1 mutants of myoglobin in the N_{Apo} state recorded in Ficoll 70 (green) and in sucrose (red). The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively.

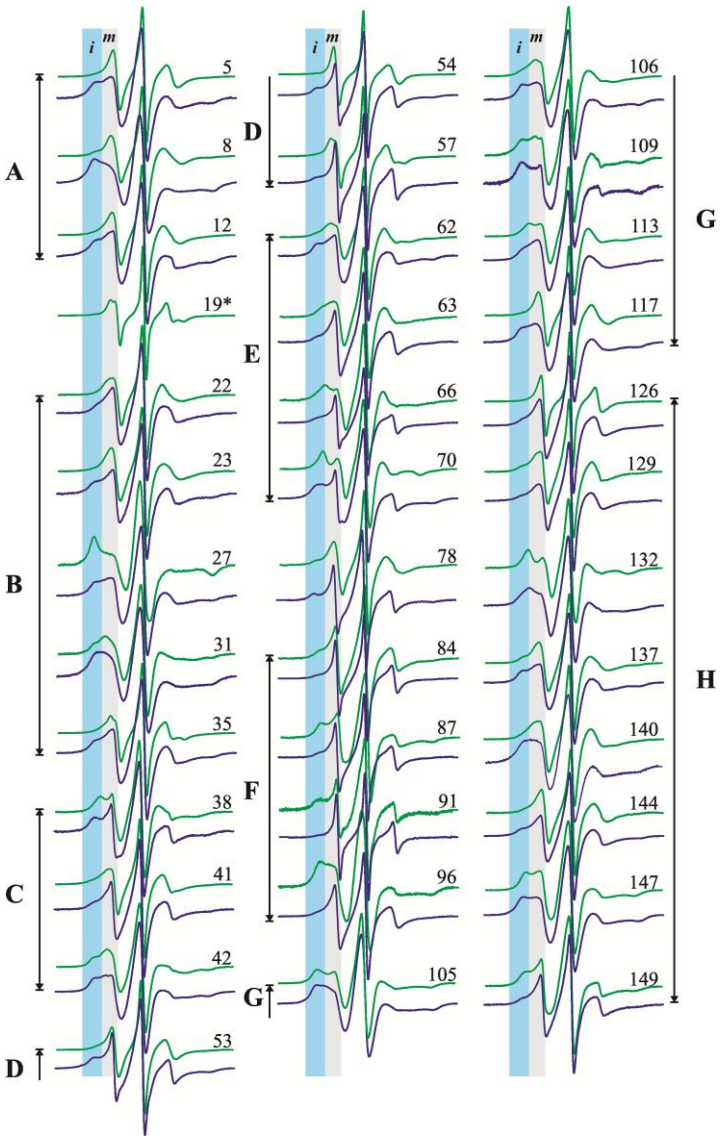


Figure 5.A.5: EPR spectra of R1 at the indicated sites in the N_{Apo} (green) and I_{MG} (blue) state. The EPR spectrum of residue A19R1 in the I_{MG} state (*) is not shown due to the presence of aggregates in the sample.

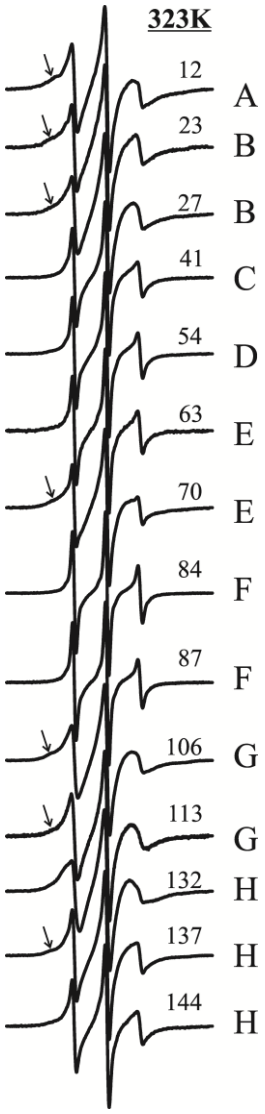


Figure 5.A.6: EPR spectra of the indicated mutants recorded at 323K. The relatively immobile spectral components observed for some of the mutants are highlighted.

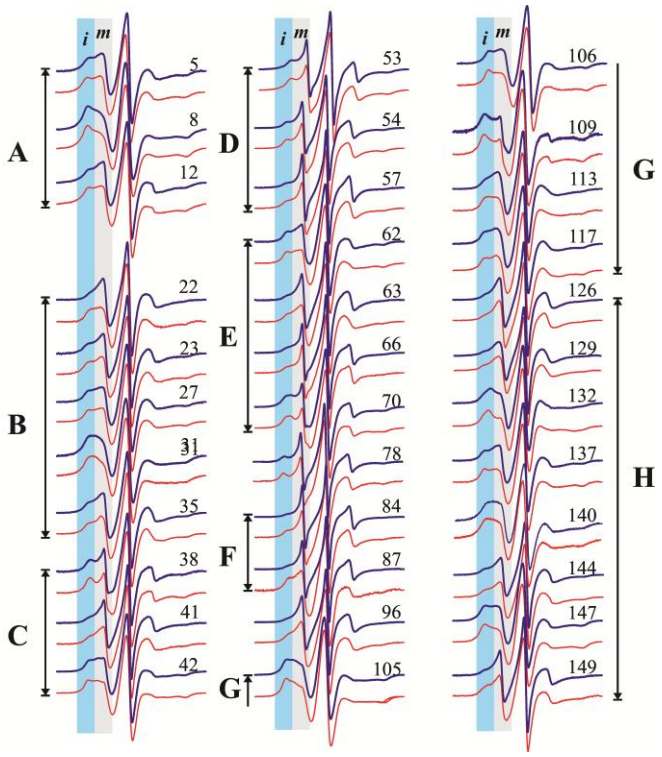


Figure 5.A.7: Mapping the conformational flexibility of myoglobin in the I_{MG} state with osmotic-perturbation SDSL. EPR spectra of R1 mutants of myoglobin in the I_{MG} state recorded in Ficoll 70 (blue) and in sucrose (red). The light blue and gray areas identify spectral intensities corresponding to relatively immobile and mobile states of the nitroxide, respectively.

5.6 References

1. Uversky, V.N., Gillespie, R.J., and Fink, L.A. 2000. Why are "natively unfolded" proteins unstructured under physiological conditions? *Proteins J.* 41:415–427.
2. Dunker, A.K., Obradovic, Z., Romero, P., Garner, E.C., and Brown, C.J. 2000. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* 11:161-171.

3. Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol*. 337:635-645.
4. Dunker, A.K., Cortese, M.S., Romero, P., Iakoucheva, L.M., and Uversky, V.N. 2005. Flexible nets: The roles of intrinsic disorder in protein interactions. *FEBS* 272:5129-5148.
5. Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., and Obradovic, Z. 2002. Intrinsic disorder and protein function. *Biochemistry* 41:6573-8582.
6. Dyson, J.H. and Wright, P.E. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol*. 6:197-208.
7. Pervushin, K., Vamvaca, K., Vogeli, B., and Hilvert, D. 2007. Structure and dynamics of a molten globule enzyme. *Nat. Struc. Mol. Biol*. 14:1202-1206.
8. Fink, A.L. 2005. Natively unfolded proteins. *Curr. Opin. Struc. Biol*. 15:35-41
9. Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. 2000. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci.* 97:8868-8873.
10. Gunasekaran, K., Tsai, C.J., Kumar, S., Zanuy, D., and Nussinov, R. 2003. Extended disordered proteins: targeting function with less scaffold. *Trends Biochem. Sci.* 28:81-85.
11. Keiderling, T.A. and Xu, Q. 2002. Unfolded peptides and proteins studies by infrared absorption and vibrational circular dichroism spectra. *Adv. Prot. Chem*. 62-111-161.
12. Fieden, C., Chattopadhyay, K., and Elson, E.L. 2002. What fluorescence correlation spectroscopy can tell us about unfolded proteins. *Adv. Prot. Chem*. 62:91-109.

13. Dyson, H.J. and Wright. P.E. 2004. Unfolded proteins and protein folding studied by NMR. *Chem. Rev*. 104:3607-3622.
14. Eliezer, D., Jennings, P.A., Dyson, H.J., and Wright, P.E. 1997. Populating the equilibrium molten globule state of apomyoglobin under conditions suitable for characterization by NMR. *FEBS Lett*. 417:92-96.
15. Columbus, L. and Hubbell, W.L. 2002. A new spin on protein dynamics. *Trends Biochem. Sci*. 27:288-295.
16. Morin, B., Bourhis, J.M., Belle, V., Woudstra, M., Carriere, F., Guigliarelli, B., Fournel, A., and Longhi, S. 2006. Assessing induced folding of an intrinsically disordered protein by site-directed spin labeling electron paramagnetic resonance spectroscopy. *J. Phys. Chem*. 110:20596-20608.
17. Belle, V., Rouger, S., Costanzo, S., Liquiere, E., Strancar, J., Guigliarelli, B., Fournel, A., and Longhi, S. 2008. Mapping α-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins* 73:973-988.
18. Brucker, E.A., Olson, J.S., Phillips, G.N., Dou, Y., and Ikeda-Saito, M. 1996. High-resolution crystal structures of the deoxy, oxy, and aquomet forms of cobalt myoglobin. *J. Biol. Chem*. 271:25419-25422.
19. Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R.M., and Schlichting, I. 1999. Crystal structures of myoglobin-ligand complexes at near atomic resolution. *Biophys. J.* 77:2153-2174.

20. Ösapay, K., Theriault, Y., Wright, P.E., and Case, D.E. 1994. Solution structure of carbomonoxy myoglobin determined from nuclear magnetic resonance distance and chemical shifts constraints. *J. Mol. Biol*. 244:183-197.

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

22. Eliezer, D., Yao, Y., Dyson, H.J., and Wright, P. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Struc. Biol*. 5: 148-155.

23. Lecomte, J.T.J., Kao, Y.H., and Cocco, M.J. 1996. The native state of apomyoglobin described by proton NMR spectroscopy: The A-B-G-H interface of wild-type sperm whale apomyoglobin. *Proteins J*. 25:267-285.

24. Cavagnero, S., Theriault, Y., Narula, S.S., Dyson, H.J., and Wright, P.E. 2000. Amide proton exchange rates for sperm whale myoglobin obtained from ^{15}N-^{1}H NMR spectra. *Protein Sci*. 9:186-193.

25. Eliezer, D., Chung, J., Dyson, J. H., and Wright, P.E. 2000. Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin. *Biochemistry* 39:2894-2901.

26. Yao, J., Chung, J., Eliezer, D., Wright, P.E., and Dyson, J.H. 2001. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry* 40:3561-3571.

27. Lietzow, M. A., Jamin, M., Dyson. J.H., and Wright, P.E. 2002. Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* 322: 655-662.

28. Meinhold, D. and Wright, P.E. 2011. Measurements of protein unfolding/refolding kinetics and structural characterization of hidden intermediates by NMR relaxation dispersion. *Proc. Natl. Acad. Sci.* 108:9078-9083.
29. Griko, Y.V., Privalov, P.L., Venyaminov, S.Y., and Kutyshenko, V.P. 1988. Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol*. 202:127-138.
30. Eliezer, D., Jennings, P.A., and Wright, P.E. 1995. The radius of gyration of an apomyoglobin folding intermediate. *Science* 270:487-488.
31. Wang, F. and Tang, X. 1996. Conformational heterogeneity and stability of apomyoglobin studied by hydrogen/deuterium exchange and electrospray ionization mass spectrometry. *Biochemistry* 35:4069–4078.
32. Jennings, P.A. and Wright, P.E. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* 262:892-896.
33. Uzawa, T., Nishimura, C., Akiyama, S., Ishimori, S., Takahashi, S., Dyson, J.H., and Wright, P.E. 2008. Hierarchical folding mechanism of apomyoglobin revealed by ultra-fast H/D exchange couples with 2D NMR. *Proc. Natl. Acad. Sci.* 105:13859-13864.
34. Hughson, F.M., Wright, P.E., and Baldwin, R. L. 1990. Structural characterization of a partially folded apomyoglobin intermediate. *Science* 249:1544-1548.
35. Lin, L., Pinker, R. J., Forde, K., Rose, G. D. & Kallenbach, N. R. (1994). Molten globular characteristics of the native state of apomyoglobin. *Nat.*
36. *Struct. Biol.* 1:447–452.
37. Nishimura, C., Dyson, H.J., Wright, P.E. 2002. The apomyoglobin folding pathway revisited: Structural heterogeneity in the kinetic burst phase intermediate. *J. Mol. Biol*. 322:483-489.

38. Felitsky, D.A., Lietzow, M.A., Dyson, J.H., and Wright. P.E. 2008. Modeling transient collapse states of an unfolded protein to provide insights into early folding events. *Proc. Natl. Acad. Sci.* 105:6278-6283.

39. Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, J.H., and Wright, P.E. 1999. Quench-flow experiments combined with mass spectrometry show apomyoglobin folds through an obligatory intermediate. *Protein Sci* 8:45-49.

40. Armstrong, B.D., Choi, J., López, C., Wesener, D.A., Hubbell, W., Cavagnero, S., and Han, S. 2011. Site-specific hydration dynamics in the nonpolar core of a molten globule by dynamic nuclear polarization of water. *J. Am. Chem. Soc.* 113:5987-5995.

41. Jennings, P.A. Stone, M.J., Wright, P.E. 1995. Overexpression of myoglobin and assignments of it amide, C^{α} and C^{β} resonances. *J. Biomol. NMR.* 6: 271-276.

42. Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. *Science* 181:223-230.

43. Chow, C., Kurt, N., Murphy, R.M., and Cavagnero, S. 2006. Structural characterization of apomyoglobin self-associated species in aqueous buffer and urea solution. *Biophys. J.* 90:298-309.

44. Griko, Y.V. and Privalov, P.L. 1994. Thermodynamic puzzle of apomyoglobin unfolding. *J. Mol. Biol.* 235:1318-1325.

45. Qu, K. B., Vaughn, J. L., Sienkiewicz, A., Scholes, C. P., and Fetrow, J. S. 1997. Kinetics and motional dynamics of spin-labeled yeast iso-1-cytochrome *c*: 1. Stopped-flow electron paramagnetic resonance as a probe for protein folding/unfolding of the c-terminal helix spin-labeled at cysteine 102. *Biochemistry* 36:2884−2897.

46. Langen, R., Cai, K., Altenbach, C., Khorana, H.G., and Hubbell, W.L. 1999. Structural features of the C-terminal domain of bovine rhodopsin: a site-directed spin-labeling study. *Biochemistry* 38:7918-7924.

47. Mohana-Borges, R., Goto, N.K., Kroon, J.G.A., Dyson, J.H., and Wright, P.E. 2004. Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.* 340:1131-1142.

48. Nishimura, C., Lietzow, M.A., Dyson, J.H., and Wright, P.E. 2005. Sequence determinants of a protein folding pathway. *J. Mol. Biol*. 351:383-392.

49. Pannier, M., Veit, S., Godt, A., Jeschke., G., and Speiss, H.W. 2000. Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn. Reson*. 142:331-340.

50. Jeschke G. and Polyhach, Y. 2007. Distance measurements on spin-labeled biomacromolecules by pulsed electron paramagnetic resonance. *PCCP* 9:1895-1910.

51. Fleissner, M.R., Brustad, E.M., Kálai, T., Altenbach, C., Schultz, P.G., and Hubbell, W.L. 2009. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci*. 106:21637-21642.

52. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011 Structure and dynamics of a conformationaly constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246.

53. Hambly, D.M. and Gross, M.L. 2005. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent accessible residues on the microsecond timescale. *J. Am. Soc. Mass. Spec.* 16:2057-2063.

54. Meinhold, D. and Wright, P.E. 2011. Measurements of protein unfolding/refolding kinetics and structural characterization of hidden intermediates by NMR relaxation dispersion. *Proc. Natl. Acad. Sci.* 108:9078-9083.

55. Columbus, L., Kálai, T., Jeko, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure. *Biochemistry* 40:3828-3846.

56. Fleissner, M.R., Cascio, D., and Hubbell, W.L. 2009. Structural origins of weakly ordered motion in spin-labeled proteins. *Protein Sci.* 18:893-908.

Chapter 6: Effect of hydrophobic core substitutions on the molecular flexibility of T4 lysozyme studied by site-directed spin labeling

6.1 Summary

It has been well-established that proteins rely on flexibility (*i.e.*, backbone motions and conformational changes) to perform their function. A series of studies have suggested that protein stability and the dynamical properties of proteins are intimately related with hydrophobic core packing. The studies presented here seek to evaluate the relationship between protein core packing, stability, and flexibility by studying the effect of native core repacking substitutions on the flexibility of a well-characterized protein using SDSL. To this end 28 core substitutions in T4L were engineered and the R1 side chain was introduced at strategically selected sites in order to monitor the local changes in backbone flexibility. The structural and thermodynamic effects of most of the core substitutions studied here have been previously characterized, but to our knowledge little information exists about the dynamical properties of those mutants in solution. With few exceptions, most of these core mutations excite either fast backbone motions (ns to ps) and/or slow conformational fluctuations (μs to ms) of the protein at the local level. Remarkably, the same type of mutation, say large-to-small substitutions, can excite motions on different time domains suggesting that the local environment may play a major role in determining the effects of the core repacking mutations. Given the role that protein motions can play in function, it would be worth exploring how core substitutions can be used in protein rational design strategies as a tool to introduce functional protein dynamics.

6.2 Introduction

The realization that intrinsic protein motions occurring in the picosecond to second time scale are essential to protein functions such as signal transduction, catalysis, and allostery has raised new questions regarding the relation between protein stability, flexibility, and function. For example, many isolated thermophilic enzymes are highly homologous to the structure and function to their mesophilic counterpart, however the thermophilic enzymes are optimally active at high temperatures (~ 70 – 125 ℃), yet they are relatively inactive at temperatures where the activity of the mesophilic enzyme is optimal (*see* Vieille and Zeikus, [1] and references therein). Structural studies on thermophilic and mesophilic homologs have revealed only minor structural changes [2], while functional assays have shown that, with some exceptions, the global flexibility of the thermophilic and mesophilic enzymes are almost identical at their optimum temperatures, but significantly different at mesophilic temperatures [3-4]. Based on these studies one mechanistic interpretation for the low activity of most thermophilic enzymes at moderate temperatures is that thermophilic enzymes are relatively "rigid" at moderate temperatures, such that conformational motions required for function are absent or damped [5], which in turn impairs their activities [4]. Based on these and others studies (*see* review by Telium et al. [6]) it has been generally assumed that there is a balance between functional protein flexibility and stability. One of the implications of the balance between protein stability and flexibility is that nature may control flexibility to optimize function and even create new functions using stabilizing/destabilizing mutations [6-7].

Studies on the relation between stability, function, and flexibility are limited but they suggest that residues that are important for function are not necessarily optimal for protein stability (*see* Schoichet et al. [8] and references therein), and mutations that disrupt the native

hydrophobic core packing of proteins can stabilize/destabilize proteins by several kcal/mol with a concomitant decrease/increase in the conformational flexibility [9-13]. The latter finding can be rationalized by careful examination of the interior packing of proteins, which reveals that although the core of most proteins are well-packed [14], the interior packing can be heterogeneous and suboptimal and is best described as "nuts and bolts in a jar" rather than a jigsaw puzzle [15]. Analysis of high-resolution structures revealed that most proteins contain cavities and pockets that can range from a few to hundreds of cubic angstroms [15], which in general are expected to be destabilizing due to the loss of van der Waals interaction [16-17]. The types of core mutations that have been introduced to evaluate the effect of core substitutions on the stability and flexibility of proteins include cavity-creating and cavity-filling mutations, substitutions that introduce strain, buried charge substitutions, and buried ion pair mutations.

The primary goal of this study is to evaluate the effects of substitutions that disrupt the native hydrophobic core packing interactions on the conformational flexibility of a well-characterized protein with SDSL. For this study T4 lysozyme (denoted T4L hereon) was selected as a model system because it has served as a paradigm to understand the relationship between structure, stability, and folding [18]. Over 700 mutants have been engineered and studied in great detail, 350 of which have been crystallized in different space groups, including a myriad of variants designed to disrupt and repack the hydrophobic core [18-24]. T4L consists of two domains as shown in Figure 6.1, namely the N-terminal domain consisting of β – sheets 1-4 and α - helix B, while the helical C-terminal domain consists of α – helix A and helices D-J. The two domains are connected by the long α-helix C. The largest hydrophobic core of the protein is located within the C-terminal domain. As depicted in Figure 6.2, the packing of the hydrophobic core in the C-terminal domain is not optimized as there are three cavities (identified I-III in

Figure 6.2), two of which are apparently empty as judged by lack of electron density in the high-resolution structure of the WT protein (PDB: 3lzm), while cavity III is occupied by a solvent molecule. With a few exceptions, most of the mutants engineered in the C-terminal domain of T4L by Brian Matthews and co-workers were designed around these cavities in order to increase [16, 22-24] or reduce their size [19]. In addition some of the substitutions created novel binding sites for hydrophobic ligands such as benzene, ethylbenzene, and xylene. Thermal denaturation studies revealed that most of these substitutions destabilized the protein by up to ~ 6 kcal/mol [23]. Notwithstanding this reduction in stability, remarkably the high-resolution structures of many of these mutants revealed that the protein is tolerant to a variety of core repacking mutations as judged by the minimal structural changes observed [25]. For instance the large-to-small mutation at site 99 (L99A) destabilized the protein by 5 kcal/mol, however the structure of the mutant protein is remarkably similar to the WT protein [16]. Careful examination of this variant showed little structural relaxation in response to the substitution, which in turn resulted in a larger than expected cavity I with a volume of 150 $Å^3$ [23]. Interestingly, other core repacking mutants showed some structural relaxation, suggesting that the structural response to cavity-creating mutations depends on the local plasticity [24]. Based on a myriad of cavity-creating mutations in the core of T4L, Matthews and co-workers found an empirical relation between the reduction in protein

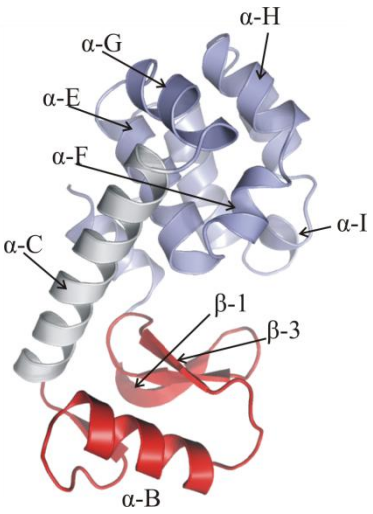


Figure 6.1: Structure of T4L. The N- and C-terminal domains of the protein are color-coded red and light blue, respectively. The interdomain helix-C is shown in gray. Some of the structural elements are indicated.

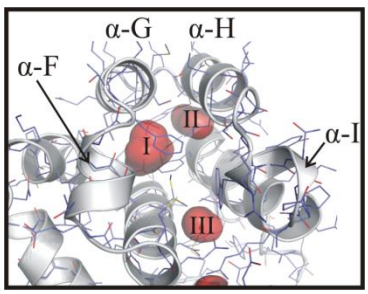


Figure 6.2: Internal cavities in the hydrophobic core of T4L.

stability and the increase in cavity size associated with a given replacement.

$$\Delta\Delta G = a + b\Delta V, \quad (6.1)$$

As shown in equation 6.1, the relationship consists of two terms. The first term denoted *a* is constant energy term that corresponds to the transfer free energy difference (*i.e.,* solvent to protein interior) between the mutated amino acid and the native residue, and the second energy term (*b*) depends on the size of the cavity created by the mutation (ΔV) and is equal to - 24 cal mol^{-1} $Å^{-3}$ [16]. This empirical relation shows that reducing the interior packing in the core of proteins not only has a hydrophobic penalty (*i.e.,* reduction in the hydrophobic effect), but also a second energy term that depends on the way the structure responds to the substitution, which in turn is determined by the local plasticity [16].

Notwithstanding the large amount of structural information obtained for these mutants, save for the L99A cavity variant [9,16,26], the conformational flexibility of most of these variants has not been characterized in solution, thus SDSL provides a means to characterize the conformational properties of these mutants in solution in order to evaluate the relationship between stability and flexibility. The types of core repacking substitutions studied in this work are shown in Figure 6.3 and include: (1) large-to-small mutations (cavity-creating), (2) small-to-large mutations (cavity filling and overpacking), (3) buried charge substitutions, and (4) compensating substitutions (packing and ionic pair). For each mutant, the R1 side chain was introduced at strategically selected sites in order to monitor the local changes in backbone flexibility.

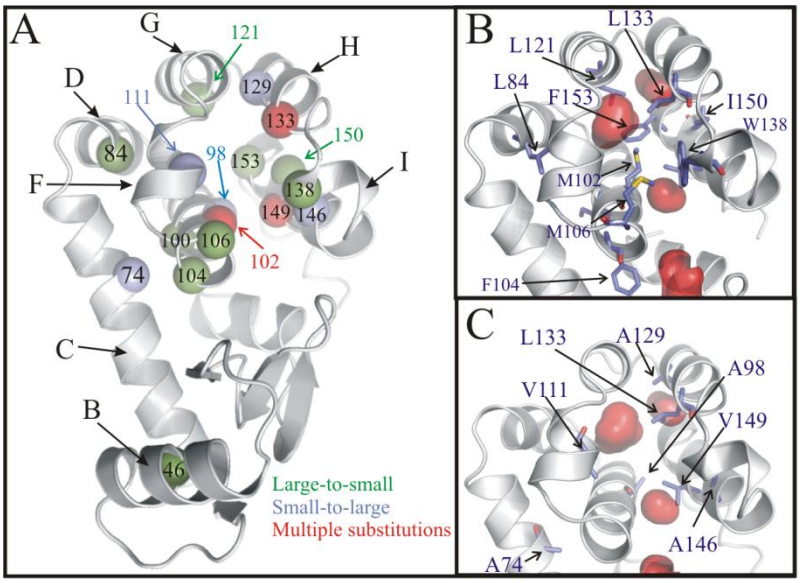


Figure 6.3: Sites where packing disrupting mutations were introduced. A) Types of substitutions introduced on each site are color coded. B) Sites selected for large-to-small mutations. C) Sites selected for small-to-large mutations. The external surfaces of the native cavities are shown in red.

6.3 Large-to-small substitutions

6.3.1 *M102A and F104A*. The M102A and F104A substitutions decrease the stability of the protein by 3.3 and 3.1 kcal/mol at pH 3, respectively [23]. The x-ray structures of both mutants were very similar to WT with root-mean-square shifts of the main chain atoms of 0.15 Å [23]. The M102A substitution resulted in an increase in the volume of the pre-existing cavity I from 34 $Å^3$ to 83 $Å^3$ as shown in Figure 6.4, which is 26 $Å^3$ less than expected assuming no structural change, however rotation and translations of the side chain of residues M106, F114, W138, and F153 reduced the size of the cavity from the expected 109 $Å^3$ to the observed 83 $Å^3$. Even though the volume of the engineered cavity is smaller than that of benzene (96 $Å^3$), it was determined that the M102A mutant does bind benzene with a K_d ~ 400μM as judged by thermal upshift assay, x-ray crystallography, and NMR spectroscopy [23, 27].

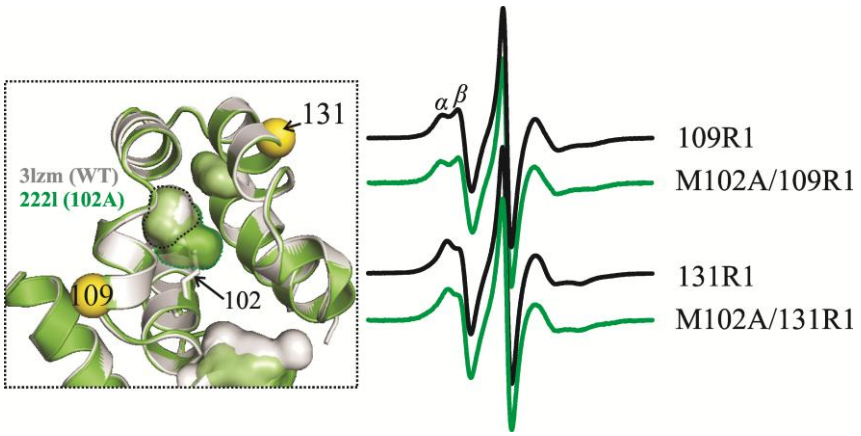


Figure 6.4: Effect of M102A mutant on the flexibility of T4L. *Left panel*: Overlay of the high-resolution structures of T4L WT (gray) and M102A mutant (green). The external surfaces of the native and mutant cavities are shown in gray and green, respectively. The pdb codes for the structure of the WT and mutant proteins are shown. The spheres at the Cα indicate the sites where R1 was introduced. *Right panel*: EPR spectra in the WT and M102A background are shown as black and green traces, respectively. The two components in the spectrum of residue 109R1 are indicated as α and β.

The F104A substitution created a crevice between the N- and C-terminal domains of the protein as depicted in Figure 6.5. The high-resolution structure showed slight shifts of the main chain atoms of residues 70-80 and 105-113 of 0.3 and 0.7 Å, respectively. Unlike the M102A substitution the engineered void is not empty as there is a new water molecule (Sol 401) that occupies the cleft and makes a hydrogen bond with the carbonyl group of residue I29 and two additional water molecules moved to the opening of the engineered crevice [23]. The F104A crevice is also a binding site for benzene as shown by x-ray crystallography.

To evaluate the effect of the M102A and F104A substitution on the flexibility of T4 lysozyme in solution, the EPR spectra of residues 109R1 and 131R1 were recorded in the M102A background and in the wild-type background, and the spectra of residues 72R1 and 109R1 were recorded in the F104A and wild-type background. For the M102A mutant, residues 109R1 of helix F and 131R1 of helix H were selected as both helices have hydrophobic residues (V111 and L133) that pack against the enlarged cavity I. The EPR spectra of residues 109R1 and 131R1 in the wild-type background are shown as black traces in Figure 6.4 and are similar to

those previously published [28]. As shown in Figure 6.4 the M102A substitution has no effect on the spectra of residues 109R1 and 131R1 indicating that the flexibility of helices F and H is not affected by the M102A mutation, although the presence of new excited states with a population lower than ca. 5 %, similar to that observed in the L99A mutant by NMR relaxation dispersion methods [29] (*see* section 1.21 in Chapter 1 of this dissertation), cannot be excluded (*see* Discussion).

For the F104A mutant, residues 72R1 of helix C and 109R1 of helix F were selected as sensor sites since both regions showed small structural shifts in response to the mutation as discussed above. The EPR spectra of residues 72R1 and 109R1 in the wild-type background and in the F104A background are shown in Figure 6.5. The EPR spectrum of residue 72R1 in the in the wild-type background is similar to that previously published [28]. The two peaks in the low field in the spectrum of residue 72R1 that might be interpreted as two components are the parallel and perpendicular components of the hyperfine extrema usually observed for z-axis anisotropic motion with an order parameter (S) $\geq$ 0.3 [30]. As shown in Figure 6.5 the F104A substitution

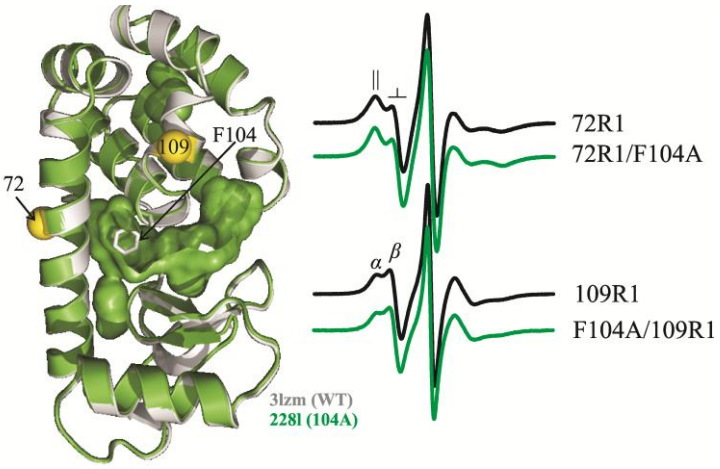


Figure 6.5: Effect of F104A mutant on the flexibility of T4L. *Left panel*: Overlay of the high-resolution structures of T4L WT (gray) and F104A mutant (green). The surface of the engineered solvent-exposed crevice is shown. The spheres at the Cα indicate the sites where R1 was introduced. *Right panel*: EPR spectra in the WT and F104A background are shown as black and green traces, respectively. The symbols ‖ and ⊥ in the spectrum of residue 72R1 indicate the parallel and perpendicular component of the hyperfine extrema. The two components in the spectrum of residue 109R1 are indicated.

has no effect on the spectrum of residue 109R1 and little effect on the spectrum of residue 72R1 as judged by the scaled mobility (Ms)[1]. For example, the difference in Ms for 72R1 in the WT and F104A background is 0.01, which is within the experimental error. However, the 2Azz' value increased by 0.8 G in the F104A mutant, which suggests that there is a slight decrease in mobility. The small increase in 2Azz' in the F104A mutant likely reflects a slight increase in the order of motion since the Ms values, which are mainly determined by rate of motion as discussed in chapter 4, are almost identical. Similar to the M102A mutant, the presence of new excited states with a population lower than ca. 5 % cannot be ruled out.

6.3.2 *L46A.* The L46A substitution is the only substitution engineered in the N-terminal domain of the T4L molecule in this study. The native leucine at position 46 is located in helix B and is solvent inaccessible. The stability and structure of the L46A mutant has been previously characterized [16], thus it provides a good model to study the local changes in flexibility by EPR spectroscopy. The L46A substitution destabilizes the protein by 2.7 kcal/mol at pH 3. The high-resolution structure of the mutant showed that the substitution created a new cavity within the N-terminal domain with a volume of 24 $Å^3$ as depicted in Figure 6.6. In addition, the structure shows small structural changes between the WT and the mutant proteins that reduced the size of the cavity by 23.5 $Å^3$ from the expected value assuming no structural relaxation [16]. The L46A cavity does not bind benzene likely because its volume is too small (25 % the volume of benzene).

To evaluate the effect of the L46A on the flexibility of the B-helix of T4 lysozyme in solution, the EPR spectrum of residue 48R1 was recorded in the L46A background and in the wild-type background. The EPR spectrum of residue K48R1 in the WT background was

[1] Scaled mobility is a measure of mobility based on the spectral central linewidth as discussed in chapter 4 of this dissertation.

previously recorded [28] and is reproduced here. As shown in Figure 6.6B, the spectrum in the L46A background is similar to that in the WT background, except for a small increase in the spectral intensity corresponding to an immobile component. The EPR spectra of residue K48R1 in the WT background was fit reasonably well using the MOMD model with a single component having a z-axis anisotropic motion with τ = 1.54 ns and S = 0.22 (Figure 6.6C), however the fit for the spectrum in the L46A background required a second component (8%) with restricted mobility (τ = 8.3 ns).

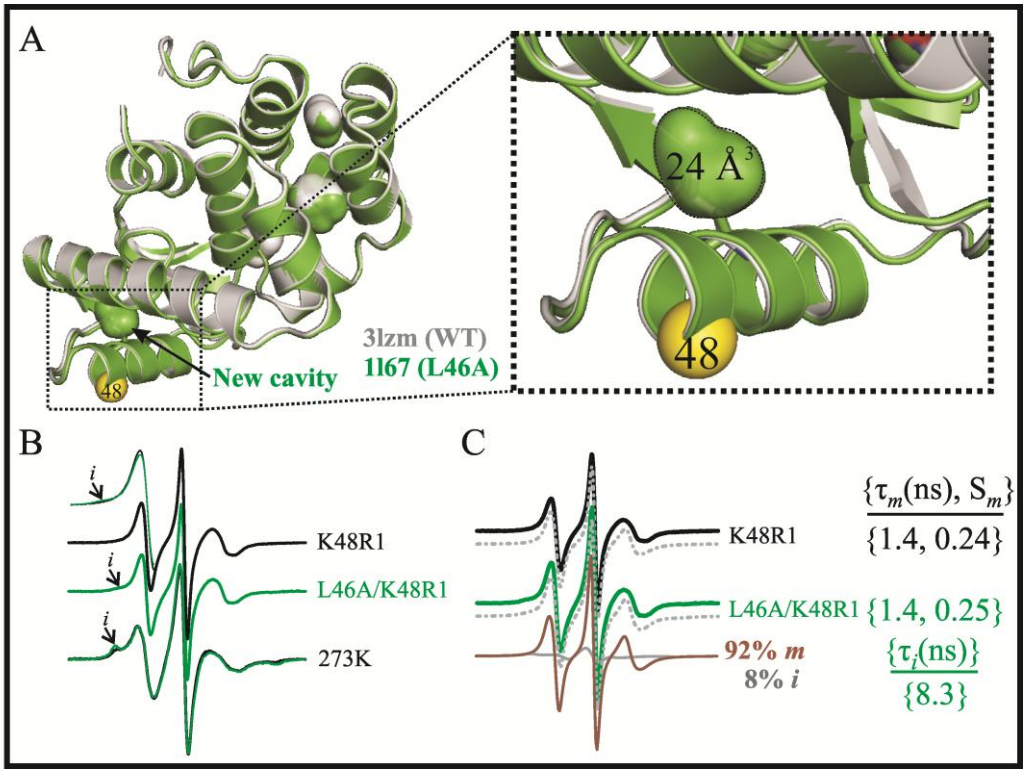


Figure 6.6: Effect of L46A mutant on the flexibility of the B-helix of T4L. A) Overlay of the high-resolution structures of T4L WT (gray) and L46A mutant (green). The surface of the engineered cavity is shown. The sphere at the Cα indicates the site where R1 was introduced. B) EPR spectra of K48R1 in the WT (black trace) and L46A (green trace) background at 295K and 273K. C) Experimental and MOMD fits (gray traces) of K48R1 in the WT and L46A background are shown. The parameters obtained from the MOMD simulations are shown.

To investigate whether the immobile component can be detected to a higher level at lower temperatures, the EPR spectra were recorded at 273K. Interestingly, the spectra of K48R1 in both the WT and L46A background at 273K showed an immobilized state, but with the L46A mutant reflecting a higher population of the immobile state as judged by the relative spectral intensities of the low field region (Figure 6.6B). For K48R1 in the WT background at low temperatures, the nitroxide must make contact with nearby residues, but the interaction must be weak that increasing the temperature effectively disrupts the interaction with the environment or reduces its population to levels below detection. At this point it is not clear if the immobile component observed in the L46A background is the same to that observed in the WT background at low temperature, but with an increased population or it is rather an entirely new conformation of the helix that introduced new interactions of the nitroxide side chain with the protein. Additional studies involving interspin distance measurements between residue 48R1 with several reference sites will be needed to distinguish these two possibilities. However it is clear that the L46A substitution affects the environment around the nitroxide side chain at position 48 at room temperature.

6.3.3 *M102A/M106A*. This double mutant was originally designed by Baldwin and co-workers [23] with the aim of creating a binding site that could bind ligands that are more polar than benzene. The M102A/M106A mutant destabilized the protein by 4.3 kcal/mol at pH 3 [23] and the high-resolution structure showed that the substitutions created a channel that connects the interior of the C-terminal domain of the protein to the solvent. The engineered channel is not empty as there is a chain of water molecules that extends from the hydrophobic core to the active site of the protein. In addition, there are substantial structural changes in the region corresponding to the F-helix as shown in Figure 6.7A, which adopts two conformations: one

similar to the WT conformation and a second non-helical S-shaped conformation [23]. Interestingly, the M102A/M106A mutant apparently is not a ligand binder due to the structural rearrangements of the F-helix.

To investigate the changes in the protein flexibility in solution of helices in the C-terminal domain due to the engineered solvent accessible channel, the EPR spectra of residues 109R1 and 131R1 were recorded and compared to that obtained in the WT background. As discussed above, the EPR spectrum of residue 109R1 in the WT background reflects two dynamic components. In the 102A/106A background the spectrum of 109R1 shows a reduction in the scaled mobility from 0.64 to 0.58 and the spectral intensity of the α-component is shifted towards the lower field as shown in Figure 6.7B.

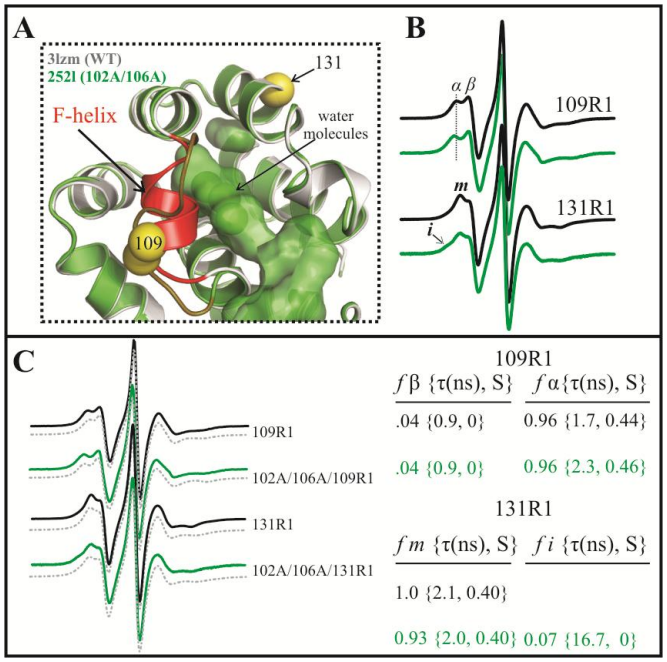


Figure 6.7: Effect of 102A/106A mutant on the flexibility of helices F and H of T4L. A) Overlay of the high-resolution structures of T4L WT (gray) and the double mutant (green). The surface of the engineered channel is shown. The spheres in the channel identify some of the water molecules inside the core. The yellow spheres at the Cα indicate the sites where R1 was introduced. The two conformations of the F-helix are shown. B) EPR spectra in the WT (black trace) and 102A/106A (green trace) background at 295K. The dashed gray line in the spectra of 109R1 is to help guide the eye. The relatively immobilized component observed in residue 131R1 in the 102A/106A mutant is indicated. C) Experimental and MOMD fits (gray traces) of 109R1 and 131R1 in the WT and 102A/106A background are shown. The parameters obtained from the MOMD fits are shown

The EPR spectrum of residue 131R1 in the 102A/106A mutant showed an immobilized spectral component not observed in the spectrum recorded in WT background. The spectra of both mutants recorded in the temperature range of 273 – 313K revealed that the spectral intensity corresponding to an immobile component in the 102A/106A mutant is higher than that of the WT background at every temperature (*see* Appendix).

To analyze the spectral changes in a more quantitative manner, the spectra of 109R1 and 131R1 in the WT and 102A/106A were fit to the MOMD model. The spectrum of 109R1 in the WT background was fit reasonably well with a 4% population of a fast isotropic component (τ = 0.9 ns) and a dominant anisotropic component with correlation time and order parameter similar to those observed in non-interacting solvent exposed sites of α–helices (τ = 1.7 ns, S = 0.44). The origin of the isotropic component in the spectrum of 109R1 in helix F in the WT background is not clear, but several studies have revealed that the short helix F has low protection factors [31], exhibits relatively high B-factor values (mean value = 32 $Å^2$), and is the most plastic region of the C-terminal domain as judged by its response to various mutations throughout the C-terminal domain [21, 23-24]. Thus it is possible that the fast component arises from a second more disordered conformer of the helix, but this remains to be proven.

A reasonably good fit of the spectrum of 109R1 in the 102A/106A mutant was obtained using the populations for the α and β components, but with a lower mobility of the nitroxide for the α component (Figure 6.7C). It is intriguing that the spectral changes observed for 109R1 with the crevice were smaller than anticipated considering the structural heterogeneity observed for the F-helix in the x-ray structure of the 102A/106A mutant [23].

The spectrum of residue 131R1 in the WT background was fit reasonably well with a single component having a z-axis anisotropic motion (τ = 2.1 ns, S = 0.4), whereas the simulation in the

102A/106A background required a new component (7 %) with lower mobility (τ = 16.7). It should be noted that the rate and order of motion of the mobile component was almost identical, within the errors of the fits for the WT and 102A/106A background suggesting that the mutation did not affect fast backbone motion of the H-helix, but it rather introduced a new state that is likely in slow exchange (lifetime > 100 ns).

6.3.4 *L121A and L133A*. The individual L→A substitution at the buried positions 121 and 133 reduced the stability of the protein by 2.7 and 3.6 kcal/mol, respectively at pH 3 [16]. The x-ray structures of both mutants were determined to high resolution by Eriksson et al. [16] and Baldwin et al. [22]. For the L121A mutant, the structure exhibits some changes in regions flanking the mutation site. As depicted in Figure 6.8 the C-terminus of the J-helix "collapses" towards the interior by about 1 Å and the side chain of residue F153 changed its conformation moving the ring of the phenylalanine side chain closer to the mutated site. Other helices also shifted to a smaller extent towards the core of the protein, which reduced the size of the engineered cavity from the expected ~149 Å^3 assuming no structural relaxation to the observed 75 Å^3. As shown in Figure 6.8 the size of the pre-existing cavity I increases from 34 Å^3 to 67 Å^3, the native cavity II disappears, and a new cavity with a volume of 8 Å^3 is created.

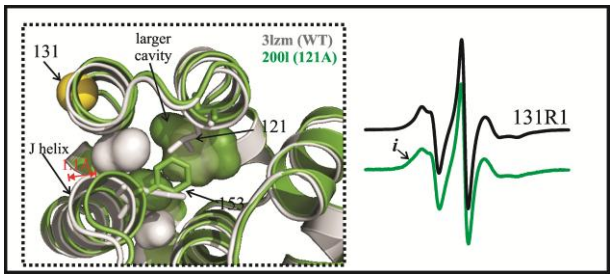


Figure 6.8: The L121A mutant. Left panel: Overlay of the x-ray structures of WT (gray) and 121A mutant (green) of T4L. The surfaces of the native and mutant cavities are shown. Some of the largest structural changes in the protein in response to the 121A mutation are highlighted. Right panel: EPR spectra of 131R1 in the WT and 121A backgrounds. The relatively immobile component observed in the spectrum of 121A is indicated.

For the L133A mutant, the structural changes are smaller than those observed in the L121A structure, but the structure still "relaxes" to reduce the size of the engineered cavity from the

expected value of 173 $Å^3$ to 142 $Å^3$ [16]. As depicted in Figure 6.9, in the L133A mutant, the native cavities I and II coalesce to form a single large cavity. It should be noted that for both mutants, the engineered cavities are empty as judged by the lack of electron density of solvent molecules inside the core in the high-resolution structures. Interestingly, neither the 121A nor the 133A mutants appear to bind benzene even though the volume of the cavity in the 133A mutant is larger than that of benzene. It has been proposed that the lack of ligand binding for these two mutants is likely due to limited flexibility of the protein and/or inappropriate shapes of the putative binding sites [23].

To evaluate changes in the flexibility of the protein within the C-terminal domain by the engineered cavities, the EPR spectra of residue 131R1 were recorded in the L121A and in the L133A backgrounds and compared to the spectrum obtained in the WT background. As shown in Figures 6.8 and 6.9, for both mutants the spectra showed a relatively immobilized component not present in the WT background. To analyze the spectral changes in a more quantitative manner, the spectrum of 131R1 in the 133A mutant was simulated using the MOMD model. The spectrum was fit reasonably well with a two-component model, consisting of a 6 % population of a component with restricted mobility (τ = 16.7 ns) indicating tertiary interaction of the nitroxide with its environment and a

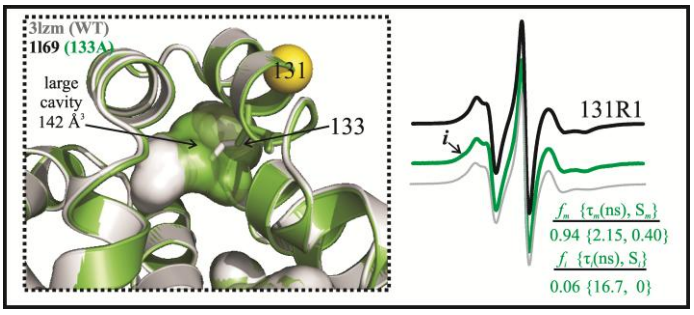


Figure 6.9: EPR study on the L133A mutant. *Left panel*: Overlay of the x-ray structures of WT (gray) and 133A mutant (green) of T4L. The external surfaces of the native and mutant cavities are shown. The sphere at the Cα indicates the site where R1 was introduced. *Right panel*: EPR spectra of 131R1 in the WT and 133A backgrounds. The relatively immobile component observed in the spectrum of 133A is indicated. The best fit for the spectrum in the 133A mutant is shown in gray. The parameters obtained from the best fit are shown in brackets.

mobile component with the same rate and order of motion as the mobile component observed in the WT background (τ = 2.15 ns, S = 0.4), which shows that L133A mutation did not affect fast backbone motion of the H-helix, but it rather introduced a new state that is in slow exchange.

6.3.5 *L121A/L133A*. The 121A/133A double mutant was originally designed with the aim of studying the correlation between the presence of relatively large internal cavities and protein stability [24]. Thermal unfolding of the double mutant showed a significant decrease in T_m of 19.6 °C compared to the WT form at pH 3, however the mutant apparently gradually unfolded over a wide range of temperature in a non-two-state manner, therefore the value of $\Delta\Delta G$ could not be reliably obtained [24]. The x-ray structure of the double mutant was solved at a resolution of 2.6 Å [24] and is shown in Figure 6.10A. The structural changes of the 121A/133A variant are similar to those obtained in the single mutants and involved a 0.8 Å inward movement of the J-helix towards the 121A site, slight rotations (~ 0.6 Å) of helices G and H, rotation of the F153 side chain, and ~ 1 Å movement of the main chain atoms of the F-helix towards the interior. In addition, cavities I and II coalesce to form a large flat-disk shape cavity with a volume of 172 Å and a small cavity of 8 $Å^3$, however the structural relaxation of the double mutant reduced the size of the large cavity by 54 $Å^3$ from the expected volume [24]. In contrast to the single 121A and 133A variants, the double mutant binds cyclic and linear nonpolar ligands such as benzene, hexane, dipropyl disulfide, and diisopropyl disulfide as judged by the thermal upshift assay [23].

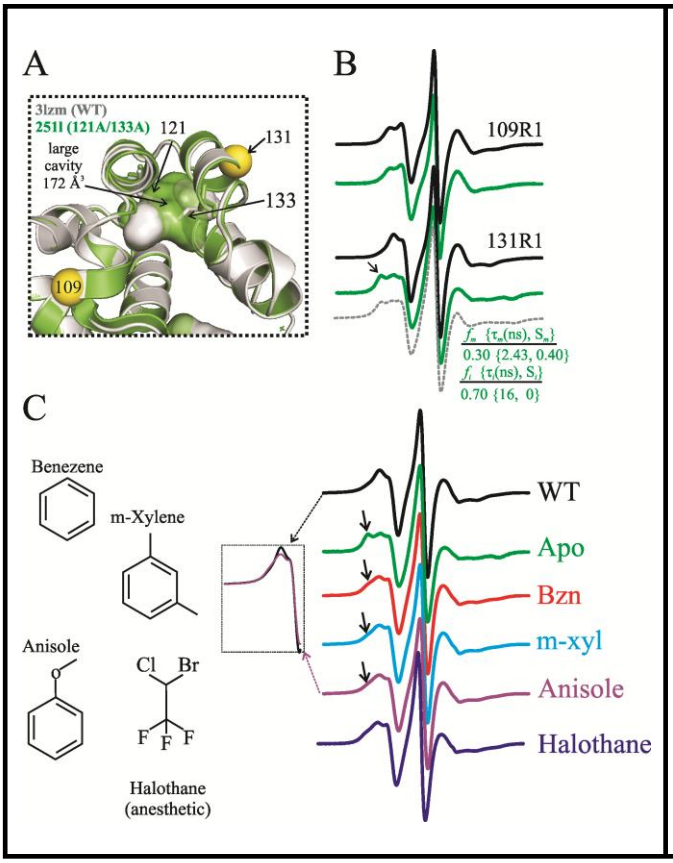


Figure 6.10: EPR studies on the L121A/L133A mutant. A) Overlay of the x-ray structures of WT (gray) and 121A/133A mutant (green) of T4L. The surfaces of the native and mutant cavities are shown. The spheres at the Cα indicate the sites where R1 was introduced. B) EPR spectra of 109R1 in the WT and 131R1 in the WT and 121A/133A backgrounds. The relatively immobile component observed in the spectrum of 121A/133A is indicated. The best fit for the spectrum in the 133A mutant is shown in gray dashed lines. The parameters obtained from the MOMD fits are shown in brackets. C) EPR spectra of 121A/131R1/133A in the presence of ligands (benzene, m-xylene, anisole, and halothane) shown in the left panel. **Binding of the general anesthetic halothane to the 121A/133A cavity mutant has not been reported prior to this work**. The EPR spectrum of 131R1 is shown for comparison. *Inset*: Overlay of the low field peaks of the EPR spectra of 131R1 in the WT and the cavity mutant in the presence of anisole.

To evaluate changes in the conformational flexibility of helices around the enlarged cavity, the EPR spectra of residues 109R1 and 131R1 was recorded in the 121A/133A background and the results are shown in Figure 6.10B. The EPR spectrum of 109R1 on helix F is identical to that obtained in the WT background, suggesting that the inward movement observed in the x-ray structure did not affect the flexibility of the helix in solution, however a new excited state of the helix populated at levels below detection (ca. < 5 %) can not be ruled out. On the other hand, the EPR spectrum of 131R1 in the 121A/133A background, revealed significant spectral changes compared to that obtained in the WT background as shown in Figure 6.10B. Specifically, a relatively immobile component is observed with similar spectral amplitude compared to that of the more mobile component. Osmotic-perturbation studies showed that the populations giving

rise to each spectral component are in slow exchange (lifetime > 100ns; *see* Appendix). Temperature-dependent EPR experiments (273 – 313 K) showed that the spectral changes remain at every temperature. In fact, the relatively immobile component observed at 273 K is also present at 313 K, which shows that the interaction giving rise to the immobile state is not disrupted by increasing temperatures up to 40 °C (*see* Appendix).

The experimental spectrum for residue 131R1 in the double mutant could be simulated with a 70 % immobile component (τ = 16 ns) and a 30 % mobile state with z-axis anisotropic motion with τ = 2.4 ns and S = 0.4. It should be noted that based on the crystal structure, the origin of the most dominant component in the spectrum of 131R1 is not clear since the slight rotation of the H-helix observed in the x-ray structure is not enough to change the environment around the nitroxide side chain. Thus, it is possible that in solution there is a second conformation of the H-helix or nearby helices, which brings the nitroxide side chain in tertiary contact with the protein environment. Such a conformation may be selected against by crystal lattice contacts, which might explain why it is not observed in the x-ray structure. Alternatively, it is possible that the nitroxide side chain at position 131 in the double mutant *induces* a second conformation of the H-helix whereby the nitroxide, which is slightly hydrophobic in character, resides inside the nonpolar cavity in order to reduce the size of the void. However the latter possibility is unlikely since site 131 is in the opposite surface of the enlarged cavity. Additional evidence that provides insights into the conformation giving rise to the immobile component in the 121A/133A variant in solution will be presented below (*see* section 6.6.1 below).

To investigate whether the binding of ligands to the 121A/133A cavity changes the dynamic equilibrium between the two states observed in residue 131R1, four aliquots of the protein

solution were incubated with either benzene, p-xylene, anisole, or halothane[2]. As a control, the 131R1 mutant in the WT background was also incubated with each of the ligands. The EPR spectrum of 131R1/133G in the holo form shows significant changes compared to that of the apo protein. Specifically, in all cases the spectral intensity corresponding to the immobile component is reduced suggesting a decrease in the population of the conformation giving rise to the immobile state. Interestingly, the spectrum of the cavity mutant protein in the presence of anisole was very similar to the WT background protein (see inset in Figure 6.10C). As expected, the EPR spectrum of 131R1 in the WT background protein showed no changes upon addition of any of the ligands (data not shown), which supports the contention that the spectral changes observed for the 121A/131R1/133A protein arise from ligand binding to the enlarged cavity.

6.3.6 *L133G*. Unlike the L133A variant, the L133G mutant was identified as a ligand binder [23]. The L133G mutation decreased the stability of the protein by 5.6 kcal/mol at pH 3. The x-ray structure of the 133G mutant was solved at high-resolution in the apo and holo-forms[3] [23] and is shown in Figure 6.11. The 133G substitution merged the native cavities I and II into a single large cavity with a volume of 176 $Å^3$, which is 18 $Å^3$ smaller that anticipated assuming no structural relaxation.

The EPR spectra of residues 109R1, 131R1, and 151R1 were recorded in the 133G background. As shown in Figure 6.11, the spectra of residues 109R1 and 151R1 are identical to that recorded in the WT background, while the spectrum of residue 131R1 showed significant changes caused by the 133G mutation analogous to the effect observed in the 121A/133A variant, but with a lower intensity. Osmotic-perturbation studies showed that the populations giving rise to each spectral component are in slow exchange (lifetime > 100ns; *see* Appendix).

[2] Binding of the general anesthetic halothane (volume = 123$Å^3$) has not been reported previous to this work.

[3] The 133G mutant binds ligands such as benzene and p-xylene at ~ mM affinities (Baldwin et al. [23]).

The spectrum of 131R1 in the 133G mutant was simulated to estimate the relative population of the immobile state and to investigate whether there are additional changes in the motion of the fast component. The spectrum was fit reasonably well with a 25 % population of an immobile component and 75 % of a mobile component with similar mobility to that obtained for 131R1 in the WT background (Figure 6.11).

To investigate whether the binding of ligands to the 133G cavity changes the equilibrium between the states giving rise to the two component spectra observed in residue 131R1, the 131R1/133G apo protein was incubated with various nonpolar ligands, including halothane. The spectral changes observed upon ligand binding were similar to those observed for the 121A/133A mutant.

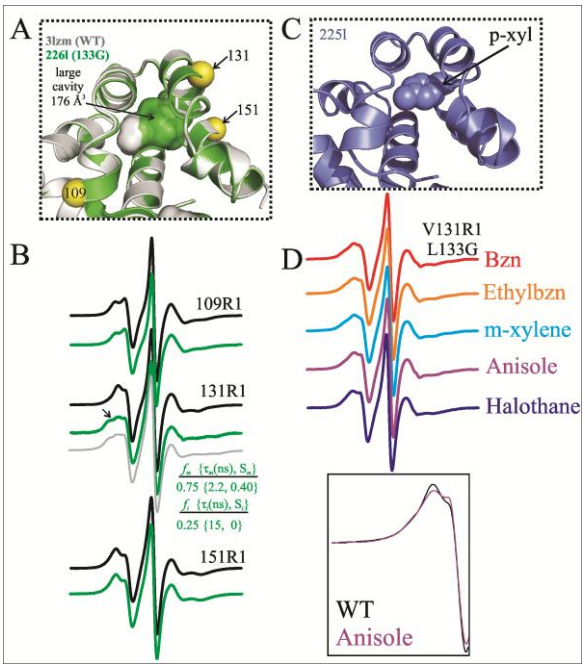


Figure 6.11: EPR studies on the L133G mutant. A) Overlay of the x-ray structures of WT (gray) and 133G mutant (green) of T4L. The external surfaces of the native and mutant cavities are shown. The spheres at the Cα indicate the sites where R1 was introduced. B) EPR spectra of residues 109R1, 131R1, and 151R1 in the WT and 133G backgrounds. The best fit for the 131R1 spectrum in the 133G variant is shown in gray. The parameters obtained from the best fit are shown in brackets. C) X-ray structure of L133G mutant bound to p-xylene (Baldwin et al. [23]). D) EPR spectra of 131R1/133G in the presence of the indicated nonpolar ligands. Binding of halothane to this mutant has not been reported previous to this work. *Inset*: Overlay of the low field lines of the spectra of 131R1 in the WT background and 131R1/133G bound to anisole.

6.3.7 *W138A*. The W138A variant was originally included in the initial screening for mutants that could potentially bind ligands [23], however it has not been as extensively characterized as the other mutants discussed above, so there is no structural information of this variant. The tryptophan at position 138 has a low solvent accessibility (3.9 %) and is localized in the short helix-I. Although there is no structural information available, modeling of the W138A mutant, suggests that the substitution should create an internal void of a volume of 233 $Å^3$ in between helices H, I, and J as depicted in Figure 6.12A. To elucidate the effect of such a reduction in the local packing around the helix I, the EPR spectrum of solvent exposed residue 140R1 was recorded in the 138A background. The EPR spectrum of 140R1 in the WT background was reported in previous studies [32] and is reproduced in this study. As shown in Figure 6.12B, there is a dramatic change in the spectrum of 140R1 in the 138A mutant compared to that obtained in the WT background. The scaled mobility value, which is determined by the central line width, is reduced from 0.8 to 0.2 and the hyperfine extrema are clearly resolved in the 138A variant, which is a hallmark for highly restricted motion of the nitroxide side chain on the nanosecond time scale indicating

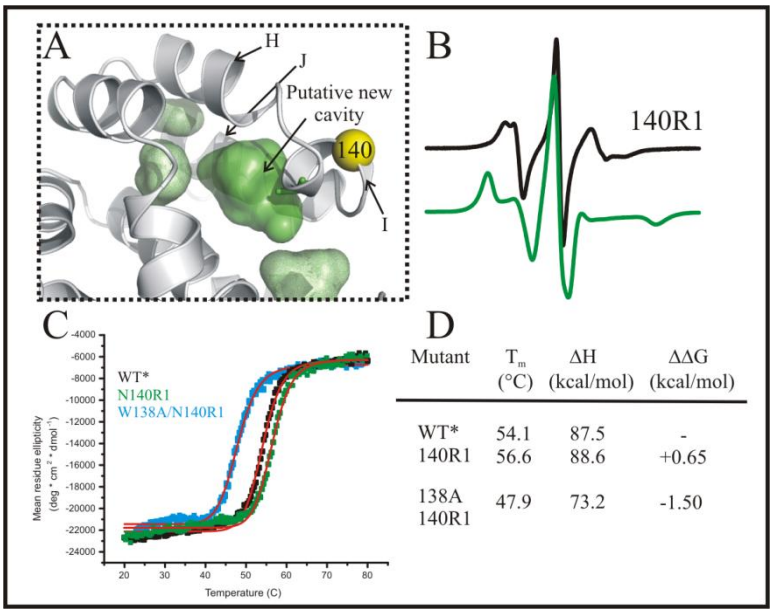


Figure 6.12: Effect of the W138A substitution on protein stability and EPR spectrum of residue 140R1. A) W138A model structure of T4L showing the surface of the putative cavity. The sphere at the Cα indicates the site where R1 was introduced. B) EPR spectra of residue 140R1 in the WT (black) and 138A (green) backgrounds. C) Thermal denaturation studies done in WT and mutant proteins. The CD experiments were performed at pH 3 as described in Chapter 8. D) Thermodynamic parameters obtained from the two-state model fits shown in panel C. ΔΔG was calculated using a ΔCp of 2.5 kcal/mol as described by Eriksson et al. [16].

extensive tertiary interaction with the environment. The spectral changes suggest that, at least, helix I undergoes significant structural changes in the 138A variant. Additional evidence that support this interpretation will be shown below.

To elucidate the changes in stability due to the W138A mutation relative to the WT and 140R1 protein, thermal denaturation studies were done for the WT, 140R1, and 138A/140R1 proteins (Figure 6.12C). Interestingly, the 140R1 mutant *increases* the stability of the protein relative to the WT protein by 0.65 kcal/mol, while the 138A/140R1 mutant *decreased* the stability of the protein by 1.50 kcal/mol relative to WT and by 2.22 kcal/mol relative to the 140R1 mutant.

6.3.8 *I100A, I150A, F153A, and V149A*. The structures and stabilities of three of these four mutants have been previously characterized [16,24], while the I150A is a new variant designed for this study. The native I100 residue is located in the center of the largely buried E-helix, has low solvent accessibility (5.6 %), and packs against residues I3, V75, and Y88. The I100A mutation decreases the stability of the protein by 3.4 kcal/mol and creates a small cavity (28 $Å^3$) between helices A, C, and E [24]. The structural changes observed in the mutant are subtle and involve slight mainchain rotation (~ 0.6 Å) of the last four residues of the C-helix (74-78) and side chain rotations of residue V75 and Y88 towards the cavity (< 1 Å).

To elucidate the local changes in flexibility due to the I100A substitution, the EPR spectrum of residue 72R1, which resides in the opposite surface of helix C that packs against the engineered cavity, was recorded in the 100A background and compared to that in the WT background (Figure 6.13A). The spectrum in the 100A background is still anisotropic, but exhibits increased mobility on the nanosecond time scale compared to that in the WT background as judged by the increase in the scaled mobility (Ms) value from 0.37 to 0.42 and a

reduction in the separation between the hyperfine extrema (2Azz') from 52.6 to 47.5. Since Ms is more sensitive to the rate of motion, while 2Azz' is more sensitive to the amplitude of motion [36], the increase in Ms and reduction of the 2Azz' values suggests an increase in the rate and amplitude of motion of the nitroxide side chain on the nanosecond time scale. Indeed, spectral simulation shows that the rate of motion decreased from 2.35 to 2.12 nanoseconds and the order of motion decreased from 0.50 to 0.42, corresponding to an increase in the angular amplitude of motion[4] of 3.2°. It has been proposed that residue 72R1 in the WT background has little contributions from fast backbone fluctuations [36], so it is likely that the increase in mobility of the nitroxide side chain for residue 72R1 in the 100A variant reflects an increase in nanosecond backbone motions of the C-helix, specifically the region nearby residue 72R1.

The native I150, located in the J-helix, has relatively low solvent accessibility (10 %), and packs against residues A130, W138, and Y139. Although structural information is not available for this variant, modeling shows that the I150A mutant should create a small new cavity or crevice (36 $Å^3$) in between helices H, I, and J. The EPR spectra of residues 131R1 and 151R1 on helices H and J, respectively was recorded in the 150A variant and is shown in Figure 6.13B. The EPR spectrum of 131R1 in the 150A mutant is dramatically different from 131R1. Specifically, the magnetic anisotropy for 131R1 is reduced in the 150A background as judged by the loss in the resolution of the parallel and perpendicular component of the A tensor. Interestingly, the scaled mobility value is similar to that obtained in the WT background (ΔMs = 0.01) suggesting a similar rate of motion. Indeed, MOMD simulation showed that a reasonable fit could be obtained with a τ= 2.2 ns, which is slightly higher to that obtained in the WT

[4] The amplitude of motion was calculated from the order parameter (S) using equation 2.38 (Chapter 2).

background (2.1 ns), and S = 0.28, which is lower than in the WT background (S = 0.4) corresponding to an increase in the angular amplitude of motion of ~ 5°.

The EPR spectrum of 151R1 in the 150A background also shows changes relative to that obtained in the WT background. Specifically, the parallel and perpendicular components of the A tensor are less well-defined indicating changes in the motion of the nitroxide. Quantitatively, the central linewidth in the I150A mutant increased from 3.43 to 3.53 G suggesting slower motion of the nitroxide in the nanosecond time scale. Indeed, spectral simulation revealed a reduction in the rate of motion of the nitroxide at site 151R1 in the I150A compared to that observed in the WT background. The spectral changes in residues 131R1 and 151R1 in the 150A variant likely reflect changes in the backbone motions of helices H and J (*see* Discussion). The thermal unfolding of the 131R1/150A mutant was compared to the 131R1 mutant and the WT protein in order to obtain the change in stability due to the I150A substitution (*see* Appendix). As expected, the large-to-small substitution reduced the stability of the protein by ~ 1 kcal/mol compared to the 131R1 background protein. It is intriguing that the reduction in stability is smaller than expected for a I→A substitution based on the hydrophobic effect (2.0 kcal/mol)[5].

[5] $\Delta\Delta G_{tr}$ from water to octanol of Ile relative to Ala is -2.0 kcal/mol (Fauchere and Pliska [35])

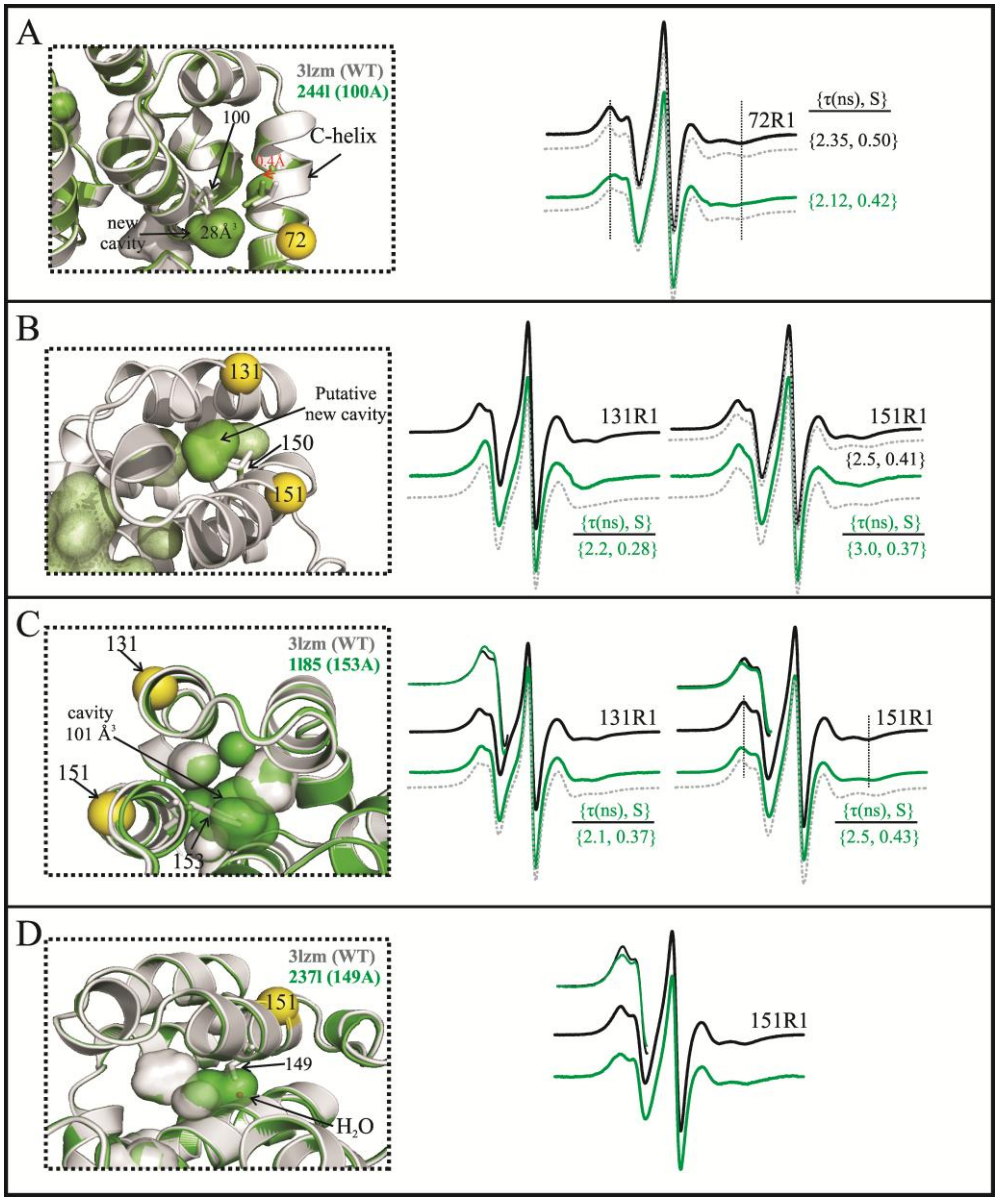


Figure 6.13: Large-to-small mutations changing the local fast backbone motions. A) I100A mutant. The ~ 0.4 Å rotation of the C-terminal residues of helix-C observed in the high resolution structure is highlighted. B) I150A mutant. A structural model showing the putative cavity is shown. C) F153A mutant. D) V149A mutant. The corresponding EPR spectra of the indicated sites in the WT (black) and mutant (green) backgrounds are shown in the right panel. The gray lines represent 2Azz'. The low field in the spectra of residues 131R1 and 151R1 in the 153A and 149A backgrounds has been amplified for clarity. The dynamic parameters obtained from MOMD simulations are shown in brackets. The pdb codes of the structures of the 100A, 153A, 149A, and WT proteins are shown (save for I150A for which there is no structural information).

The F153A mutation decreases the stability of the protein by 3.5 kcal/mol [16] and the high-resolution structure shows that the substitution increases the size of cavities I and II by a total volume of 79 $Å^3$. The most significant structural relaxation due to the mutation involves movement of the C-terminal end of helix J towards the core (0.8 Å) and rotation of the side chain of residue M102; both structural changes reduced the size of the engineered cavity [16]. The EPR spectra of residues 131R1 and 151R1, reporting on the flexibility of helices H and J, respectively showed subtle changes in the mobility of the nitroxide side chain (Figure 6.13C). The low field of spectrum of 131R1 in the 153A variant has a higher spectral amplitude reflecting slightly lower order than that obtained in the WT background as shown in Figure 6.13C, while the spectrum of residue 151R1 in the mutant reflects a small reduction in the mobility of the nitroxide as judged by reduced scaled mobility ($\Delta Ms = 0.04$), increased in the hyperfine extrema ($\Delta 2Azz' = 3G$), and by spectral simulation. As discussed above, the C-terminal region of the J helix "collapses" towards the core to reduce the size of the cavity, so it is likely that the slight reduction in the motion of the nitroxide for residue 151R1 in the 153A background reflects a local reduction on the nanosecond backbone motions of the J-helix.

The V149A mutation in helix J, reduces the stability by 3.2 kcal/mol [24] and creates a small internal cavity between helices A and J (14 $Å^3$). The high-resolution structure revealed little structural changes and that the cavity is completely filled by a water molecule that forms hydrogen bonds with several internal residues [24]. The EPR spectrum of residue 151R1 in the 149A variant was recorded to evaluate whether there are any changes in the mobility of the J-helix compared to WT protein. As shown in Figure 6.13D there is a subtle change in the low field of the spectrum in the 149 variant compared to that in the WT background suggesting that

there is a small, but detectable change in the backbone motions of the J-helix in the 149A mutant. It should be noted that unlike some of the large-to-small substitutions discussed above, the 100A, 150A, 153A, and 149V variants did not reveal new spectral components not observed in the WT background, but rather changes in the local dynamics on the nanosecond time scale.

6.3.9 *L99A/F153A*. The 99A/153A mutant is one of the most destabilizing cavity mutants in T4L; at pH 3, its stability is 8.3 kcal/mol lower than that of the WT protein [16]. The structural changes observed in the double mutant protein are a combination of those observed in the structures of the individual variants and thus the structure is very similar to that of the F153A mutant since the L99A mutant showed little structural relaxation [16]. The double mutant increased the volume of the pre-existing cavities I and II by a combined 207 $Å^3$ as depicted in Figure 6.14 creating a benzene-binding site.

To evaluate the changes in flexibility of this unstable variant, the EPR spectra of residues 131R1 and 151R1 were recorded and compared to those obtained in the WT background. As shown in Figure 14 the EPR spectra of both residues showed subtle changes compared to those obtained in the WT background. Specifically, for the 131R1 mutant, the spectral changes are analogous to those observed in the 153A variant (Figure 6.13C). Quantitatively, the scaled mobility value increased from 0.53 to 0.58, which suggests that there is an increase in the mobility of the nitroxide at site 131 in the nanosecond time scale. The spectrum of 151R1 in the 99A/153A mutant exhibits a slight reduction in the spectral intensity corresponding to the parallel component of the hyperfine extrema (*see* || symbol in Figure 6.14). Quantitatively, the mutant showed a 6 % reduction on the scaled mobility value and 1 G increase in the hyperfine splitting value (*see* Appendix), both of which suggest a reduction in the overall mobility of the nitroxide side chain in the nanosecond time scale. As in the F153A variant, the reduction in the

mobility of the nitroxide at site 151R1 in the 99A/153A protein may reflect a local reduction on the nanosecond backbone motions of the J-helix due to the inward movement of the J-helix observed in the high-resolution structure of the double mutant.

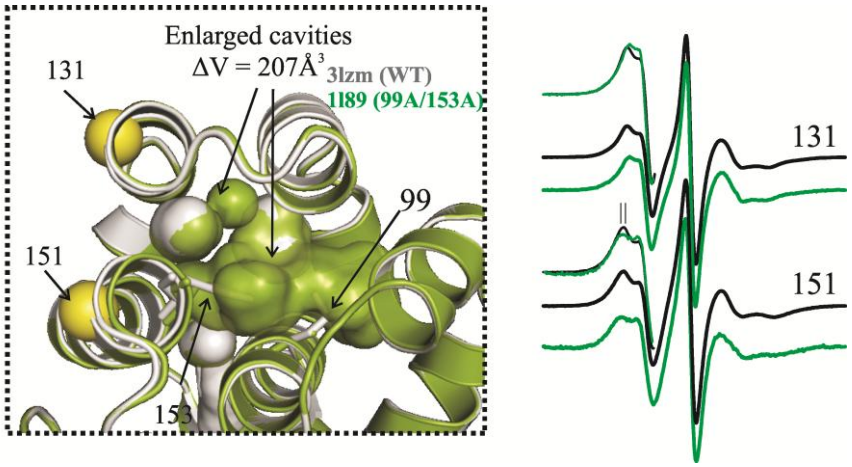


Figure 6.14: EPR studies on the L99A/F153A variant. *Left panel*: Overlay of the high-resolution structures of the WT (gray) and double mutant (green). The surfaces of the native and mutant cavities are shown. *Right panel*: EPR spectra of residues 131R1 and 151R1 in the WT (black) and mutant (green) backgrounds. The low field intensities of the four spectra have been amplified for clarity.

6.3.10 *L84A*. The L→A substitution at site 84 decreases the stability of the protein by 3.9 kcal/mol and is different from the other large-to-small substitutions discussed above as it created a tunnel with both ends open to the solvent as depicted in Figure 6.15 [24]. The crystal structure of this variant revealed significant structural relaxation in response to the mutation involving the F-helix, which moves about 2 Å and becomes less ordered than in the WT structure as judged by the changes in the electron density of the backbone [24].

To monitor the changes in the flexibility of the F-helix in *solution*, the EPR spectrum of residue 109R1 was recorded in the L84A background. As shown in Figure 6.15 the EPR spectrum showed dramatic changes in the mobility of the nitroxide side chain compared to that in the WT background. Qualitatively, in the low field it is clear that the spectral intensity of the α component is significantly reduced. To obtain a more quantitative picture into the changes in the

motion of the nitroxide for this variant, the EPR spectrum was fit to the MOMD model. The spectrum was fit reasonably well with a two-component model with the same population, but with a lower S for the alpha component compared to that obtained in the WT background indicating an increase in the amplitude of the nitroxide motion of 4° (Figure 6.15). The spectral changes observed upon arising from the L→A susbtitution can arise either from the 2 Å backbone motion, which may change the environment around the R1 side chain for site 109 or from changes in the backbone motions of the helix. Since none of the two components in the spectrum obtained in the WT or L84A variants reflect any tertiary interaction with the environment and the high-resolution structure suggests that the helix becomes more disordered in the L84A mutant, the latter possibility is likely. Osmotic-perturbation studies showed significant changes in the mobility of residue 109R1 in the L84A background, which is a hallmark of slow exchange (*see* Appendix).

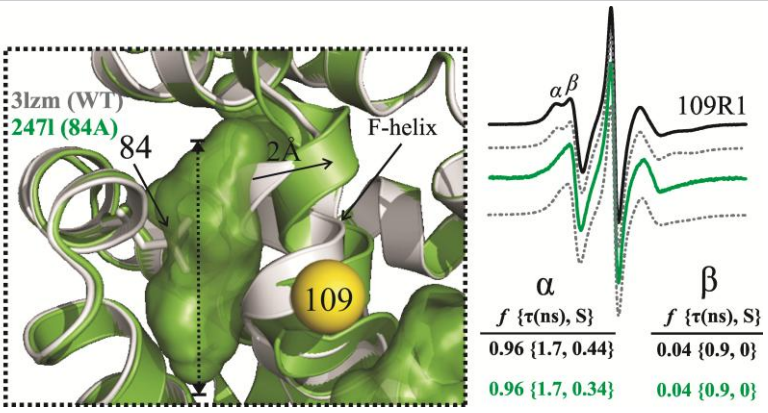


Figure 6.15: EPR studies on the L84A variant. *Left panel*: Overlay of the high-resolution structures of the WT (gray) and the L84A mutant (green). The external surface of the engineered tunnel is shown. The 2 Å movement of the F-helix is indicated. *Right panel*: EPR spectra of residue 109R1 in the WT (black) and mutant (green) backgrounds. The fits of the experimental data are shown in gray. The parameters obtained from the MOMD fits are shown.

6.4 Small-to-large substitutions

Many small-to-large substitutions have been engineered by Brian Matthews and co-workers [19,21,22,38] to evaluate the thermodynamic and structural effects of strain and cavity-filling substitutions on T4L. However, save for the A98V mutant [31], the effect of such mutations on the flexibility of the protein in solution has not been investigated. Thus in this study the local changes in protein flexibility due in some of those small-to-large variants designed by Matthews' group and additional newly designed variants designed for this dissertation (Figure 6.16) were evaluated by EPR spectroscopy and the results are shown below.

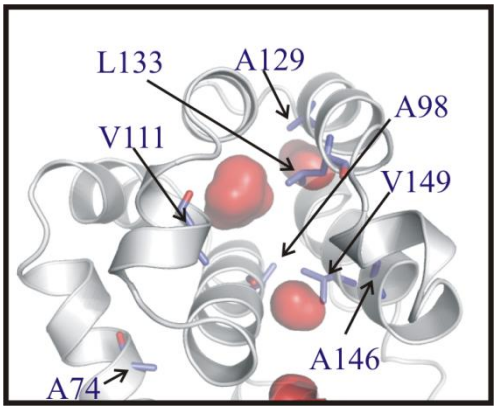


Figure 6.16: Sites selected for small-to-large substitutions. The exterior surfaces of the native cavities are shown in red.

6.4.1 *Cavity-filling mutants A129V, A129M, L133F*: The small-to-large substitutions on sites 129 and 133 were originally designed to improve the protein packing within the hydrophobic core of T4L since the natives A129 and L133 residues project towards the native cavities I and II, respectively as shown in Figure 6.16 [19,26]. Thermal stability studies showed a slight reduction in the stability of these three variants ($\Delta\Delta G \leq 1.9$ kcal/mol) compared to the WT protein and the crystal structures revealed that the larger side chains were accommodated with subtle structural change in the structure [19,26]. For the 129V and 133F variants, the substitutions fill the pre-existing cavities II and I, respectively but decrease the stability of the protein. The authors concluded that the increase in native packing, which was expected to increase the stability of the protein by 24 cal/mol per additional $Å^2$ buried in the core, was offset by strain energy introduced by the bulkier residue [19]. In the case of the A129M variant, the overall cavity volume was actually greater (24 $Å^3$) than that of the WT protein due to some

structural changes that resulted in a larger cavity I and created a small new cavity as shown in Figure 6.18A [26].

To study whether these substitutions affect the local fast backbone motions of the H-helix, which is the helix bearing the substitution, the EPR spectra of residue 131R1 were recorded in the 129V, 133F, and 129M backgrounds. As shown in Figure 6.17, the 131R1 spectrum in the 133F background shows spectral intensity in the region corresponding to a relatively immobilized state of the nitroxide, however the lineshape of the more mobile component is very similar to that in the WT background indicating no changes in the fast backbone motion. Osmotic-perturbation studies showed that the populations giving rise to each spectral component are in slow exchange (lifetime > 100ns; *see* Appendix). Conversely, the EPR spectra of residue 131R1 in the 129V and 129M background reveal subtle, but appreciable changes in the fast backbone motions of the helix as shown in Figure 6.18A. Qualitatively, the spectral intensity in the low field became sharper and slightly narrower indicating changes in the motion of the nitroxide side chain. The scaled mobility if both mutants increased by ~ 0.05 indicating increase in the motion of the nitroxide. To get a more quantitative understanding of the changes in the motion of the nitroxide for the 129M variant, the EPR spectrum was fit with a single component model as shown in Figure 6.18A. Spectral simulation for the A129M variant revealed a decrease in the correlation time from 2.2 to 1.6 ns and a slight increase in the order parameter from 0.40 to 0.45 compared to the spectrum in the WT background, which suggests a faster rate of motion of the H-helix, but of

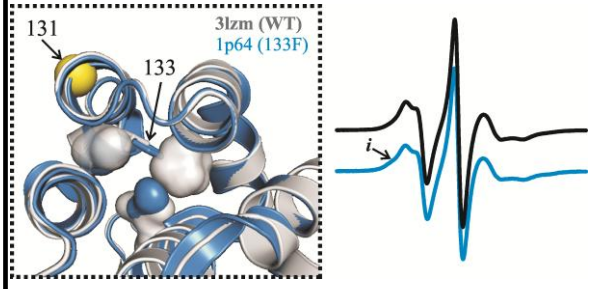


Figure 6.17: Left panel: Overlay of the WT (gray) and L133F mutant (blue) structures. Right panel: EPR spectra of residue 131R1 in the WT and 133F backgrounds are shown in black and blue, respectively.

narrower angular amplitude (2°). It should be noted that both the side chains of the valine and methionine residue at site 129 fill the pre-existing cavity II, which is located right below the H-helix, thus the reduction of the angular motion of the helix may be related to the increase in the local hydrophobic packing of the helix.

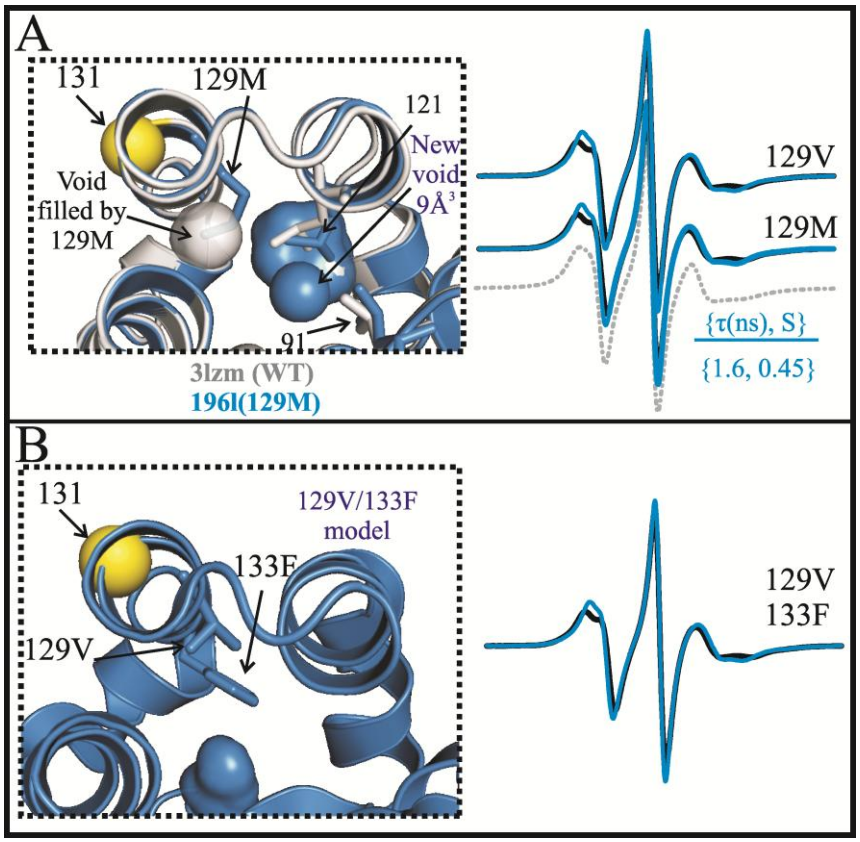


Figure 6.18: EPR studies of cavity-filling mutants. A) A129V and A129M. *Left panel*: Overlay of the WT (gray) and 129M (blue) structures. The surfaces of the internal cavities are color-coded. *Right panel*: EPR spectra of residue 131R1 in the WT, 129V, and 129M proteins are shown. The fit of the spectrum in the 129M variant is shown in gray. The parameters obtained from the simulations are shown in brackets. B) A129V/L133F. *Left panel*: Structural modeling of the double mutant shows no internal cavities. *Right panel*: EPR spectra of residue 131R1 in the WT and 129V/133F proteins are shown.

To further increase the hydrophobic packing below the H-helix, the 129V/133F double mutant was engineered. Although this mutant has not been previously characterized, modeling shows that the double mutation should fill the two pre-existing cavities present in the WT protein as depicted in Figure 6.18B. The thermal stability of the double mutant in the 131R1 background was compared to that of the 131R1 protein alone in order to determine the relative stability (*see* Appendix). The 129V/133F mutant slightly destabilized the protein by about 0.3 kcal/mol in the 131R1 background. Interestingly, the EPR spectrum of the double mutant is almost identical to that of the 129V mutant alone and, unlike the spectrum of the 133F variant, there is not spectral intensity in the region corresponding to an immobile state. Thus, the flexibility of the H-helix in the double mutant is dominated by the effects of filling the pre-existing cavity II.

6.4.2 *A98V and V149L*. Both of these mutants were selected for this study as they overpack the core near the J-helix. The structure and stability of the A98V variant was investigated in previous studies [38], which revealed that the mutation decreased the stability of the protein by 3.2 kcal/mol. The native A98 residue is in van der Waals contact with several residues of the J-helix, namely V149, T152, F153. The structural changes upon the substitution include a 0.6 Å outward movement of the J-helix and a bending angle for the same helix of 3.4° compared to the WT protein [38]. In addition to these structural and thermodynamic studies, Anderson et al. [31] measured the exchange rates for every residue in order to characterize the dynamic properties of this variant in solution. Interestingly, the exchange rates measurements revealed dramatic changes in the protection factors in the A98V mutant compared to the WT protein, particularly for those backbone amides in the C-terminal domain of the protein, which showed that the effect

of the mutation propagated far from the site of the mutation, but were limited to the domain bearing the mutation.

The V149L mutation is a new variant engineered for this dissertation because structural modeling shows that the leucine residue should overpack the region just below the J-helix. The EPR spectra of residue 151R1, reporting on the dynamics of the J-helix, were recorded for both mutants. As shown in Figure 6.19, the A98V mutation has no effect on the backbone motions of the J-helix as judged by the lack of any detectable spectral changes. Qualitatively, the scaled mobility and 2Azz' values for the A98V mutant were the same as those in the WT background protein. The EPR spectrum of 151R1 in the 149L variant showed little changes compared to that obtained in the WT background, which suggests that neither of these two mutations affects the local dynamics of the J-helix. It should be noted that the lack of spectral changes observed for the A98V mutant is not incompatible with the significant changes in exchange rates observed by NMR [31] because the time domain and dynamic processes that each technique probes are very different (*e.g.*, nanosecond to picosecond *versus* second or longer).

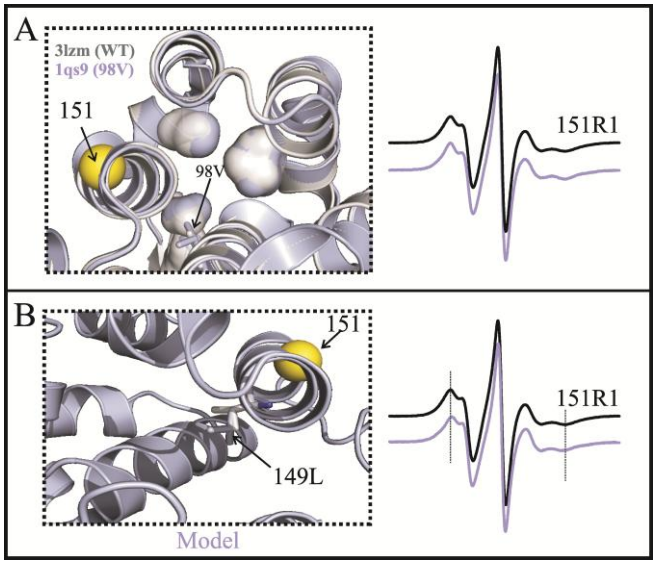


Figure 6.19: EPR studies in overpacking mutants of T4L. A) A98V. *Left panel*: Overlay of the WT (gray) and A98V (lightblue) structures. *Right panel*: EPR spectra of residue 151R1 in the WT, A98V proteins are shown. B) A129V/L133F. *Left panel*: Structural modeling of the V149L mutant. *Right panel*: EPR spectra of residue 151R1 in the WT and 149L proteins are shown. The dotted lines are used to help guide the eye.

6.4.3 *V111I*. The native valine at position 111 is one of the two residues of the short F-helix that packs against the hydrophobic core of the protein. In the original studies the V→I

substitution was designed to alter the packing of the hydrophobic core of the protein [21]. The V111I variant was slightly more destabilized relative to the WT protein ($\Delta\Delta G$ = 0.85 kcal/mol), however the x-ray structure revealed a substantial change in the conformation of the F-helix, which involved a 1.5 Å outward movement of the helix away from the core as shown in Figure 6.20 and a loss of *intra*helical hydrogen bond between the carbonyl group of residue 109 and the amide of the residue 113 [21]. To evaluate the changes in flexibility of the F-helix in solution, the EPR spectrum of residue 109R1 in the V111I background was recorded. As shown in Figure 6.V111I, there is a dramatic change in the spectrum obtained in the 111I mutant compared to that in the WT background. Qualitatively, the spectral intensities in the lower field corresponding to the α and β components in the WT background merge into a single peak. The scaled mobility value increases from 0.64 to 0.72 indicating faster motion of the nitroxide side chain in the V111I variant. The spectrum was fit reasonably well with a two-component model with the same population, but with higher rate and lower order for the α component compared to that in the WT background. Specifically, the rate of motion of the α component increases from 0.59 GHz to 0.67 GHz, which is in agreement with the increased in the scaled mobility value. It should be noted that the V111I mutant decreased the size of the native cavity I (note changes in the surface of the cavity I shown in Figure 6.20), which should be stabilizing, however the loss of the intrahelical hydrogen bond apparently played a major role in destabilizing the protein and it is likely responsible for the changes in the backbone motions of the helix detected by EPR[6].

[6] The changes in the EPR spectra in the WT and 111V proteins are unlikely to arise due to the outward movement of the F-helix since the EPR spectrum of 109R1 in the G113A mutant, in which the helix also moves (Nicholson et al. [39]), showed no spectral changes (data not shown).

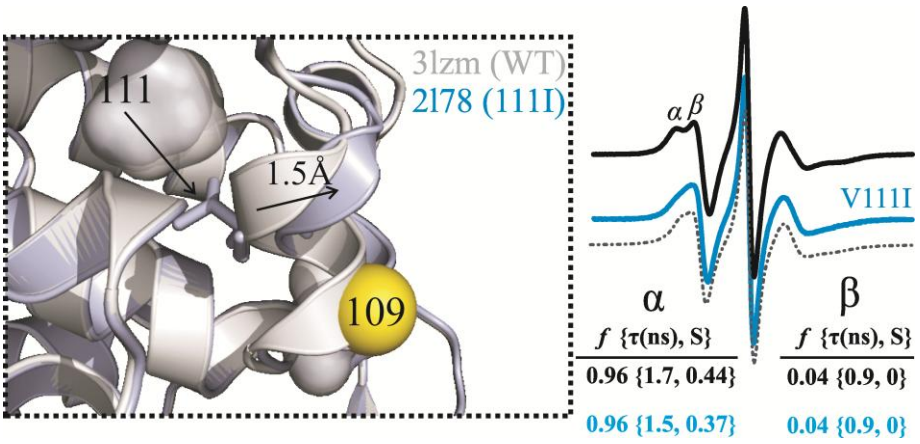


Figure 6.20: Changes in backbone flexibility of the F-helix due to the V111I mutation. *Left panel*: Overlay of the structures of T4L in the WT (gray) and V111I (lightblue) proteins. The outward movement of the F-helix is highlighted. *Right panel*: EPR spectra of residue 109R1 in the WT (black) and V111I backgrounds (lightblue). The fit of the spectrum in the 111I mutant is shown in gray dashed lines. The parameters obtained from the MOMD simulations are shown.

6.4.4 *A74M*. The native alanine at position 74 is located in the interdomain helix C, has low solvent accessibility (3%), and packs against hydrophobic residues I100, V103, and F104 of the E-helix. Although the structural effects of this substitution has not been previously characterized, structural modeling shows that any larger side chain at position 74 would overpack this hydrophobic region and possibly affect the dynamics of the C-helix. As mentioned above, previous studies have shown that the interdomain C-helix is well ordered and has relatively damped fast backbone fluctuations [31,36]. Thus in order to evaluate the effect of overpacking mutations on the overall stability of the protein and flexibility of the C-helix, an alanine-to-methionine substitution was engineered at site 74 and the R1 side chain at site 72 was used as a sensor of changes in local backbone motions.

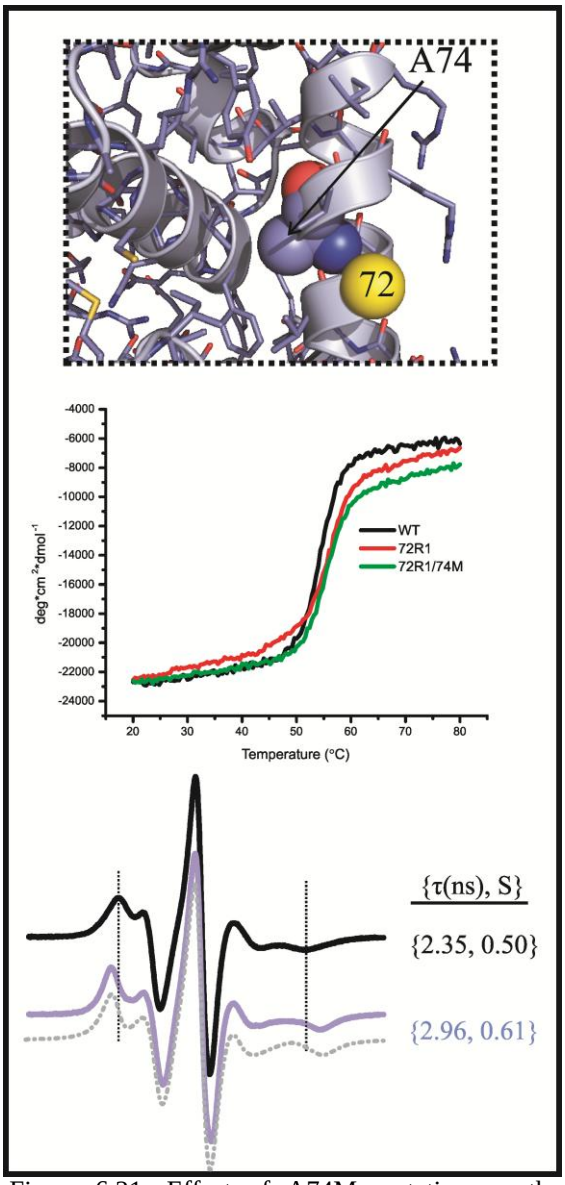


Figure 6.21: Effect of A74M mutation on the stability and flexibility of T4L. *Top panel*: Ribbon diagram showing the site of the substitution and the sites where R1 was introduced (sphere at Cα). *Middle panel*: Thermal denaturation studies. *Bottom panel*: EPR spectra of residue 72R1 in the WT (black) and A74M (light blue) backgrounds. The best fit for the A74M spectrum is shown.

Thermal denaturation studies were done in the WT, 72R1, and 72R1/74M variants to elucidate the changes in stability due to the A74M mutation (Figure 6.21). The thermal stability of the D72R1 is similar to that published previously [40]. The D72R1 mutation increased the stability of the protein relative to the WT protein by 0.42 kcal/mol. Remarkably, the 72R1/74M variant is more stable than the WT protein and has similar stability compared to the 72R1 variant, suggesting that the overpacking mutation does not affect the overall stability of the protein. The EPR spectrum of residue 72R1 in the 74M mutant, reporting on the local dynamics of the C-helix, revealed significant differences compared to the spectrum obtained in the WT background. Although both spectra exhibit anisotropic lineshape, the spectrum of the overpacking mutant is broader and the spectral intensity corresponding to the parallel component of the hyperfine extrema is sharper and moves outwards indicating slower mobility (*see* dotted lines in Figure 6.21). Quantitatively, the scaled mobility value decreases from 0.37 to 0.29 and the hyperfine splitting values increases from 52.6 to 58.5 G, which suggest a decrease in the rate and increase in the order of motion of the

nitroxide side chain. Indeed, MOMD simulation of both spectra revealed an increase in the correlation time ($rate^{-1}$) and order parameter of 72R1 in the A74M variant. Specifically, the rate of motion decreased from 0.42 to 0.34 GHz and the order parameter increased from 0.50 to 0.61 corresponding to a decrease in the angular amplitude of the motion of the nitroxide side chain of ~ 5º. The origins for the reduction in the mobility of the nitroxide in the A74M mutant are not entirely understood, however previous studies have shown that the EPR spectra of sites exhibiting strongly ordered z-axis anisotropic motion of the nitroxide (S > 0.50), such as the 72R1/74M variant, generally reflect tertiary interaction with the environment [36,42-44]. Although there is no structural information of this variant, the dramatic increase in the order of motion of the nitroxide could arise from local structural rearrangement of the protein that brings the nitroxide ring into tertiary contact with nearby side chains. Since the nitroxide side chain at site 72 in the WT background is solvent exposed as judged by the spectral lineshape, the structural relaxation giving rise to slower motion of the nitroxide might involve segmental rotation of the C-terminal region of the C-helix, which would bring the nitroxide closer to the protein environment. Alternatively, the slower motion of the nitroxide at site 72R1 in the A74M mutant might be a reflection of further reduction on the nanosecond backbone fluctuations of the helix C. Future studies involving interspin distance measurements will be done to distinguish these two possibilities.

6.4.5 *A146V*. As shown above in Figure 6.12, the short helix I is very sensitive to changes in the local packing. To elucidate the effect of overpacking the hydrophobic core near this short helix, a larger side chain at position 146 was engineered. As depicted in Figure 6.22, the native alanine at position 146 packs against residues W138 and Y139 of helix I and structural modeling shows that any larger side chain can not be accommodated without structural changes due to

steric clash. To determine the effect of the A146V mutation on the overall stability of the protein, thermal denaturation studies were done on the 140R1/146V mutant and compared to that obtained in the 140R1 protein. The thermal denaturation studies revealed that the A146V mutation decreased the stability of the protein by 1.68 kcal/mol relative to the 140R1 background (*see* Appendix). Assuming that the observed change in $\Delta\Delta G$ has additive contributions from the difference in transfer free energy for the A→V substitution[7] ($\Delta\Delta G_{tr}$) and a reorganization term arising from strain energy[8] ($\Delta\Delta G_{reorg}$), the strain energy from overpacking can be estimated. Using the experimental $\Delta\Delta G$ value obtained for the 140R1/146V mutant (- 1.68 kcal/mol), the inferred strain energy is estimated to be about 2.9 kcal/mol, which is partially compensated by the favorable hydrophobicity term from the A→V mutation.

The EPR spectrum of residue 140R1 in the 146V mutant was recorded and compared to that obtained in the WT background. The spectrum of the 140R1/146V mutant showed spectral intensity in the region corresponding to a relatively immobile state of the nitroxide side chain not observed in the spectrum of the 140R1 protein at 295K. At 295K, the spectrum of residue 140R1 can be simulated with a single component having z-axis anisotropic motion with τ = 1.2 ns and S = 0.41 as shown in Figure 6.22C, suggesting no tertiary interaction with the environment. However, temperature dependent EPR studies shows that at temperatures below 293K, a relatively immobile component is resolved (see arrows in Figure 6.22D) indicating that there is a second state of the nitroxide side chain or the protein that places the nitroxide in tertiary contact with other groups of the protein.

[7] $\Delta\Delta G$ and $\Delta\Delta G_{tr}$ are expressed as mutant minus WT. $\Delta\Delta G_{tr}$ for the A→V substitution is + 1.2 kcal/mol (Xu et al. [24]).

[8] Strain energy in this case refer to both backbone and side chain strain.

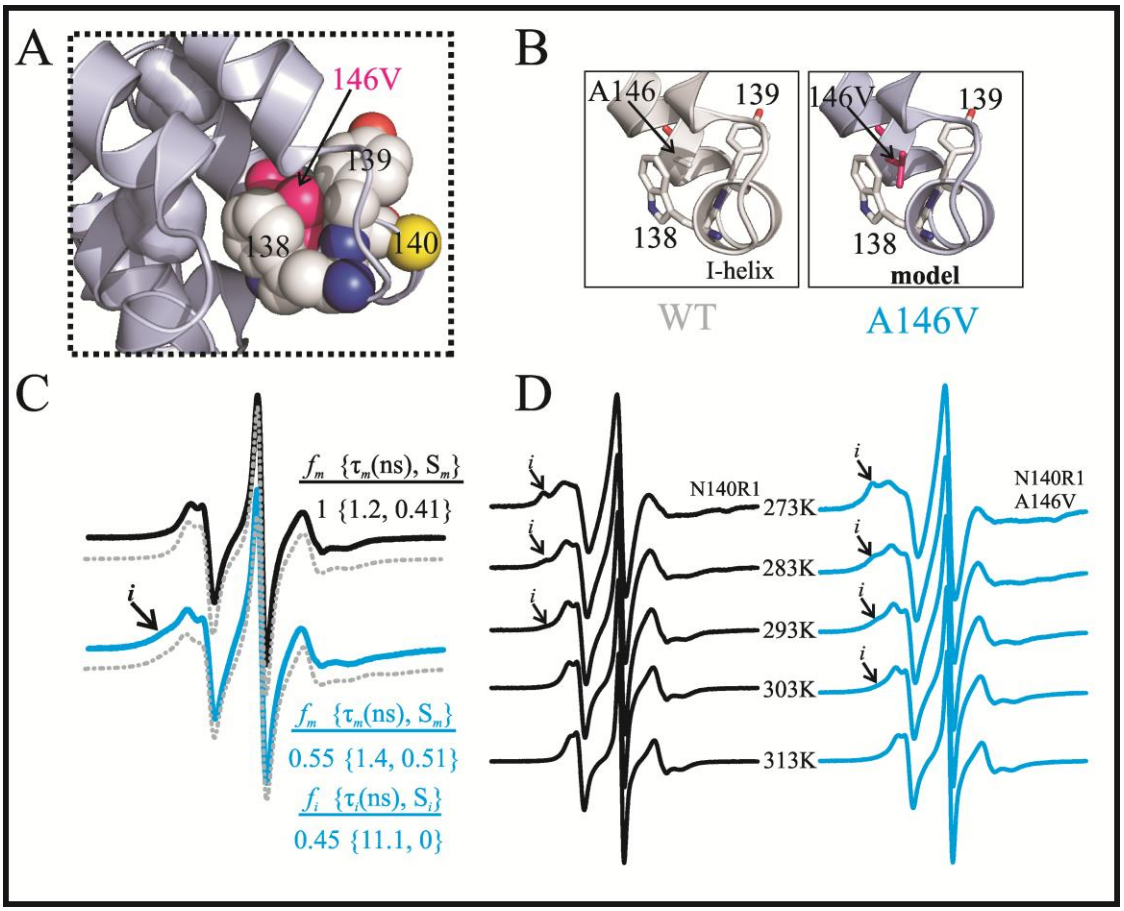


Figure 6.22: EPR studies on the overpacking A146V mutant. A) Model of T4L A146V protein showing the steric clash of the engineered 146V side chain with nearby residues. The yellow sphere at the Cα show the site where R1 was introduced. B) The alternative hydrophobic packing of helix I is shown. C) EPR spectra of residue 140R1 in the WT (black) and A146V (light blue) proteins. MOMD fits of the corresponding spectra are shown in gray. D) Temperature dependent EPR studies. The relatively immobile states of the nitroxide are indicated.

However, the interaction must be sufficiently weak that thermal energy effectively competes with it, thus reducing the population of the immobile component to levels below detection (ca. < 5%) at 295K. Temperature-dependent studies of 140R1/146V mutant also shows an increase in the population of the relatively immobile component at low temperature, however the interaction(s) giving rise to the immobile state must be stronger that than in the WT background as the population remain in detectable levels up to 303 K. It should be noted that the population of the immobile state at low temperatures in the 140R1/146V protein is much higher than that observed in the 140R1 variant as judged by the relative spectral intensities. The EPR

spectrum of the 140R1/146V mutant at 295K can be satisfactorily simulated with 45 % of an immobile component (τ = 11.1 ns) and a weakly ordered component reflecting slower nanosecond motions (τ = 1.4 ns and S = 0.50) compared to the weakly ordered component in the 140R1 spectrum. The reduced mobility of the more mobile state of the 140R1/146V mutant suggests that overpacking the region adjacent to the helix-I reduced the fast backbone motion of the helix likely due to steric constrains of its hydrophobic surface. However, based on the EPR spectral changes and the temperature-dependence studies, it is not clear whether the immobile component observed in the 146V mutant is a *new* state introduced by the mutation or is the same state observed for 140R1 at low temperatures, but at a higher population. Additional experiments involving DEER spectroscopy will be required to distinguish these two possibilities.

6.5 Compensating mutations

6.5.1 *Cavity compensating L121A/A129M.* As mentioned above, the individual L121A and A129M substitutions destabilized the T4L protein relative to WT protein [16,22] with the former destabilizing the protein due to the energetic cost of increasing the size of an existing cavity (24 cal mol^{-1} $Å^{-3}$) and the latter due to introduced torsional strain. Baldwin et al. [22] realized that since residues L121 and A129 are adjacent to one another, the combined variant 121A/129M (denoted "size-switch" variant) could possibly compensate for the destabilizing effect of the individual mutations (Figure 6.23B). Indeed, thermodynamic studies showed thermodynamic compensation of the double mutant compared to the single variants. Interestingly the compensation was not complete since the "size-switch" mutant was less stable than the WT protein by 1 kcal/mol [22]. To elucidate the structural changes of the 121A/129M mutant relative to the individual mutants and the WT protein, the x-ray structure of the double mutant was determined at a resolution 0f 1.85 Å [22]. The structure of the 121A/129M variant showed

structural changes analogous to those observed in the 121A variant and include a ~ 1 Å inward movement of the C-terminal region of the J-helix and a side chain shift of residue F153 as depicted in Figure 6.23A. The overall size of the internal cavities was similar to WT, but distributed differently as the pre-existing cavity II disappeared and the cavity I increase in size. Based on the combined structural and thermodynamic studies done on this variant and in others "size-switch" variant in T4L, the authors concluded that the packing of the hydrophobic core of T4L is finely tuned, thus changes in the hydrophobic packing can be partially compensated only to a limited extent by a mutation of adjacent residue(s) [22].

To evaluate whether there is a compensation in the local flexibility of the "size-switch" variant relative to the individual mutation, the EPR spectrum of residue 131R1, reporting on the flexibility of the H-helix, was recorded in the 121A/129M variant and compared to the spectra obtained in the WT, 121A, and 129M backgrounds. As discussed above, the individual 121A and 129M substitutions exert different effects on the lineshape of residue 131R1, for instance, the spectrum of the 121A mutant revealed a new immobile state not observed in the spectrum obtained in the WT background, while the spectrum in the 129M variant reflects faster nanosecond motion with a more restricted amplitude. As shown in Figure 6.23C, the spectrum of 131R1 in the 121A/129M protein showed only subtle differences compared to the spectrum obtained in the WT background (*see* lower field peaks). Comparison of the spectral lineshape of the 4 variants revealed that the "size-switch" variant partially compensates the effects of the 121A and 129M individual mutations (Figure 6.23D). Specifically the "size-switch" variant does not show the relatively immobile state observed in the 121A mutant and the spectral intensity in the lower field is reduced relative to the 129M variant and is more similar to the WT (*see* amplified low field peak in Figure 6.23D). In addition, the 2Azz' values, which are smaller for

the 129M variant, moved closer to the values obtained in the WT background in the double mutant (*see* Appendix). These results indicate that in addition to the partial thermodynamic and structural compensation of the 121A/129M mutant, there is also a partial compensation in the local flexibility of the H-helix by this double mutant.

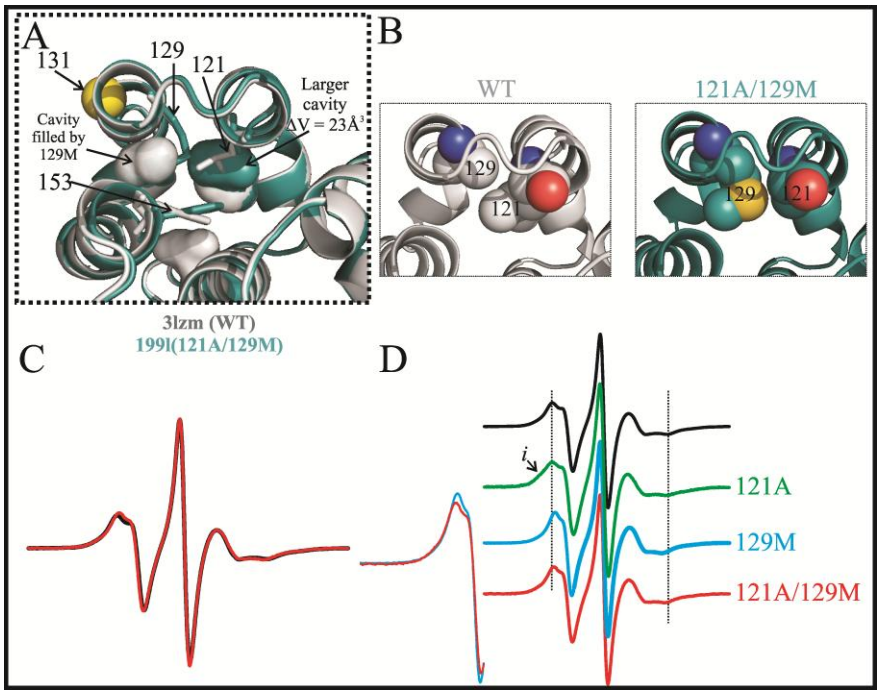


Figure 6.23: EPR studies on the 121A/129M variant of T4L. A) Overlay of the high-resolution structures of the WT (gray) and 121A/129M mutant (cyan). B) Ribbon diagram of each structure showing a CPK representation of the mutated residues. C) EPR spectra of 131R1 in the WT and 121A/129M backgrounds are shown in black and red, respectively. D) EPR spectra of 131R1 in the WT, 121A, 129M, and 121A/129M proteins. The dashed lines indicate the hyperfine splitting. The low field peaks of the spectra in the 129M and 121A/129M mutants were amplified for clarity.

6.5.2 *Buried polar compensating 102K/133D substitution*: The myriad of structures of soluble proteins have shown that most charged and polar residues are usually localized on the surface of the protein, while nonpolar residues are distributed towards the protein interior forming "waxy" cores. This empirical observation has been rationalized in terms of evolutionary pressure to enhance the stability of the protein by keeping ionizable groups away from the interior due to the high energetic cost (unfavorable ΔG_{tr}) of transferring a charge group from a

high dielectric environment (water: ε ~ 80) to the relatively low dielectric environment in the interior of proteins (ε ~ 2-8) [45]. However, for many proteins ionizable buried residues play essential roles in protein function (*e.g.*, buried residues E11 in T4L and D25 in HIV-1 protease) [8,46]. Thus it has been proposed that for such cases the microenvironment around the polar amino acid must have evolved to mitigate the energetic cost of burying a functionally important ionizable amino acid [47]. One of the implications of this assumption is that burial of an isolated charge in a completely hydrophobic environment should be energetically costly and deleterious to the structure and functionality of proteins.

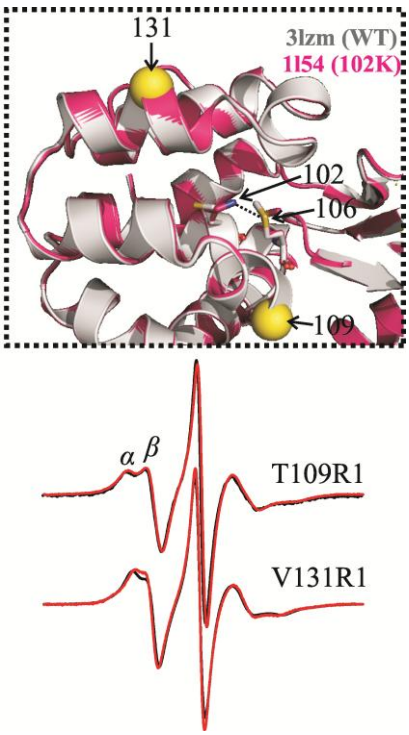


Figure 6.24: EPR studies on the buried charge mutant M102K. Overlay of the high-resolution structures of the WT and M102K proteins [20]. The hydrogen bond between residues 102 and 106 in the 102K protein is shown. EPR spectra of 109R1 and 131R1 in the WT (black) and mutant (red) proteins are shown.

To elucidate the structural and energetic cost of engineering a charged group in the interior of a protein, Dao-pin and coworkers [20] introduced two buried charged mutants (M102K and L133D) in the interior of T4L. Both mutations were created in sites with low solvent accessibilities (SASA < 1%) and located in the hydrophobic core of the C-terminal domain of the protein. The engineered charged residues have similar size and shape as the native hydrophobic residues (isosteres), thereby eliminating the possible effects of introducing new cavities and strain into the structure [20]. Although both mutations were highly destabilizing ($\Delta\Delta G \geq 5.7$ kcal/mol near neutral pH), remarkably both variants folded at neutral pH [20]. The M102K protein was the more unstable of the two variants, however it exhibited 35% of enzymatic

activity compared to the WT protein, while the L133D mutant showed little activity (4%). The high-resolution structure of the M102K protein showed that the engineered lysine was indeed buried inside the core of the protein and formed a (N-H•••S) hydrogen bond with M106 (Figure 6.24). Surprisingly there was little structural changes in response to the M102K mutation compared to the WT protein (RMSD = 0.14 Å), however the B-factor values for residues 105-114 were significantly higher than those obtained in the WT structure suggesting higher mobility for this region [20]. Apparently, the L133D variant did not crystallize, thus the structural properties of this variant could not be elucidated.

To elucidate the changes in flexibility in solution of some of the core helices due to these buried charge substutions, the EPR spectra of residues 109R1 and 131R1, reporting on the flexibility of the F- and H- helices, respectively were recorded in the M102K variant (Figure 6.24). It is intriguing that although the high-resolution structure of the M102K protein showed significantly higher B-factor values ($\Delta > 40$ Å^2) for the F-helix, the EPR spectrum of 109R1 was the same as that obtained in the WT protein. Interestingly, the EPR spectrum of residue 131R1 showed subtle changes in the mobility of the nitroxide in the 102K variant compared to that obtained in the WT background as judged by the spectral changes in the low field. Quantitatively, the scaled mobility value for 131R1 increases from 0.54 to 0.61, which indicates higher mobility of the nitroxide side chain in the 102K protein. It should be noted the changes observed for 131R1 in the M102K variant likely reflect subtle changes in the fast backbone motion of the H-helix. The spectrum of residues 131R1 was also recorded in the L133D variant to study the local

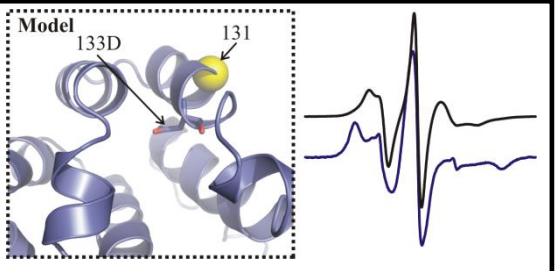


Figure 6.25: EPR studies on the buried charge mutant L133D. Model of the L133D mutant. EPR spectra of 131R1 in the WT (black) and mutant (blue) proteins are shown.

changes in backbone flexibility (Figure 6.25). Interestingly, the spectrum of residue 131R1 in the L133D protein showed dramatic spectral changes compared to the spectra obtained in the WT background (Figure 6.25). Specifically, the spectrum reflects two (or more) spectral components with a dominant highly immobilized component reflecting extensive tertiary interaction of the nitroxide side chain with the protein environment. Osmotic-perturbation studies showed that the populations giving rise to each spectral component are in slow exchange (lifetime > 100 ns; *see* Appendix). The data suggest that the 133D mutation introduced conformational flexibility in the μs-ms time scale and likely cause local structural changes at or in the vicinity of H-helix, whereby the nitroxide side chain at site 131 becomes buried or partially buried. At this point it is not entirely understood the type of the structural changes of this variant (*i.e.*, local unfolding or axial rotation of the H-helix) or the scope of it (*i.e.*, whether it extends to other helices of the C-terminal domain), thus future CW studies involving additional sensor sites in combination with DEER spectroscopy studies will undoubtedly shed light onto the type and scope of structural changes due to the L133D substitution.

In order to elucidate whether the effects observed in the L133D variant can be compensated with a buried group of opposite charge, the double mutant M102K/L133D was engineered. As mentioned above, the individual variants have the similar shape and size as the native residue they replaced and structural modeling suggests that both ionizable groups can form hydrogen bond and/or salt bridge to the close proximity to each other[9] as depicted in Figure 6.26. Previous studies showed that the thermal stability of this double mutant was much lower than the WT protein [31] indicating little thermodynamic compensation. Thermal denaturation

[9] In the model, the L133D was placed at the same coordinates as the native leucine, while the M102K side chain was modeled as in the 102K structure except for the dihedral angle between Cδ and Cε, which was changed from 29° to 150° (as defined by Lovell et al. [48]) to bring it in close proximity (3.2 Å) to the L133D side chain.

studies in the pH range of 3 to 6.5 done in this dissertation reveal that the double mutant was less stable than the WT background protein in the pH range studied (*see* Appendix). The lack of thermodynamic compensation of the M102K/L133D mutant in T4 lysozyme is in agreement with computational [49] and experimental studies [50] on other proteins, which showed that hydrophobic interactions (denoted "hydrophonic bridges" by Hendsch and Tidor [49]) of amino acids with the same size and shape as ionic amino acids (isosteric replacements) provide more stabilizing energy than buried salt bridges due to the high energetic cost of dessolvation of charged residues, which usually is not fully compensated by the favorable bridge and protein energy terms. In fact, Waldbruger et al. [50] found that the stability of the Arc repressor of bacteriophage P22 could be increased by replacing a native buried salt-bridge triad (R31, E36, R40) with several hydrophobic triads using combinatorial mutagenesis. Remarkably, some of the newly designed mutants of the Arc repressor were not only more stable than the WT protein, but also retained full activity. Notwithstanding the improvement in protein stability of "hydrophobic bridges" compared to buried polar pairs, it has been proposed that buried polar interactions may be important for providing specificity and for modulating conformational flexibility [49-51].

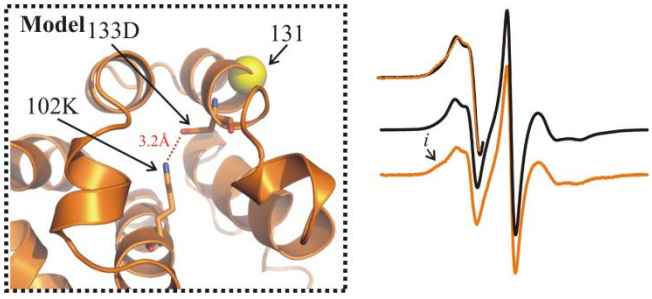


Figure 6.26: M102K partially compensates the L133D mutation. A model of the 102K/133D mutant is shown in the left panel. EPR spectra of 131R1 in the WT (black) and the double mutant (orange) proteins are shown. The relatively immobile component observed in the 102K/133D mutant is shown.

Despite the absence of thermal compensation in the 102K/133D ionic pair, remarkably, the EPR spectrum of residue 131R1 in the double mutant (102K/133D) revealed a partial compensation in the local structure and flexibility of the H-helix due to the 102K mutation (Figure 6.26). Specifically, the EPR spectrum reflects two dynamic components, but unlike the

spectrum in the 133D background, the more mobile component is dominant (as judged by spectral intensities) and is qualitatively similar to the mobile component observed in the WT background. However a relatively immobile component is still observed in the 102K/133D variant, which is not observed in the spectra obtained in the WT and 102K background proteins, suggesting that there is residual μs–ms flexibility from the L133D mutation. It is possible that the compensating M102K mutation shifts the equilibrium between the conformations giving rise to the immobile (new) and mobile components (native-like) of the nitroxide side chain towards the mobile conformer. Precedents for buried electrostatic interactions modulating conformational changes have been previously identified in rhodopsin by EPR spectroscopy [51]. Alternatively, it is possible that the conformation giving rise to the immobile component observed in the M102K/L133D variant is different to that observed in the L133D variant. Additional studies will be needed to distinguish these possibilities. At this point it is not clear whether at neutral pH, the compensation arises due to ionic interaction (salt-bridge) or hydrogen-bond formation between the polar residues, since it is not known whether the ionizable groups are charged at the experimental conditions (pH 6.8). Regardless of the nature of the putative interaction and origin of the immobile component in each case, it is clear that although there is little thermodynamic compensation from the second buried ionizable substitution, there is certainly a dynamic compensation detected by EPR spectroscopy.

6.6 Structural changes due to the L121A/L133A and W138A mutations studied by DEER spectroscopy.

As shown above, the large-to-small variants L121A/L133A and W138A showed the most striking effects on the local flexibility of the protein as judged by the dramatic spectral changes observed compared to the WT background. In both cases, the spectra in the WT background at

295K reflect a single dynamic component of the nitroxide side chain (*see* Figures 6.10D and 6.12B). However in the cavity-creating variants a relatively immobilized state of the nitroxide for residues 131R1 and 140R1 became the dominant spectral component suggesting that local structural rearrangements of the H and I helices changed the environment around the nitroxide side chain for the 121A/133A and W138A mutants, respectively. It should be noted the x-ray structure of the 121A/133A mutant does not give insights into the structural origins of the immobile component observed in the EPR spectrum of residue 131R1, while structural information on the W138A mutant is not available. To get additional insights into the origins of the more immobile component observed in the EPR spectra of residues 131R1 and 140R1 in the 121A/133A and W138A backgrounds, respectively, double electron-electron resonance (DEER) spectroscopy was employed for distance mapping in the WT and mutant protein.

6.6.1 *Interspin distance measurements in the WT and L121A/L133A background proteins*: To measure the distances between spin labels in T4L, three pairs of R1 mutants were constructed in the WT and mutants backgrounds; namely N68R1/V131R1, D89R1/V131R1, T109R1/V131R1. The interspin distance distributions of the doubly labeled mutants N68R1/V131R1 and T109R1/V131R1 of T4L in the WT background was published previously [40]. The doubly labeled mutant D89R1/V131R1 in the WT background, reporting on the distance between helices E and H, was constructed and studied by Dr. Mark R. Fleissner (unpublished results) and the original data of this mutant and the N68R1/V131R1 and T109R1/V131R1 variants in the WT background (gray traces in Figure 6.27) were kindly provided by Dr. Mark R. Fleissner and Dr. Wayne L. Hubbell for comparison purposes.

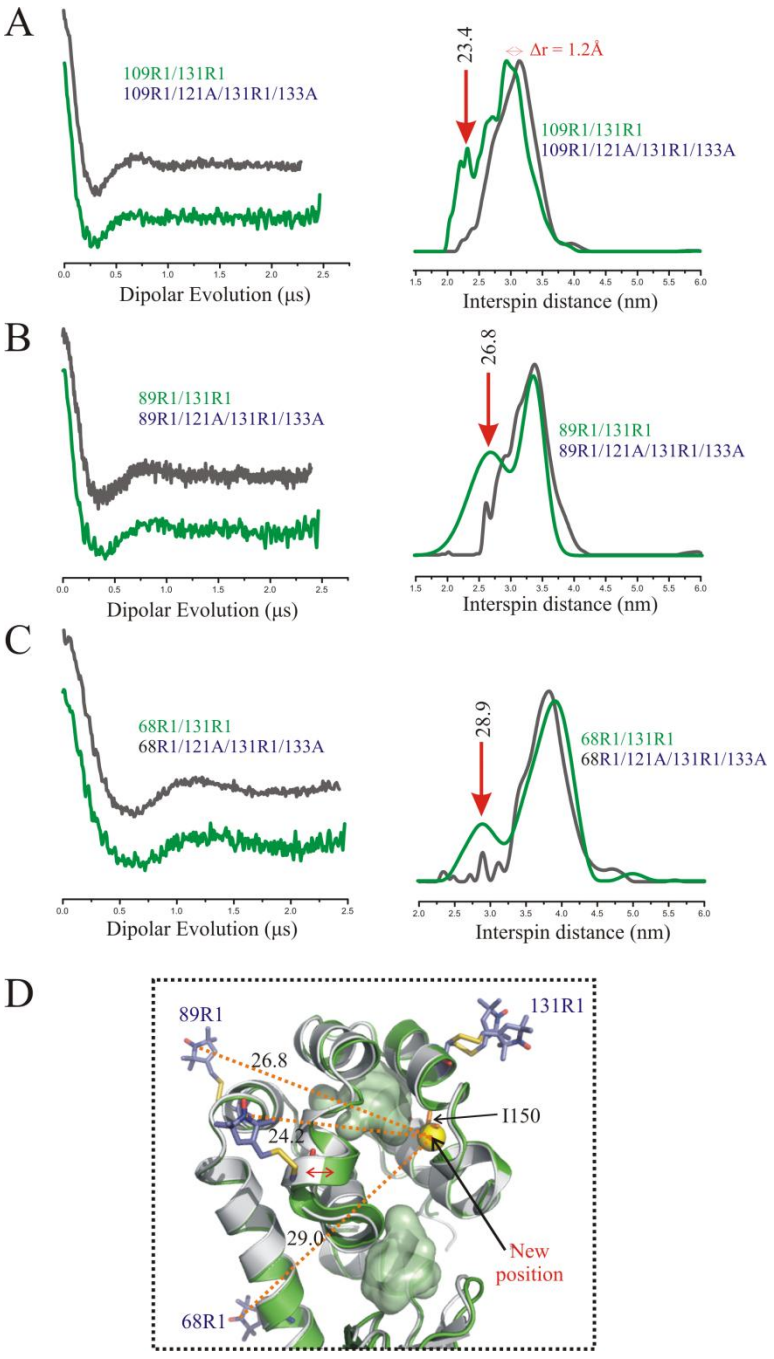


Figure 6.27: DEER studies done in the L121A/L133A mutant. A) Background substracted dipolar evolution and corresponding distance distribution for 109R1/131R1 mutant in the WT (dark gray) and 121A/133A (green) backgrounds. B) Background substracted dipolar evolution and corresponding distance distribution for 89R1/131R1 mutant in the WT (dark gray) and 121A/133A (green) backgrounds. C) Background substracted dipolar evolution and corresponding distance distribution for 68R1/131R1 mutant in the WT (dark gray) and 121A/133A (green) backgrounds. D) Ribbon diagram of T4L showing the putative spatial location of residue 131R1 in the 121A/133A background.

It should be emphasized that the interspin distance distributions reflect spatial distribution in the position of the nitroxide and not uncertainty in the distance measurements. As shown in panels A-C in Figure 6.27, in all cases the interspin distance distributions in the 121A/133A background are multimodal. Remarkably the most probable interspin distances between residues 68R1-131R1, 89R1-131R1, and 109R1-131R1 in the 121A/133A mutants are similar to those obtained in the WT background, although a small decrease in the most probable distance (Δr = 1.2 Å) is observed for the T109R1-V131R1 sample. The shifts in the most probable distance for residues 109R1-131R1 in the 121A/133A likely arise due to the structural relaxation observed in the x-ray structure of the 121A/133A mutant [24], which involved a ~ 1 Å inward movement of the main chain atoms of the F-helix towards the interior as depicted in Figure 6.27D. As discussed above, the CW spectrum of residue 109R1 in the 121A/133A was indistinguishable to that obtained in the WT background suggesting a similar environment for the nitroxide side chain at site 109 in both mutants. Interestingly, the most significant changes in the interspin distance of the doubly labeled mutants in the WT and 121A/133A backgrounds are the distance distributions. Specifically, shorter interspin distances ($\Delta r \geq 8$ Å) of significant intensities were observed for each of the three doubly labeled mutants in the 121A/133A background (*see* red arrows in Figure 6.27A-C), suggesting a change in the spatial distribution of the nitroxide. It is tentatively assumed that the observed changes in the spatial distribution of the nitroxide arise mainly from residue 131R1 due to the CW spectral changes observed at that site in the 121A/133A mutant and the fact that residues N68R1, D89R1, and T109R1 are relatively far from the sites of the substitution. Remarkably, comparison of the distance distribution of the mutants in the WT and 121A/133A backgrounds, reveal that the shorter distances are also present in the WT background proteins, but at a significantly lower intensity raising the interesting possibility

that the 121A/133A mutation did not create a *new* state, but it rather *shifts a pre-existing equilibrium* towards a second conformation, which may be exist at a very low population in the WT background.

To map the position of the nitroxide at site 131R1 from the peak corresponding to the **second** most probable distance in each doubly labeled mutant, the nitroxide side chain for sites 68R1 and 109R1 were modeled using the empirically determined spatial probability distribution for the nitroxide at these two sites [40]. At site 89R1, the nitroxide side chain was modeled using the canonical $\chi 1$, $\chi 2$, and $\chi 3$ angles observed for R1 in crystal structure of spin-labeled proteins [44], while $\chi 4$ and $\chi 5$ dihedral angles were optimized to avoid steric clashes and be in agreement with the *most probable* interspin distance with the nitroxide at site 131R1[10] in the WT background.

Trilateration using the second most probable interpsin distance in the distance distributions of the three pairs revealed the alternative location of the nitroxide at site 131R1 giving rise to the second most probable distances (depicted as a yellow sphere in Figure 6.27D). Interestingly, the new putative location placed the nitroxide at site 131R1 in a partially buried environment with extensive tertiary contacts with several residues of helices I and J; in fact the sphere, representing the spatial location of the nitroxide at site 131R1, is occupying the same position as the side chain of residue I150 suggesting that there must be some structural adjustments of the J helix, perhaps involving axial rotation, that allows the nitroxide at site 131 to occupy the same position usually occupied by residue I150 of the J-helix. It is should be emphasized that simple rotameric shift of the nitroxide side chain at site 131R1 without any structural adjustments of helices H and J can not account for the alternative location of residue 131R1 in the 121A/133A variant. As

[10] The crystal structure of the spin labeled protein V131R1 (Fleisner et al. [44]) was used for the positions of the nitroxide side chain at site 131R1.

shown above, the spectrum of residue 131R1 in the 121A/133A protein reflects a highly immobilized component indicative of extensive tertiary interaction of the nitroxide side chain with the protein environment that likely arises from this alternative location inferred from the distance distributions. Due to the limited amount of data, alternative modes of motion that may account for spectral changes and distance distribution observed for 131R1 in the 121A/133A mutant cannot be ruled out (see Discussion).

6.6.2 *Interspin distance measurements in the WT and W138A background proteins*. To measure the structural changes in the I-helix in the W138A mutant two pairs of R1 mutants were constructed in the WT and mutants background, namely N68R1/N140R1 and D89R1/N140R1. The N68R1/N140R1 mutant in the WT background was originally designed and studied by Dr. Mark R. Fleissner (unpublished) and the data was kindly provided for comparison purposes. As shown in Figure 6.28, there are significant differences in the intesrpin distance distribution for both doubly labeled mutants in the W138A variant compared to those obtained in the WT background. For example, the most probable distance for the 68R1/140R1 labels in the WT background is 35 Å, while the same in the W138A background is 29 Å. In addition, the distance distribution for the 68R1/140R1 in the WT background protein is much broader than that obtained in the 138A mutant. The distance distribution of 68R1/140R1 in the 138A mutant is clearly multimodal and interestingly, the additional peaks of distances observed in the W138A variant are also observed in the WT background, but at a much higher intensity. It should be noted that the most probable interspin distance for the 68R1/140R1 mutant in the W138A mutant is also observed in the WT background protein (*see* dashed lines).

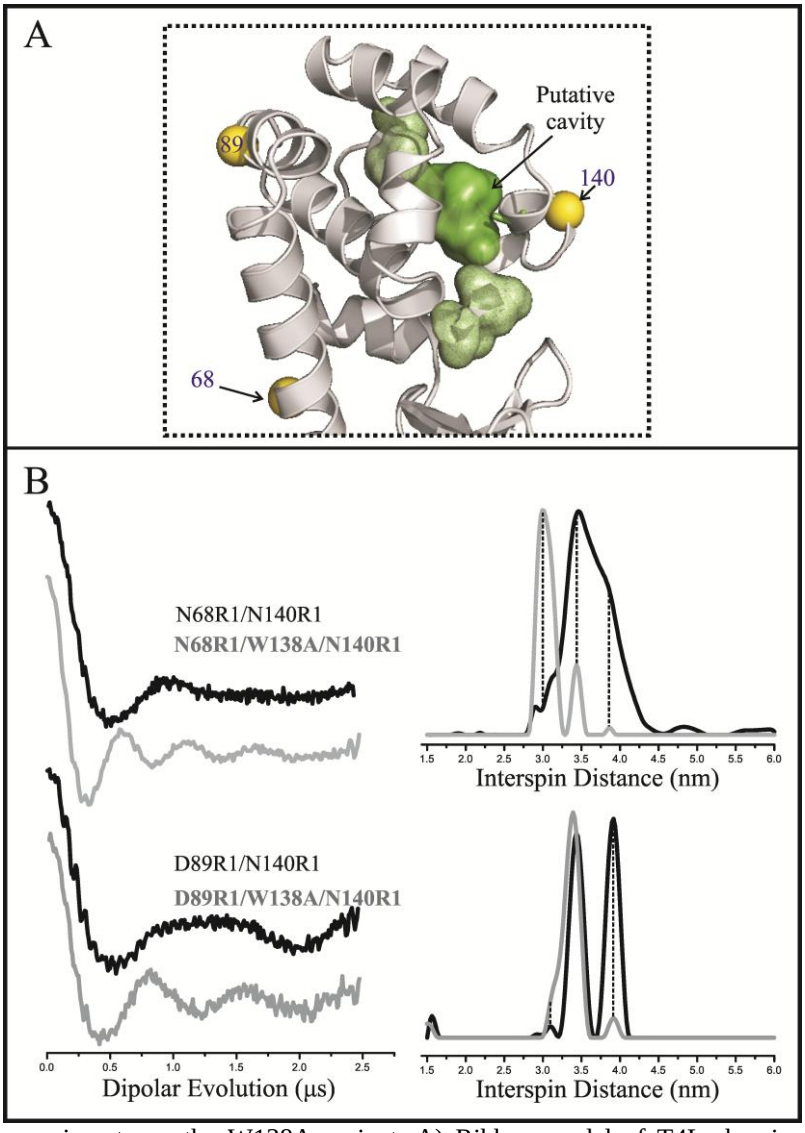


Figure 6.28: DEER experiments on the W138A variant. A) Ribbon model of T4L showing the putative cavity created by the W138A mutation. The yellow spheres at the Cα indicated the sites where R1 was introduced. B) Background subtracted dipolar evolution and corresponding distance distribution for N68R1/N140R1 and D89R1/N140R1 mutants in the WT (gray) and W138A (green) proteins. The dashed lines are used to help guide the eye.

Interestingly, the distance distribution of the doubly labeled mutant 89R1/140R1 in the WT background is bimodal with peaks of almost equal intensities, which reveals spatial delocalization of the nitroxide that can arise either from R1 rotameric equilibria or local conformational exchange. Both of these phenomena are expected to give rise to multicomponent

EPR spectra (*see* Chapter 3), however the CW EPR spectra for D89R1 [28] and N140R1 (Figure 6.22C) at 295K reflect a single spectral component. As discussed above, EPR experiments at temperatures below 293K on the N140R1 mutant in the WT background showed that a second highly immobilized component is populated (Figure 6.22D), thus it is possible that the immobile component observed at low temperatures in the CW EPR spectrum gives rise to one of the two most probable distances observed in the interspin distance distribution. The possible origins of the two most probable interspin distances and the immobile component observed in the 140R1 at low temperatures will be discussed below (*see* Discussion). It is unlikely that the spatial delocalization of the nitroxide for the 89R1/140R1 in the WT background arise from residue 89R1, since it is located in helix D, which is a well-ordered region of the protein with low crystallographic thermal factor [53]. In fact previous EPR studies showed that the CW EPR spectrum of residue D89R1 reflects relatively restricted and ordered motion suggesting limited backbone flexibility [28].

The distance distribution of the doubly labeled mutant 89R1/140R1 in the W138A background shows dramatic changes compared to that obtained in the WT background (Figure 6.28B). Specifically, in contrast to the distribution obtained in the WT background, the distance distribution exhibits a dominant peak corresponding to a most probable interspin distance of 34 Å, while a second peak of low intensity (< 5 % of intensity of the dominant peak) is observed corresponding to a interspin distance of 39 Å. Remarkably, the two peaks observed in the W138A variant correspond to similar interspin distances observed in the WT protein, but at a significantly different ratio suggesting that the W138A mutation did not introduce a new state not present in the WT protein, but it rather *shifts the equilibrium* towards the state giving rise to the shorter interspin distance (34 Å).

Future experiments using a conformationally constrained nitroxide side chain (*see* discussion) and a double resonance experiment that monitors spatial reorientations of a helix under physiological conditions [41] will be done to determine whether the bimodal distance distribution observed for the 89R1/140R1 pair represents two positions of helix I or two rotamers of 140R1.

6.7 Discussion and Conclusions:

The primary goal of this study was to evaluate the effects of substitutions that disrupt the native hydrophobic core packing interactions on the molecular flexibility of a protein with SDSL. T4 lysozyme was selected as a model system for this study since a myriad of core repacking substitutions have been engineered and characterized with CD spectroscopy and x-ray crystallography by Brian Matthews and co-workers [16,19,21-24]. Save for the L99A cavity creating mutation and the A98V overpacking mutation [9,16,26,31], the conformational flexibility of most of these variants has not been characterized in solution, thus SDSL provided a means to characterize the changes in the local flexibility for some of these mutants. In addition to the previously engineered mutants, new core mutations that were expected to disrupt the native packing in the core of T4 lysozyme were also designed in this study. The types of core repacking substitutions studied in this work include: large-to-small mutations (cavity-creating), small-to-large mutations (cavity filling and overpacking), buried charge substitutions, and compensating substitutions (packing and ionic pair).

With few exceptions, most of these substitutions excite either fast backbone motions (ns to ps) and/or slow conformational fluctuations (μs to ms) of the protein at the local level. Remarkably, the same type of mutation, say large-to-small, can excite motions on different time domains suggesting that the local environment may play a major role in determining the effects of the core repacking substitution. For example, the spectrum of residue 131R1, reporting on the flexibility of the H-helix, in the L121A/L133A mutant showed two states in slow exchange (lifetime > 100 ns), while the same in the I150A mutant reveal changes in the ns to ps motion of the backbone. In this regard, it is not possible to distinguish the type of substitution (*i.e.*, overpacking *versus* cavity-creating) by the changes in the conformational flexibility observed.

The context-dependent dynamic response to core substitutions observed in this study may be a general trait in globular proteins [54]. For example, core repacking substitutions in the SH3 domain of Fyn kinase [11] and the computational designed variant of protein Gβ1 [55] exert different dynamic responses to the proteins. Specifically, the core repacked variant of protein Gβ1 showed similar ns to ps time scale motions as the WT protein, but exhibits μs to ms fluctuations not present in the WT protein [55]. Conversely core substitutions in the SH3 domain had no effects on μs to ms fluctuations, but they rather enhance backbone flexibility on the ns to ps time scale [11]. It is likely that in these cases the local environment (*e.g.,* overall packing, tertiary interactions, strain energy) may have played a major role in determining the type of fluctuations enhanced by the core substitutions. It should be noted that the study presented here focused primarily on effects at the local level (same helix) of these repacking substitutions, however in the few cases where nonlocal effects were studied it was observed that for certain substitution the observed changes in the conformational flexibility were local (*e.g.,* 133G), but for others the changes in flexibility propagated along several helices of the C-terminal domain of the protein (*i.e.,* 102A/106A, I150A, 99A/153A).

It should be realized that even in cases where no spectral changes were observed from the substitution (*i.e.,* M102A, F104A, A98V, and V149L), it is possible that the mutation introduced new excited states of the protein, but at levels below detection (ca < 5 %) for EPR spectroscopy. One such example previously identified by NMR spectroscopy is the L99A mutant in T4L [9]. This large-to-small substitution in the hydrophobic core of the protein created a large cavity able to bind over 50 nonpolar ligands [56]. Remarkably, the high-resolution structure showed virtually no structural changes of the mutant compared to the WT protein [16], however relaxation dispersion experiments identified a short-lived low populated (3%) excited state

introduced by the L99A mutation, which apparently involved a second conformation of the short F-helix and nearby regions [26,29]. Such a conformation is expected to be too low to be detected by EPR spectroscopy. Indeed previous EPR studies showed no changes in the EPR spectra of residue T109R1 (helix F) in the L99A compared to the WT [57]. As discussed in Chapter 1, Akasaka have shown that pressure perturbation can be employed as a means to populate low-lying excited states of lower partial molar volume [58-59], thus high pressure EPR should provide a way to study the excited state created by the L99A mutant. Indeed, pioneering high-pressure EPR studies done by McCoy [57] showed significant differences in the spectra of residue T109R1 in the L99A mutant compared to that obtained in the WT background at elevated pressures. Specifically, the EPR spectrum of residue 109R1 in the L99A background at elevated pressures revealed a highly immobilized state of the nitroxide side chain not observed in the WT background, which suggests that for the 99A variant, high pressure populated a new state. Interestingly similar results were obtained for other sites along the C-terminal domain of the protein, while R1 in the N-terminal domain of the protein showed little changes at high pressures between the WT and the L99A protein [57] suggesting that the dynamic effects of the L99A mutation are essentially confined to the C-terminal domain, which is in agreement to NMR relaxation studies [9]. From the EPR data, McCoy et al. [57] estimated that the state giving rise to the immobile component was ~ 2 kcal/mol above the ground state at atmospheric pressure (corresponding to a population of 3.4 % at 1 atm), which is the same order of magnitude as that obtained from the relaxation dispersion experiments suggesting that the new state observed in high pressure EPR experiment is the same as that observed by relaxation dispersion. Similar high-pressure EPR experiments are underway for some of the variants studied here in

order to evaluate whether the mutations introduced short-lived excited states similar to those observed in the L99A mutant.

6.7.1 *Origins of the spectral changes of newly designed variants I150A and W138A.* Although structural information of the I150A variant is not available, the EPR spectral changes of residues 131R1 and 151R1 observed in the I150A mutant likely arise from changes in the fast backbone motions of helices H and J and not from disruption of tertiary interaction of the nitroxide with the environment (*ie.*, side chain of residue I150). This conclusion is based on careful examination of the high-resolution structures of the V131R1 and T151R1 proteins solved at 100K and 291K [60], which showed that the nitroxide side chains at these two sites are not in tertiary contact with the side chain of residue I150 as depicted in Figure 6.29. Previous studies showed that although the spectrum of residues 131R1 in the WT protein reflects a single component at room temperature, a relatively immobile state appear at low temperatures arising from tertiary interaction with the protein environment likely with the hydrophobic patch formed by residues A130 and I150 [60]. Temperature-dependent EPR studies done in this study on the 150A mutant revealed that the immobile component observed for 131R1 at low temperature in the WT background is also present in the I150A background (*see* Appendix) suggesting that the nitroxide side chain at site 131R1 does not interact with the side chain of the native residue I150, at least beyond the Cβ. In addition, the spectral changes, due to the I150A mutation, remain even at high temperatures

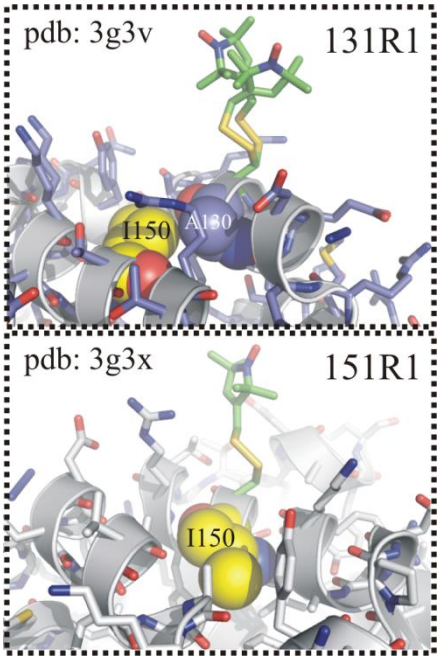


Figure 6.29: High-resolution structures of spin-labeled proteins 131R1 and 151R1 (Fleissner et al., 2009). The side chain of residues A130 and I150 are shown as cpk.

(40°C), where the putative hydrophobic interaction is not expected to contribute to the spectra, thus it is unlikely that the spectral changes observed in the 150A mutant arise from disruption of the tertiary interaction of the native isoleucine with the nitroxide side chain, but it rather arise from changes in the fast backbone motion of helices H and J due to the reduction of hydrophobic contact of the helix. It is interesting that the I150A mutation exerts different dynamic effects on the motion of 131R1 and 151R1, reporting on the dynamics of helices H and J, respectively. Specifically, the I150A mutation slightly reduced the rate of motion and increased the amplitude of motion of the nitroxide at site 131R1 by 5°, while the rate motion of residue T151R1 is reduced in the 150A protein. It is possible that the reduction in mobility of the J helix arises from local structural relaxation (*i.e.*, as inward movement of the helix) in response to the I150A mutation in order to partially fill the void created in the I150A mutation, whereby the motion of the helix becomes more restricted. It should be noted that such a structural relaxation had been observed in the structures of the cavity-creating mutants 121A and 153A studied here [16] and a reduction of the mobility of the nitroxide for 151R1 was observed in the latter mutant (*see* Figure 6.30 below).

The CW spectral changes and the dramatic changes in the interspin distance distribution observed for the W138A variant deserve comment. Although atomic resolution structure of the W138A mutant is not currently available, structural modeling showed that such substitution should create a cavity adjacent to helix I. The EPR spectrum of residue 140R1, which was used to report on the local changes in the dynamics of helix I, revealed that the environment around the nitroxide side chain changed from one in which the nitroxide was solvent-exposed to an environment of extensive tertiary contact in the W138A variant as judged by the significant reduction in the mobility of R1. Interspin distance measurements between residue 140R1 with

two reference sites that are distant from the site of the mutation showed significant structural changes in the helix I, specifically in both cases the most probable interspin distance is 5 Å shorter than that observed in the WT protein. Since the CW spectrum of residue 140R1 in the W138A mutant is dominated by a highly immobilized component, it is likely that the shorter interspin distance correspond to the protein conformation giving rise to the more immobile component. The new substate in the W138A mutant may be the result of a local structural rearrangement that fills the cavity with a side chain, perhaps with tyrosine 139. Preliminary studies have shown that the W138A mutant does not bind any of the ligands that bind to the 121A/133A and the 133G mutant (data not shown). Binding of other ligands such as p-cresol[11], indole, and indene to the W138A cavity mutant will be investigated in future studies. The origin of the bimodal distance distribution in the 89R1/140R1 pair in the WT background will be determined in future studies by using the conformationaly constrained bifunctional spin label (RX) shown in Figure 6.30 in combination with a pulsed EPR experiment (ELDOR) described by Fleissner and coworkers [41] that monitors the spatial reorientation of helices at physiological temperatures.

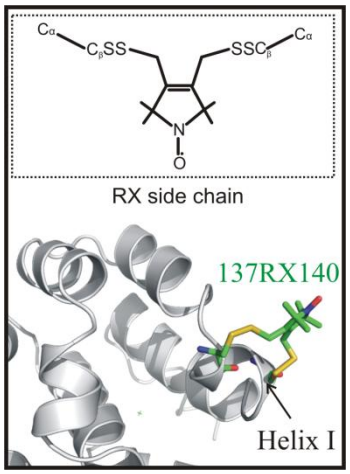


Figure 6.30: Introduction of bifunctional spin label RX in helix I to investigate helix fluctuation in the μs time scale. *Top panel*: RX side chain generated after reaction of a protein variant containing two cysteines (*i* and *i* +3 or *i* and *i* + 4) with the bifunctional spin label. Bottom panel: Model of the RX side chain introduced in helix I *via* sites 137 and 140.

6.7.2 *Origins of spectral changes for the L121A/L133A mutant*. Interestingly, the L121A/L133A variant may be a second example of a core mutation affecting relative populations of confomers. Even though the crystal structure showed subtle structural changes in the 121A/133A mutant compared to the WT protein [24], the EPR spectrum of residue 131R1 in

[11] p-cresol resembles the side chain of tyrosine, thus if Y139 fills the cavity in the apo state as suggested in this dissertation, the cavity mutant should be able to bind p-cresol.

this mutant showed significant changes in the mobility of R1 compared to that obtained for the WT protein, which suggests changes in the conformational flexibility of helix H in solution. Specifically, at 295K a relatively immobilized state of the nitroxide was observed in the L121A/L133A protein, which was not observed in the spectrum obtained in the WT protein. Osmotic perturbation studies showed that the populations giving rise to each spectral component in the L121A/L133A are in slow exchange (lifetime > 100ns; *see* Appendix). Remarkably, the addition of ligands changed the equilibrium of the two states observed for 121A/131R1/133A towards the mobile component. Notwithstanding the spectral changes observed in the 121A/133A protein, the interspin distance distribution obtained between residue 131R1 with three reference sites showed similar most probable distances compared to the WT protein (Figure 6.27), which would not explain the dramatic spectral changes observed in the CW spectral lineshape of residue 131R1. However, careful examination of the distance distribution revealed significant changes in the distance distribution between the R1 pairs in the WT and mutant proteins. Specifically, the distance distribution in the 121A/133A proteins showed significant intensities corresponding to shorter interspin distances (Δr = 8Å; *see* arrows in Figure 6.27). Remarkably, the shorter distances are also observed in the WT protein, but at a significantly lower intensities, which raises the interesting possibility that the 121A/133A mutation did not introduce a new state, but it rather shifted a pre-existing equilibrium. The distance distributions obtained from DEER spectroscopy were used to get an idea of the spatial localization of the nitroxide in the conformation that gives rise to the immobile state for residue 131R1 in the 121A/133A mutant. The preliminary analysis suggests that the nitroxide may be localized in the interface between helices H and J, specifically in the region where the native I150 resides. Giving the limited amount of data available at this point, alternative scenarios that may account

for the appearance of the immobile component for residue 131R1 cannot be completely ruled out. For example, subtle rotation of nearby helices G and/or J towards helix H, where 131R1 resides, can bring the nitroxide ring into tertiary contacts with side chains of the rotating helix. Thus, additional experiments involving interspin distance measurements and CW lineshape analysis for R1 introduced at sites in helices G and J are underway.

6.7.3 *R1 side chain as a sensor of backbone motions in stable helices*. This work highlights the utility of the R1 side chain as a sensor of even subtle changes in fast backbone motions arising from core repacking substitutions. The substitutions that changed the mobility of the nitroxide on the nanosecond time scale are shown in Figure 6.30 and Table 6.1. Previous studies based on molecular dynamic simulations of R1 at solvent exposed sites of α–helices, suggested that the site-to-site variations in the rate and order of motion arise primarily from tertiary interaction with the local environment and not from differential fast backbone motions [61-62]. However, in the present study many of the mutations that changed the rate and order of motion of the R1 side chain are completely buried and even in different helices to that where the R1 side chain was attached. As an example, the changes in the rate and order of motion observed for 72R1 and 131R1 in the I100A and A129M mutants, respectively cannot be explained by disruption of tertiary interaction of the nitroxide with its environment, but rather by changes in the backbone motions since both substitutions are in the opposite surface of the helix to which

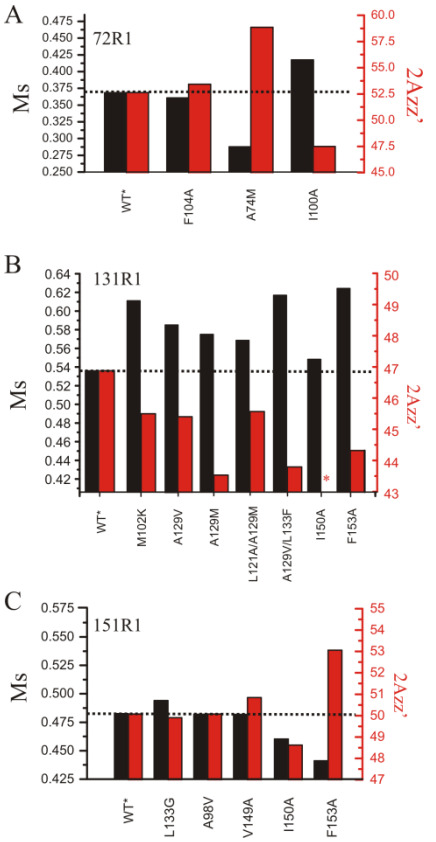


Figure 6.31: R1 as a sensor of changes in backbone motions. Scaled mobility (black) and 2Azz' (red) values for R1 in the WT* and mutant proteins are shown. 2Azz' for 131R1/150A (*) was unresolved.

the R1 is attached as depicted in Figure 6.31. It should be noted that in some cases the core substitutions clearly introduced tertiary interactions of the nitroxide side chain with its environment (*i.e.*, 133G, 121A/133A, 146V, 133D), but in such cases the tertiary interactions were usually revealed as an immobile component in the EPR spectra.

One exception appears to be the 72R1/A74M mutant, which showed a significant reduction in the mobility and amplitude of motion of R1 in the overpacking mutant. Specifically, in this case the weakly ordered motion ($S \leq 0.50$) of 72R1 observed in the WT protein becomes strongly ordered (S = .61) in the A74M mutant, which corresponds to a decrease in the amplitude of motion of 5º. Given the lack of structural information for this mutant, the origins for the reduction in the mobility of the nitroxide are not entirely understood, however previous studies have shown that the EPR spectra of sites exhibiting strongly ordered z-axis anisotropic motion of the nitroxide ($S \geq 0.50$), such as the 72R1/74M variant, generally reflect tertiary interaction with the environment [36,42-44], thus it is likely that the dramatic increase in the order of motion of the nitroxide arises from some local structural rearrangement of the protein (*i.e.*, segmental rotation of the C-terminal region of the C-helix) that introduce tertiary interactions of the protein with the nitroxide side chain.

Table 6.1: Mutants reflecting changes in the nanosecond motions of the R1 side chain as judged by MOMD simulations.

Sample	τ_m (ns)	S_m
72R1	**2.35**	**0.50**
72R1/74M	2.96	0.61
72R1/I100A	2.12	0.42
131R1	**2.05**	**0.40**
L129M/131R1	1.63	0.45
131R1/I150A	2.22	0.28
131R1/F153A	2.10	0.37
151R1	**2.47**	**0.41**
I150A/151R1	2.98	0.37
151R1/F153A	2.46	0.43

However, the slower motion of the nitroxide at site 72R1 in the A74M mutant could also be a reflection of further reduction on the nanosecond backbone fluctuations of the helix C. Additional studies involving interspin distance measurement with DEER spectroscopy will be essential to distinguish between these two possibilities.

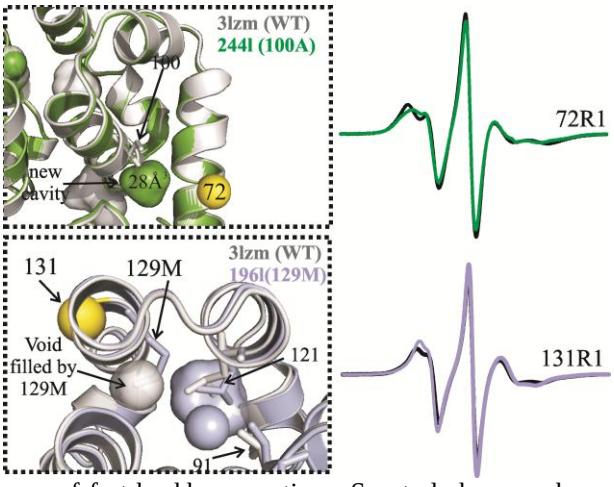


Figure 6.32: R1 side chain as sensor of fast backbone motions. Spectral changes observed for residues 72R1 and 131R1 in the 100A and 129M mutants, respectively.

6.7.4 *Implications of core substitutions in protein function and design*. Most globular proteins have interior packing densities similar to those of crystalline solids (Richards, 1974). However, careful examination of protein core packing suggests that interior packing can be heterogeneous and suboptimal and is best described as "nuts and bolts in a jar" rather than a jigsaw puzzle [15]. Analysis of high-resolution structures revealed that most proteins contain cavities and pockets that can range from a few to hundreds of cubic angstroms [15], which in general are expected to be destabilizing due to the loss of van der Waals interaction [16-17]. One implication of this is that improving the packing of the protein interior may be a general way to increase protein stability. To this end several groups have successfully increased the stability of several proteins by optimizing the interior packing [10,63]. One question that arises from these studies is that if suboptimal protein packing has an energetic cost, why has evolution not selected against such packing defects given the importance of protein stability? One interpretation for the prevalence of packing deficiencies in proteins may be that suboptimal core packing may play an important role in protein function, thus an increase in protein packing can lead to disruption of the functional properties of proteins. For example, the internal cavities present in the core of

myoglobin are essential for providing a pathway for oxygen migration from the solvent to the buried binding site [64]. In addition to providing internal pathways for ligand migration, core packing defects have been proposed to play important roles in functional protein dynamics, such as in facilitating conformational switching and allostery [10,65-66]. Thus if suboptimal packing is important for protein flexibility, which in turn is essential for function, it is possible to design proteins that are too stable and rigid for the incorporation of the desired functional property. Results from this work and many others [10-12,29,54-55] clearly show that substitutions that disrupt the native hydrophobic core packing interactions can excite fast backbone motions (ns to ps) and/or slow conformational fluctuations (μs to ms) of proteins and even change the energy landscape whereby new states with distinct functional properties are populated [26]. Thus given the role that these motions can play in protein function, it would be worth exploring how these substitutions can be used in protein rational design strategies as a tool to introduce functional protein dynamics.

6.8 Appendix

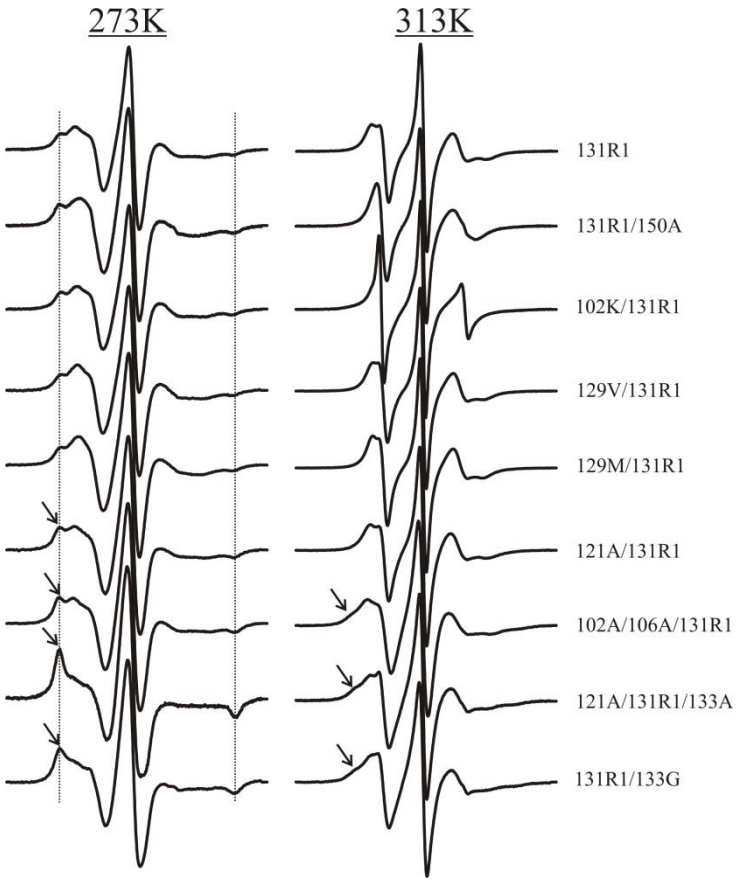


Figure 6.A.1: Temperature-controlled EPR studies of 131R1 in the WT and mutant proteins

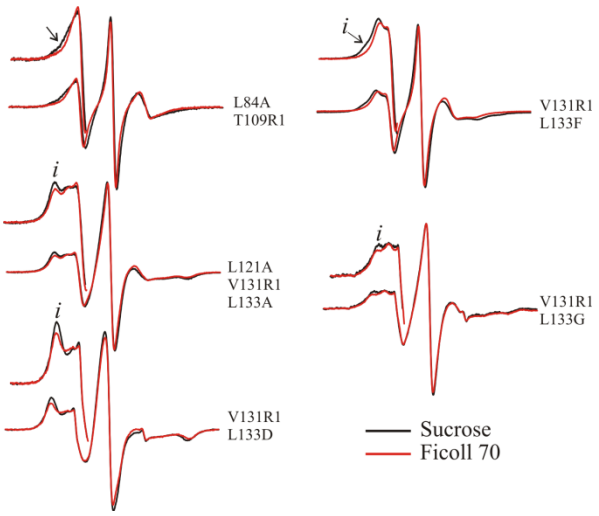
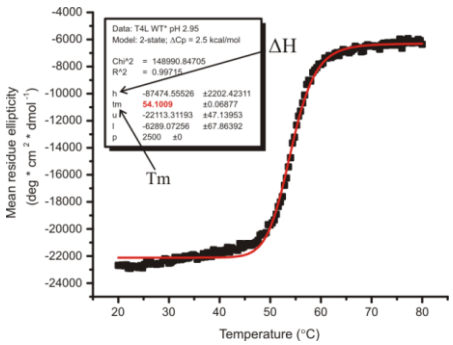


Figure 6.A.2 Osmolyte-perturbation studies on selected hydrophobic core mutants

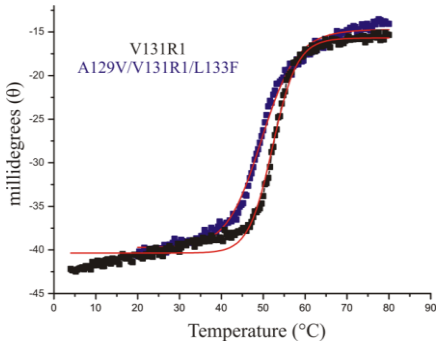
Table 6.A.1 Scaled mobility (Ms) and hyperfine splitting values (2Azz') of R1 at the corresponding sites in the WT and mutant proteins.

Sample	Central linewidth (δ)	δ^{-1}	Ms	2 Azz' (G)
48R1	2.47	0.41	0.80	
L46A/48R1	2.52	0.40	0.78	
72R1	3.99	0.25	0.37	52.63
72R1/F104A	4.03	0.25	0.36	53.43
72R1/A74M	4.51	0.22	0.29	58.85
72R1/I100A	3.73	0.27	0.42	47.49
109R1	2.88	0.35	0.64	
L84A/109R1	2.90	0.35	0.63	
M102A/109R1	2.97	0.34	0.61	
M102A/M106A/109R1	3.07	0.33	0.58	
F104A/109R1	2.78	0.36	0.67	
109R1/V111I	2.66	0.38	0.72	
131R1	3.22	0.31	0.54	46.89
M102K/V131R1	2.97	0.34	0.61	45.51
M102A/M106A/131R1	3.37	0.30	0.50	
L121A/131R1	3.50	0.29	0.47	
A129V/131R1	3.05	0.33	0.59	45.42
A129M/131R1	3.08	0.32	0.57	43.54
L121A/A129M/131R1	3.11	0.32	0.57	45.57
131R1/L133A	3.14	0.32	0.56	
L121A/131R1/L133A	3.75	0.27	0.41	
131R1/L133G	3.51	0.29	0.47	
131R1/L133F	3.44	0.29	0.48	
A129V/131R1/L133F	2.95	0.34	0.62	43.78
131R1/I150A	3.18	0.31	0.55	
131R1/F153A	2.92	0.34	0.62	44.33
L99A/V131R1/F153A	3.07	0.33	0.58	46.34
140R1	2.47	0.40	0.80	43.04
W138A/140R1	5.22	0.19	0.20	61.73
140R1/A146V	2.71	0.37	0.70	46.17
151R1	3.43	0.29	0.48	50.07
133G/151R1	3.38	0.30	0.49	49.91
A98V/151R1	3.43	0.29	0.48	50.07
V149A/151R1	3.44	0.29	0.48	50.84
V149L/151R1	3.15	0.32	0.56	48.78
I150A/151R1	3.53	0.28	0.46	48.62
151R1/F153A	3.62	0.28	0.44	53.06
L99A/T151R1/F153A	3.59	0.28	0.45	51.17

A) *T4L WT**



B) *T4L V131R1 & A129V/V131R1/L133F*



C) *T4L N140R1 & N140R1/A146V*

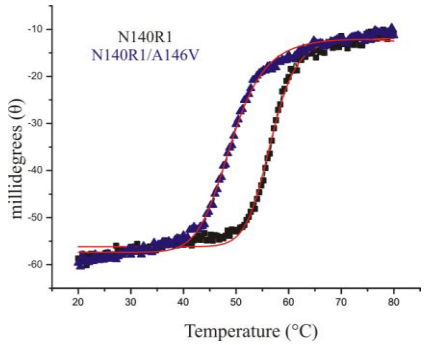


Figure 6.A.3 Temperature denaturation studies on core mutants of T4L. The denaturation curves were fit on Origin 6 using equations described in Chapter 8.

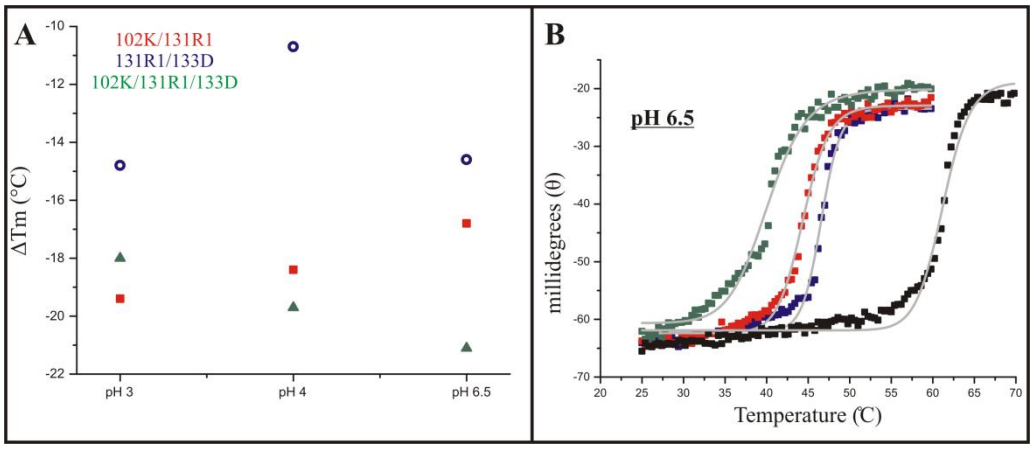


Figure 6.A.4: Temperature denaturation studies of buried polar variants of T4L at different pH. A) Change in the Tm ($Tm^{mutant} - Tm^{V131R1}$) as a function of pH. The buffers used for each pH are the same as reported in Dao-pin et al., 1991. B) Temperature-denaturation studies of the 131R1 (black), 102K/131R1 (red), 131R1/133D (blue), and 102K/131R1/133D (green) at pH 6.5. At pH 6.5, the thermal unfolding was not reversible for any of the 4 mutants (*see* Table below).

Table 6.A.2 Reversibility of R1-labeled mutants at different solvent pH

pH	Sample	% Reversibility†
4	131R1	82
	102K/131R1	92
	131R1/133D	47
	102K/131R1/133D	86
6.5	131R1	46
	102K/131R1	50
	131R1/133D	45
	102K/131R1/133D	55

†Reversibility was assessed by restoration of the signal upon cooling.

Table 6.A.3: Thermal stability of the spin-labeled T4L core mutants

Sample	*ΔH (kcal/mol)*	*ΔTm (℃)*	*ΔΔG (kcal/mol)*
WT*	87.50	-	-
D72R1	88.30	+1.6	0.42
D72R1/A74M	72.12	+1.4	0.30
D72R1/I100A	63.28	-5.5	-1.18
N140R1	88.58	+2.5	0.65
N140R1/A146V	60.91	-5.0	-1.03
W138A/N140R1	73.18	-6.2	-1.53
V131R1	72.47	-1.8	-0.41
V131R1/I150A	56.93	-6.7	-1.34
A129V/V131R1/L133F	49.18	-4.4	-0.74
M102K/V131R1	18.71	-21.2	-2.48
V131R1/L133D	35.31	-16.6	-2.56
M102K/V131R1/L133D	21.00	-19.8	-2.37
V131R1	108.00	-	-
M102K/V131R1	37.90	-18.4ǂ	-3.01ǂ
V131R1/L133D	49.03	-10.7ǂ	-1.88ǂ
M102K/V131R1/L133D	39.51	-19.7ǂ	-3.39ǂ

WT* refers to cysteine-less pseudo wild-type T4L. $\Delta Tm = Tm^{mutant} - Tm^{WT*}$, except for the mutants labeled with the ǂ symbol. Tm of the WT* protein is 54.1℃ at pH 3.0 [ΔG (25℃) = -11.1 kcal/mol]. ΔH is the enthalpy of unfolding at pH 3.0 (save for the shaded mutants, which were denatured at pH 4). ΔΔG is the free energy of unfolding of the mutant at the Tm of the WT* protein. A constant ΔC_p = 2.5 kcal/mol*deg [16] was used in most calculations, save for the buried polar variants for which ΔC_p = 1.8 kcal/mol*deg was used as in Dao-pin et al., 1991. ǂ $\Delta Tm = Tm^{mutant} - Tm^{V131R1}$ and ΔΔG is the free energy of unfolding of the mutant at the Tm of the V131R1 protein. Tm of the V131R1 protein is 61.3℃ at pH 4.0

Table 6.A.4: Dynamic parameters obtained from MOMD simulations

Sample	f_i	τ_i **(ns)**	S_i	f_m	τ_m **(ns)**	S_m
48R1	0	-	-	1	1.4	0.25
L46A/48R1	.08	8.3	-	.92	1.4	0.24
72R1	0	-	-	1	2.35	0.50
72R1/74M	0	-	-	1	2.96	0.61
72R1/I100A	0	-	-	1	2.12	0.42
109R1	.96	1.7	0.44	.04	0.9	-
L84A/T109R1	.96	2.3	0.33	.04	0.9	-
M102A/M106A/109R1	.96	2.3	0.46	.04	0.9	-
109R1/V111I	.96	1.5	0.37	.04	0.9	-
131R1	0	-	-	1	2.05	0.40
M102A/M106A/131R1	.07	16.7	-	.93	2.00	0.40
L129M/131R1	0	-	-	1	1.63	0.45
131R1/L133A	.06	16.7	-	.94	2.15	0.39
131R1/133G	.25	15.0	-	.75	2.17	0.40
L121A/131R1/L133A	.70	16.0	-	.30	2.43	0.40
131R1/I150A	0	-	-	1	2.22	0.28
131R1/F153A	0	-	-	1	2.10	0.37
140R1	0	-	-	1	1.21	0.41
140R1/A146V	.45	11.1	-	.55	1.40	0.51
151R1	0	-	-	1	2.47	0.41
I150A/151R1	0	-	-	1	2.98	0.37
151R1/F153A	0	-	-	1	2.46	0.43

Table 6.A.5: Interspin distance distributions from DEER spectroscopy

Mutant	Helices	Most probable R1-R1 distance (Å)	Width of distance distribution‡ (Å)
N68R1/V131R1	C-H	38.1	10.3
N68R1/121A/131R1/133A		(28.9)[23%] 39.1[74%]	11.0
D89R1/V131R1	D-H	33.6	
D89R1/121A/131R1/133A		(26.8)[36%] 33.6[64%]	
T109R1/V131R1	F-H	31.4	
T109R1/121A/131R1/133A		(23.4)[20%] & 29.8[37%]	
D89R1/N140R1	D-I	34.4[52%], 39.0[48%]	3.1, 2.4
D89R1/W138A/N140R1		33.9	4.6
N68R1/N140R1	C-I	34.5	
N68R1/W138A/N140R1		29.9[77%] & 34.3[23%]	3.6, 2.1

‡ The width of the distribution was taken as the range of the interspin distance above 16 % of the peak intensity.
The relative percentage was taken as the individual intensity of the peak divided by the total intensity of all peaks.
Data kindly provided by Dr. Mark R. Fleissner (Fleissner et al. [40]; unpublished)

6.9 References

1. Vieille, C., and Zeikus, G.J. 2001. Hyperthermophilic enzymes: Sources, uses, and molecular mechanism of thermostability. *Micr. Mol. Biol. Rev*. 65:1-43.
2. Petsko, G.A. 2001. Structural basis for thermostability in Hyperthermophilic proteins, or "there is more than one way to skin the cat". *Meth. Enzym*. 334:469-478.
3. Vihinen, M. 1987. Relationship of protein flexibility to thermostability. *Prot. Engin*. 1:477-180.
4. Závodszky, P., Kardos, J., Svingor, A., and Pestko, G.A. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci.* 95:7406-7411.
5. Wolf-Watz, M., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E.Z., and Kern, D. 2004. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struc. Mol. Biol.* 11:945-949.
6. Telium, K., Olsen, J.G., Kragelund, B.B. 2011. Protein stability, flexibility, and function. *Bioch. Bioph. Acta* 1814:969-976.

7. Tokuriki, N, and Tawfik, D.S. 2009. Protein dynamism and evolvability. *Science*, 324:203-207.
8. Schoichet, B.K., Baase, W.A., Kuroki, R., and Matthews, B.W. 1995. A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci.* 92:452-456.
9. Mulder, F.A.A., Hon, B., Muhandiram, D.R., Dahlquist, F.W., and Kay, L.E. 2000. Flexibility and ligand exchange in a buried cavity mutant of T4 lysozyme studied by multinuclear NMR. *Biochemistry* 39:12614-12622.
10. Lee, C., Maeng, J., Kocher, J.P., Lee, B., and Yu, M.H. 2001. Cavities of α1-antitrypsin that play structural and functional roles. *Protein Sci.* 10:1445-1453.
11. Mittermaier, A., and Kay, L.E. 2004. The response of internal dynamics to hydrophobic core mutations in the SH3 domain from the Fyn tyrosine kinase. *Protein Sci.* 13:1088-1099.
12. Whitley, M.J., Zhang, J., and Lee, A.L. 2008. Hydrophobic core mutations in Cl2 globally perturb fast side-chain dynamics similarly without regard to position. *Biochemistry*, 19:8566-8576.
13. Gabellieri, E., Balestreri, E., Galli, A., and Cioni, P. 2008. Cavity-creating mutations in *Pseudomonas aeruginosa* Azurin: Effects on protein dynamics and stability. *Biophy. J.* 95:771-781.
14. Richards, F.M. 1974. The interpretations of protein structures: total volume, group volume distributions, and packing density. *J. Mol. Biol.* 82:1-14.
15. Liang, J., and Dill, K.A. 2001. Are proteins well-packed? *Biophy. J.* 81:751-766.

16. Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P., and Matthews, B.W. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* 255:178-183.
17. Vlassi, M., Cesareni, G., and Kokkinidis, M. 1999. A correlation between the loss of hydrophobic core packing interactions and protein stability. *J. Mol. Biol*. 285:817-827.
18. Baase, W.A., Liu, L., Tronrud, D.E., Matthews, B.W. 2010. Lessons from the lysozyme of phage T4. *Protein Sci.* 19:631-641.
19. Karpusas, M., Baase, W.A., Matsumura, M., and Matthews, B.W. 1989. Hydrophobic packing in T4 lysozyme probed by cavity-filling mutations. *Proc. Natl. Acad. Sci.* 86:8237-8241.
20. Dao-pin, S., Anderson, D.E., Baase, W.A., Dahlquist, F.W., and Matthews, B.W. 1991. Structural and thermodynamic consequences of burying a charge residue within the hydrophobic core of T4 lysozyme. *Biochemistry*, 30:11521-11529.
21. Hurley, J.H., Baase, W.A., and Matthews, B.W. 1992. Design and structural analysis of alternative hydrophobic core packing arrangements in bacteriophage T4 lysozyme. *J. Mol. Biol.* 234:1143-1159.
22. Baldwin, E., Xu, J., Hajiseyedjavadi, O., Baase, W.A., and Matthews, B.W. 1996. Thermodynamic and structural compensation in "Size-switch" core repacking variants of bacteriophage T4 lysozyme. *J. Mol. Biol.* 259:542-559.
23. Baldwin, E., Baase, W.A., Zhang, X., Feher, V., and Matthews, B.W. 1998. Generation of ligand binding sites in T4 lysozyme by deficiency-creating substitutions. *J. Mol. Biol.* 277:467-485.

24. Xu, J., Baase, W.A., Baldwin, E., and Matthews, B. 1998. The response of T4 lysozyme to large-to-small substitutions within the core and its relation with the hydrophobic effect. *Protein Sci.* 7:158-177.

25. Matthews, B.W. 1995. Studies of protein stability with T4 lysozyme. *Adv. Prot. Chem*. 46:249-277.

26. Bouvignies, G., Vallurupalli, P., Hansen, D.F., Correia, B.E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D., and Kay, L.E. 2011. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature*, 477:111-114.

27. Feher, V.A., Baldwin, E., and Dahlquist, F.W. 1996. Access of ligands to cavities within the core of a protein in rapid. *Nat. Struct. Biol.* 3:516-521.

28. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35:7692-7704.

29. Mulder, F.A.A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. 2001. Studying excited states of protein by NMR spectroscopy. *Nature Struct. Biol*. 8:932–935.

30. Columbus, L., and Hubbell., W.L. 2004. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* 43:7273-7287.

31. Anderson, D.E., Lu, J., McIntosh L., and Dahlquist, F.W. 1990. The folding, stability, and dynamics of T4 lysozyme: A pesprective using nuclear magnetic resonance. In *NMR of Proteins*. Ed. by Clore M. and Gronenborn, A. New York: Macmillan Press; pp 258–304.

32. López, C.J., Fleissner, M.R., Guo, Z., Kusnetzow, A.K., and Hubbell, W.L. 2009. Osmolyte-perturbation reveals conformational equilibria in spin-labeled proteins. *Protein Sci.* 18:1637-1652.

33. Damodaran, S., and Song, K.B. 1986. The role of solvent polarity in the free energy of transfer of amino acids side chains from water to organic solvents. *J. Biol. Chem*. 261:7220-7222.

34. Daura, X., Hünenberger, P.H., Mark,A.E., Querol, E., Avilés,F.C., and Gunsteren W.F.V. 1996. Free energies of transfer of trp analogs from chloroform to water: Comparison of theory and experiment and the importance of adequate treatment of electrostatic and internal interactions. *J. Am. Chem. Soc.* 18:6285-6294.

35. Fauchere, J.L., and Pliska, V. 1983. Hydrophobic parameters pi of amino acid side-chains from the partitioning of N-acetyl-amino-acids amides. *Eur J. Med. Chem*. 18:369-375.

36. Columbus, L.M. 2001. Investigating backbone and side chain dynamics of α-helices in the nanosecond regime by site-directed spin labeling. PhD thesis, University of California, Los Angeles.

37. Columbus, L., Kálai, T., Jeko, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in alpha-helices:analysis by variation of side chain structure. *Biochemistry* 40:3828-3846.

38. Liu, R., Baase, W.A., and Matthews, B.W. 2000. The introduction of strain and its effects on the structure and stability on T4 lysozyme. *J Mol. Biol*. 295:127-145.

39. Nicholson, H., Tronrud, D.E., Becktel, W.A., and Matthews, B.W. 1992. Analysis of the effectiveness of proline substitutions and glycine replacements in increasing the stability of phage T4 lysozyme. *Biopolymers* **32**, 1431-1441.

40. Fleissner, M.R., Brustad, E.M., Kálai, T., Altenbach, C., Schultz, P.G., and Hubbell, W.L. 2009. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci.* 106:21637-21642.
41. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011 Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246.
42. Langen, R., Oh, K.J., Cascio, D., and Hubbell, W.L. 2000. Crystal structures of spin labeled T4 lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry,* 39:8396-8405.
43. Lietzow, M.A., and Hubbell, W.L. 2004. Motion of spin label side chains in cellular retinol-binding protein: Correlation with structure and nearest-neighbor interactions in an antiparallel β–sheet. *Biochemistry* 43:3137-3151.
44. Fleissner, M.R. 2007. X-ray structures of nitroxide side chains in proteins: a basis for interpreting distance measurements and dynamic studies by electron paramagnetic resonance. PhD thesis, University of California, Los Angeles.
45. Gilson, M.K., Rashin, A., Fine, R., and Honig, B. 1985. On the calculation of electrostatic interactions in proteins. *J. Mol. Biol*. 183:603-516.
46. Huang, L. and Chen, C. 2010. Autoprocessing of human immunodefficiency virus type 1 protease miniprecursor fusion in mammalian cell. *AIDS Reas. Ther*. 7:27-37.
47. Isom, D.G., Cannon, B.R., Castañeda, C.A., Robinson, A., and García-Moreno, B. E.2008. High-tolerance for ionizable residues in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci.* 105:17784-17788.

48. Lovell, S. C., Word, M.J., Richardson, J.S., and Richardson, D.C. 2000.The penultimate rotamer library. Proteins, 40:389-408.

49. Hendsch, Z.S., and Tidor, B. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* 3:211-226.

50. Waldburger, C.D., Schildbach, J.F., and Sauer, R.T. 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* 2:122-128.

51. Lam, S.Y., Yeung, R.C.Y., Yu, T., Sze, K., and Wong, K. 2011. A rigidifying salt bridge favors the activity of thermophilic enzyme at high temperatures at the expense of low temperature activity. *PLOS Biology,* 9:1-8.

52. Kim, J.M., Altenbach, C., Kono, M., Oprian, D.D., Hubbell, W.L., and Khorana, G.H. 2004. Structural origins of constitutive activation in rhodopsin: Role of the K296/E113 salt bridge. *Proc. Natl. Acad. Sci.* 34:12508-12513.

53. Weaver, L.H. and Matthews, B.W. 1987 Structure of bacteriophage T4 lysozyme refined at 1.7A resolution. *J. Mol. Biol.* 193:189-199.

54. Clarkson, M.W. and Lee, A.L. 2004. Long-range dynamic effects of point mutations propagate through side chains in the serine protease inhibitor eglin c. *Biochemistry* 43:12448-12458.

55. Crowhurst, K.A. and Mayo, S.L. 2008. NMR-detected conformational exchange observed in a computationally design variant of Gβ1. *Prot. Eng. Des. Sel.* 21:577-587.

56. Morton, A. and Matthews, B.W. 1995. Specificity of ligand binding in a buried nonpolar cavity in T4 lysozyme. *Biochemistry* 34:8576-8588.

57. McCoy, J.J. 2011. High pressure EPR on spin labeled proteins. PhD thesis, University of California, Los Angeles.

58. Li, H. and Akasaka, K. 2006. Conformational fluctuations of proteins revealed by variable pressure NMR. *Bioch. Biophy. Acta* 1764:331-345.

59. Akasaka, K. 2006. Probing conformational fluctuation of proteins by pressure perturbation. *Chem Rev.* 106:1814–1835.

60. Fleissner, M.R., Cascio, D., and Hubbell, W.L. 2009. Structural origin of weakly ordered nitroxide motion in spin-labeled proteins. *Protein Sci.* 18:893-908.

61. Sezer, D., Freed, J.H., and Roux, B. 2008. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine α-helix. *J. Phys. Chem.* 112:5755-5767.

62. Seezer, D., Freed, J.H., and Roux, B. 2009. Multifrequency electron spin resonance of a spin-labeled protein calculated from molecular dynamics simulations. *J. Am. Chem. Soc.* 131:2597-2605.

63. Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'bried, R., Sturtevant, J.M., and Regan, L. 1996. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Sci.* 5:1584-1593.

64. Brunori, M, and Gibson, Q.H. 2001. Cavities and packing defects in the structural dynamics of myoglobin. *EMBO Reports*, 2:674-679.

65. Kadirvelraj, R., Sennett, N.C., Polizzi, S.J., Weitzel, S., and Wood, Z.A. 2011. Role of packing defects in the evolution of allostery and induced-fit in human UDP-glucose dehydrogenase. *Biochemistry,* 50:5780-5789.

66. Ogata, K., Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., Nakamura, H., Nishimura, Y., Ishii, S., and Sarai A. 1996. The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and *trans*-activation. *Nat. Struct. Biol.* 3: 178-187.

Chapter 7: Application of site-specific immobilization of spin-labeled proteins for EPR studies

7.1 Summary

The immobilization of proteins on solid surfaces for the characterization of their functional properties *in-vitro* and for high-throughput functional proteomics has become an increasingly important area of research. However, the number of biophysical techniques capable of studying the conformational and structural properties of immobilized proteins and protein complexes on solid matrices are limited. SDSL in combination with EPR spectroscopy is well-suited for the study of the structure and dynamics of soluble and membrane-bound proteins immobilized on solid matrices. However, the nonspecific nature of most commonly used immobilization strategies can be detrimental to the structural integrity of proteins, and in some cases, it can lead to immobilized proteins exhibiting much lower activities than those in solution. In this work a new strategy was employed that enabled site-specific immobilization of spin-labeled proteins such that the nitroxide side chain and functionally important regions of the spin-labeled protein are oriented away from the molecular surface of the matrix. EPR studies demonstrated that the strategy employed here effectively immobilizes the protein on the X-band EPR time scale and revealed that the structural and conformational properties of a spin-labeled protein coupled site-specifically to a solid support were similar to those in solution. These studies will be fundamental for the development of a flow system whereby the structural and dynamical effects of other molecules (*e.g.*, ligands, substrates, proteins, osmolytes) can be assayed sequentially on the same spin-labeled protein.

7.2 Introduction

The immobilization of proteins onto solid surfaces has become increasingly important in the area of protein science, particularly in the development of protein biochips for high-throughput analysis of protein-protein, protein-ligand, and protein-nucleic acid interactions, and for probing expression levels [1-2]. Protein immobilization on solid supports has also been shown to be useful for single-molecule enzymology studies [3] and to be highly effective in (1) increasing protein stability [4], (2) preventing protein aggregation, and (3) increasing the yield of functional proteins during *in-vitro* refolding studies [5-6]. Protein immobilization methods include nonspecific random adsorption (through hydrophobic or ionic interaction), nonspecific covalent coupling *via* amine, carboxyl, and thiol groups, bioaffinity noncovalent binding, and site-specific covalent coupling to functionalized matrices [2]. By far the most common strategy for immobilization of protein into solid supports involves coupling to activated matrices, such as CNBr-acvitated sepharose, *via* the primary amine groups in lysine side chains (Figure 7.1). EPR studies have shown that spin-labeled proteins attached to CNBr sepharose *via* lysine residues are indeed immobilized on the X-band EPR time scale (*see* Chapter 3 and [7]) and have similar structures compared to proteins in solution [8-9]. Attachment *via* native residues and nonspecific random adsorption have broad applicability and are straightforward to use since they do not require modification of the native protein, however they generally involve heterogeneous multiple point attachment, which leads to random orientation of the protein in the matrix. The nonspecific orientation of the protein in the matrix can result in a potential loss of activity arising from steric hindrance of the functional important protein surface by the matrix or from coupling *via* functionally important residues (*e.g.*, catalytic lysines). In addition, nonspecific adsorption to hydrophobic matrices has been shown to induce protein denaturation [10-11].

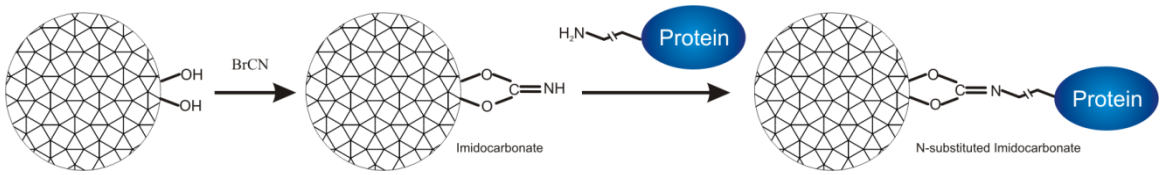


Figure 7.1: Commonly used strategy for protein immobilization *via* amino groups.

In recent years, site-specific coupling of proteins to solid surfaces has become the method of choice for protein immobilization due to the ability to orient the protein on the solid surface such that active sites and functionally important regions are positioned away from the molecular surface of the matrix in order to preserve the biological activity. Indeed, recent studies have shown that proteins and antibodies immobilized in an oriented fashion exhibit activities similar to proteins in solution and higher than those proteins randomly oriented [12-14]. Site-specific protein immobilization can involve covalent and noncovalent binding. Due to the inherent high specificity and reversibility, coupling of proteins through affinity tags such as the hexahistidine (His_6) to nickel-coated matrices is among the most commonly used site-specific noncovalent immobilization strategy. However the coupling *via* His_6 tag is not very strong (K_d = 1 μM) and can be affected by certain additives (*e.g.*, EDTA), pH, and ionic strength of the buffer [15]. Alternatively, noncovalent protein coupling between site-specifically biotinylated proteins to matrices functionalized with avidin/streptavidin has become a more sophisticated approach to immobilize proteins to solid surfaces due to the high specificity and stability of the biotin-streptavidin interaction [16].

Even though recent studies have shown that proteins oriented in solid surfaces *via* site-specific attachment retain their activities [16], direct studies on the conformational and structural properties of immobilized proteins and protein complexes on solid matrices remains a challenge for many biophysical methods (*e.g.*, CD spectroscopy, x-ray crystallography). SDSL in

combination with EPR spectroscopy is well-suited for the studies of immobilized proteins and protein complexes, since it does not rely on optical absorption and, save for magnetic resins, it is insensitive to the chemical groups often present in commonly used resins, while at the same time it is highly sensitive to proteins bearing nitroxide side chains, which enables the study of even low amounts of spin-labeled proteins (~ 50 pmol) in the matrix. In addition, SDSL is not limited to the study structure and dynamics of immobilized proteins, but it can also be used to study protein-protein interaction between soluble and membrane-bound spin-labeled proteins with proteins immobilized on solid matrices. A recent study showed that EPR spectroscopy can also be used as a tool to probe the degree of immobilization of proteins on the solid surface [7].

In order to study the structural and conformational properties of spin-labeled proteins and protein complexes coupled to solid supports, it is important to develop a system that enables site-specific immobilization of spin-labeled proteins such that the nitroxide side chain and functionally important regions are oriented away from the matrix. The platform for coupling of spin-labeled proteins to the solid support matrix for EPR application should be highly specific, stable under various experimental conditions (*i.e.*, temperature, pH, ionic strength), unreactive to commonly used nitroxide reagents (thiol-unreactive), and highly efficient under mild conditions.

In this work genetically encoded bioorthogonal[1] functional groups were introduced in T4 lysozyme (T4L) for site-specific modification with a "tag" and subsequent immobilization of the protein onto a functionalized resin. In this strategy, a genetically modified bioorthogonal[2] suppressor tRNA/aminoacyl-tRNA synthetase pair, generated by directed evolution methods in Peter Schultz's laboratory [17], were used to specifically incorporate an unnatural amino acid to

[1] Bioorthogonal refers to non-native chemical functionality

[2] Not endogenous to the organism used for recombinant expression (*E. coli*). The tRNA/tRNA synthetase was derived from archaea *Methanococcus jannaschii*, thus it does not cross-react with the endogenous synthetases.

recombinant T4L in response to the amber (TAG) codon (Figure 7.2). Using this strategy more than 30 unnatural amino acids have been incorporated into recombinant proteins expressed in *E. coli*, yeast, and mammalian expression systems (*see* review by Xie and Schultz [18]). The unnatural amino acid *p*-acetylphenylalanine (denoted p-AcPhe hereon) was introduced at solvent-exposed sites of T4L site-specifically for selective modification of the unique keto group with a tag for subsequent immobilization. Recent studies showed that the T4L bearing p-AcPhe at solvent exposed sites can be expressed at relatively high levels and the unnatural amino acid has no effects on the structure of the protein as judged by the similar backbone fold of the 131p-AcPhe protein (shown in Figure 7.3) and the WT protein [19]. Among the myriad of unnatural amino acids that can be incorporated site-specifically, p-AcPhe was selected as its unique keto group has been shown to react readily under mild conditions with hydroxylamine and hydrazyde compounds to form stable oxyme and hydrazone linkages, respectively as depicted in Figure 7.4 [20-21].

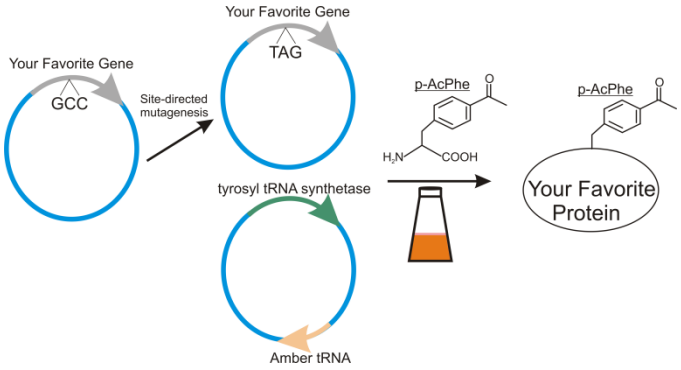


Figure 7.2: Strategy for genetic incorporation of unnatural amino acids to the protein of interest.

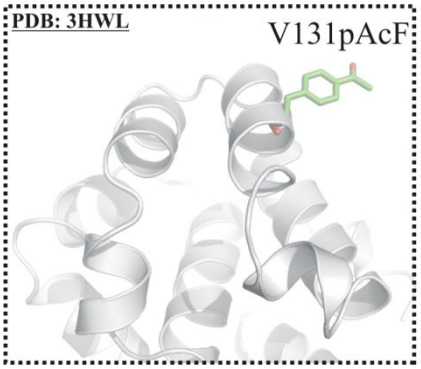


Figure 7.3: High-resolution structure of T4L bearing p-AcPhe unnatural amino acid at site 131 [19].

The main advantage of this strategy is the ability to design single or multiple attachment points in different regions of the protein. In this study, the extent of immobilization of T4L using bioorthogonal chemistry was compared to that coupled nonspecifically in CNBr-activated

sepharose and the conformation of the protein in the solid support was investigated by comparing the EPR spectra of spin-labeled T4L in the solid support to those in solution.

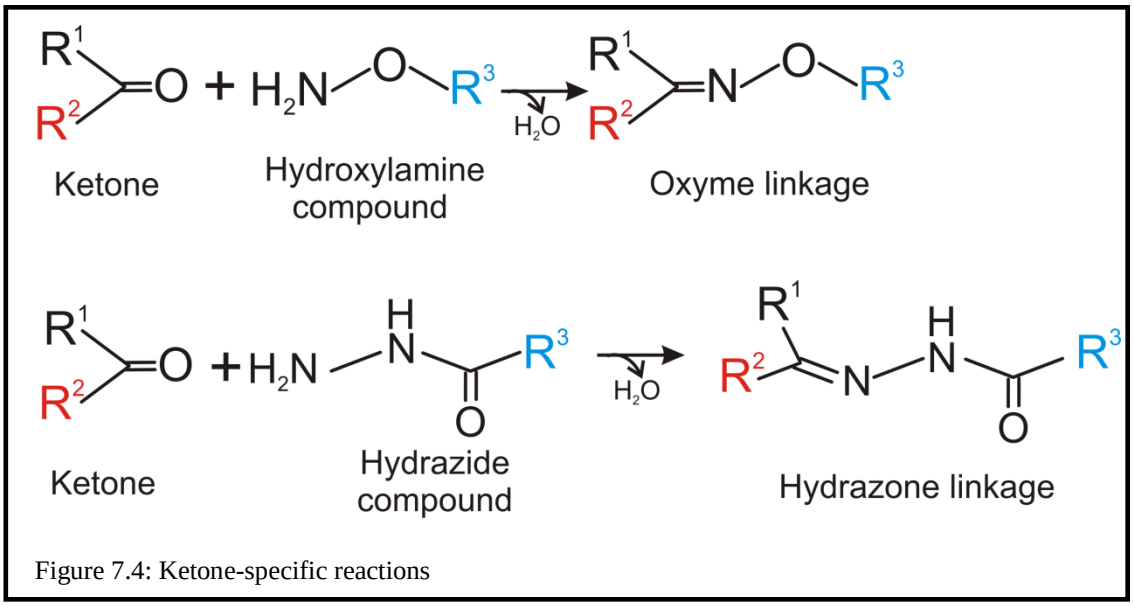


Figure 7.4: Ketone-specific reactions

7.3 Results

7.3.1 *Efficiency of biotinylation of T4L bearing p-AcPhe and immobilization of the biotinylated protein on streptavidin beads.* To modify T4L with a "tag" for site-specific immobilization, a series of T4L mutants bearing the p-AcPhe side chain at a single and multiple sites were generated. To immobilize the protein into a solid surface two strategies were employed based on the reactions shown in Figure 7.4. The first involved a direct attachment of the p-AcPhe mutants to commercially available hydrazide beads and the second involved noncovalent immobilization of T4L based on the biotin/streptavidin interaction. Oriented covalent immobilization of oxidized glycoproteins to hydrazide beads has been demonstrated and has the advantage of requiring mild buffer conditions [22-23], however the binding efficiency of T4L containing p-AcPhe to hydrazide beads was very low (≤ 11 %), even after 5 days of incubation at 37 °C (*see* Appendix). Given the low efficiency of coupling, the immobilization strategy *via* hydrazone-linkage was not pursued any further.

Due to the relatively high yield, stability, and bioothogonality of the oxyme-ligation (depicted in Figure 7.4), a "tag" containing a hydroxylamine group was used for modification of the p-AcPhe side chain and subsequent immobilization of the protein into a functionalized matrix. To this end, mutant proteins were incubated with a biotinylated hydroxylamine derivative to generate a biotin side chain site-specifically for immobilization on commercially available streptavidin beads (Figure 7.5).

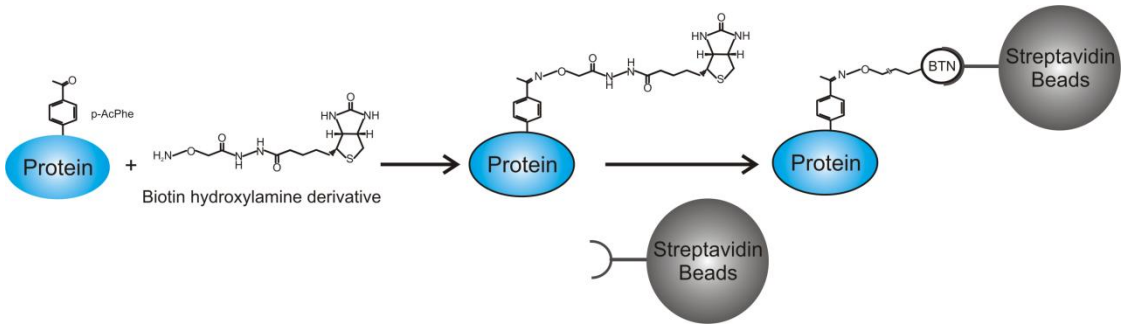


Figure 7.5: Site-specific immobilization strategy for protein bearing genetically encoded p-AcPhe unnatural amino acid.

This strategy takes advantage of the high affinity between biotin and streptavidin (K_d~10^{-15} M), which is one of the strongest noncovalent interactions known [24]. In addition, biotin/streptavidin interaction is highly stable even under harsh conditions (*e.g.*, low pH, 4 M GdnHCl, 60 °C) due to the high intrinsic stability of streptavidin [16,25]. Streptavidin exist as a homotetramer and contains four binding sites for biotin, two of which are in the same molecular surface as shown in Figure 7.6, which enables up to 2 potential binding sites for protein immobilization on the streptavidin beads.

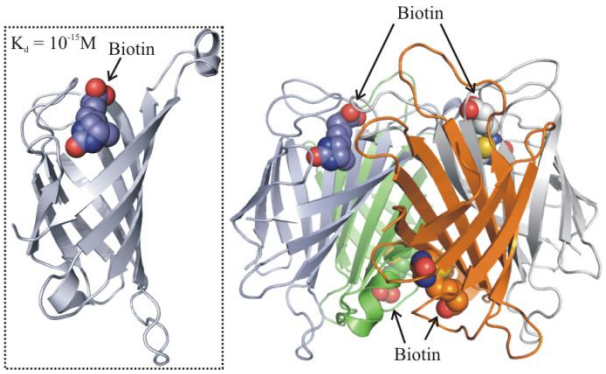


Figure 7.6: High-resolution structure of streptavidin bound to biotin. Each monomer is color coded for clarity. *Inset* The structure of one of the monomer is shown.

To evaluate the efficiency and specificity of biotinylation of T4L bearing p-AcPhe, untreated protein and protein incubated with 10-fold molar excess of biotinylated hydroxylamine were analyzed by MALDI-TOF mass spectrometry. A representative data set is shown in Figure 7.7A. The experimental mass of the unlabeled protein was 18679 Da, while the mass of the same protein incubated with biotin shifted by 314.48 Da, which is within 1 Da of the expected molecular mass of singly biotinylated T4L. A small spectral intensity corresponding to the mass

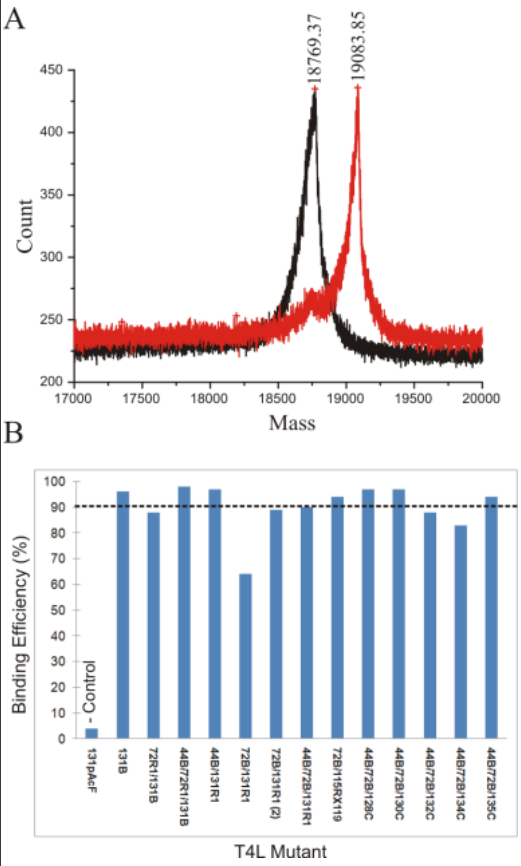


Figure 7.7: Biotinylation and site-specific immobilization on streptavidin beads is highly efficient. A) MALDI-TOF MS analysis of T4L V131p-AcPhe before (black) and after (red) incubation with biotinylated hydroxylamine. B) Binding efficiency of biotinylated T4L mutants to streptavidin beads. The dashed line indicates the average binding efficiency.

of unlabeled protein was observed in the protein incubated with biotinylated hydroxylamine, which corresponds to less than 10 % based on integration of the two observed peaks.

After biotinylation, the modified proteins were incubated with streptavidin beads and the coupling efficiency was evaluated by UV_{280} absorption. As a negative control, a protein bearing p-AcPhe was incubated with the streptavidin beads to test for nonspecific coupling. As shown in Figure 7.7B, singly and doubly biotinylated T4L mutants coupled to streptavidin resin at high efficiency (65 – 100%), while the nonbiotinylated protein coupled with less than 3 % efficiency. These results combined demonstrate that the biotinylation and subsequent immobilization of T4L proteins bearing p-AcPhe is specific and highly efficient.

7.3.2 *Study of the extent of immobilization of spin-labeled T4L mutants on streptavidin beads by EPR.* As discussed in previous chapters, the X-band EPR spectra of spin-labeled proteins have contributions from three dynamic modes, namely: internal motions of the nitroxide side chain (τ_{int}), overall rotational diffusion of the protein (τ_r), and backbone motions (τ_{bb}). To compare the contribution of τ_r to the EPR spectra of spin-labeled proteins immobilized on the streptavidin beads, the contributions from τ_{int} and τ_{bb} have to be minimized. To this end, a nitroxide derivative with constrained internal motions with respect to the protein (denoted R8 hereon) [26] was introduced at site D72 of T4 lysozyme, which has been shown to have minimal contributions from τ_{bb} [27]. To a good aproximation, the motion of the nitroxide for this label (shown in Figure 7.8A) at site 72 measures the rotational diffusion of the protein as a whole. For comparison purposes, a bifunctional spin label that generates a cross-linked side chain, which has been shown to have highly constrained internal motions [7], was also used to probe the extent of immobilization of site-specifically attached T4L (Figure 7.8B). To evaluate the relative contributions from τ_r to the EPR spectra of 72R8 coupled to streptavidin beads, the hyperfine splitting values (2Azz') were measured and compared to those values obtained for the same mutant in buffer, in 30 % w/w sucrose, and in immobilized protein *via* lysine residues in CNBr-activated sepharose beads. The latter immobilization strategy has been shown to be highly efficient in eliminating the contribution of τ_r to the spectra of spin-labeled proteins (*see* Chapter 3 and Fleissner et al. [7]), thus it serves as a good reference point to compare with site-specific attchment in streptavidin beads. The spectra of singly (131B) and doubly biotinylated (44B/131B) proteins were compared to one another to probe whether there is a difference in the 2Azz' values between proteins attached *via* one or two sites. It should be noted that the

biotinylation sites were strategically designed so that the spin label is oriented away from surface of the matrix.

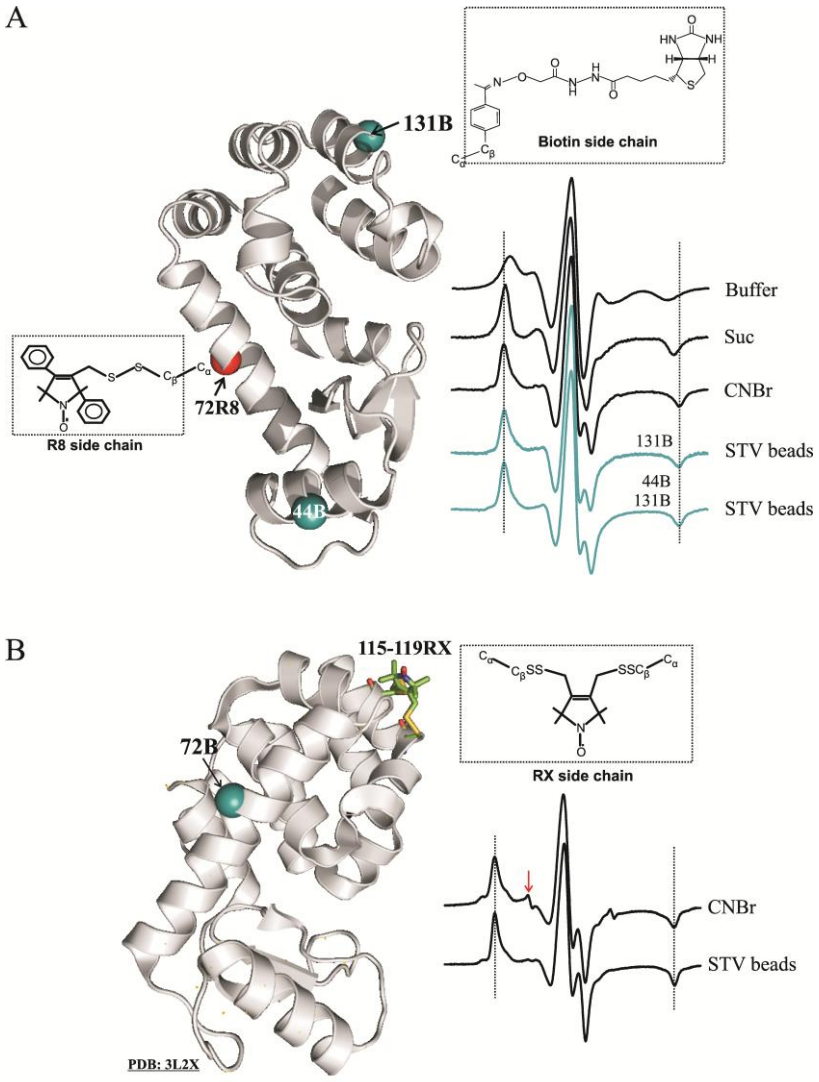


Figure 7.8: Site-specific coupling of biotinylated T4L to streptavidin beads effectively reduces the rotational diffusion of the protein. A) EPR spectra of protein bearing the R8 side chain at site 72 at the indicated conditions. The black traces are used to represent T4L D72R8 in the WT background. The teal traces indicate spin-labeled biotinylated protein. The biotinylation sites are indicated by teal spheres at the Cα. The dotted lines (2Azz') are used to help guide the eye. B) EPR spectra of T4L 115-119RX randomly attached on CNBr beads and site-specifically attached on streptavidin beads. The red arrow in the spectrum of 115-119RX in CNBr is used to indicate a small amount of free spin label in the sample.

As anticipated, the EPR spectrum of 72R8 in buffer showed the smaller 2Azz' values reflecting the highest contribution from τ_r (Figure 7.8A). The EPR spectrum in 30% w/w sucrose exhibits a 5.3 G increase in the hyperfine splitting value due to attenuation of the overall rotational diffusion of the protein. Interestingly, the EPR spectrum of 72R8 immobilized on CNBr-sepharose revealed a further increase in the hyperfine splitting value compared to that obtained in sucrose (Δ2Azz' ~ 2.7 G), which shows that 30% w/w sucrose does not completely eliminate the contribution from τ_r to the X-band EPR spectra of spin-labeled proteins. The EPR spectra of singly and doubly biotinylated protein immobilized site-specifically in streptavidin beads showed the same 2Azz' values, within the experimental error (*see* Appendix), both of which are quatitatively similar to that obtained in CNBr-activated sepharose. It should be emphasized that it is not possible to establish whether the 44B/72R8/131B mutant attached to the resin *via* the two biotinylated sites (44B and 131B).

Regardless of whether the doubly biotinylated protein attached through one or two sites, the data showed that, remarkably, a single site-specific attachment in T4L is sufficient to effectively immobilize the protein in the X-band EPR time scale to a similar extent to that in CNBr sepharose. To investigate the extent of immobilization using the bifnctional spin label, a different mutant bearing the p-AcPhe amino acid at site 72 was engineered. In this case, the bifunctional spin label was introduced at sites 115 and 119 to generate a cross-linked nitroxide side chain denoted RX (Figure 7.8B). As mentioned above, the RX side chain has been shown to exhibit highly constrained internal motions in spin-labeled proteins [7]. The mutant 115-119RX was selected because its structure has been determined at high-resolution, which revealed a highly localized nitroxide side chain as a result of its low τ_{int} [7]. The EPR spectrum of 115-119RX immobilized CNBr-sepharose is shown in Figure 7.8B and is similar to that published

previously [7]. The highly mobile component observed in the spectrum (*see* arrow) arises from a small amount of free spin (< 1 %) present in the sample. The spectrum of 72B/115-119RX immobilized on streptavidin beads is qualitatively similar to that in CNBr sepharose. Quantitatively, the 2Azz' values in both immobilization strategies are the same within the experimental error (*see* Appendix), which supports the contention that the protein is indeed immobilized to a similar extent to that in CNBr-sepharose beads.

To compare whether the extent of immobilization is site- and region- specific, three new mutants were designed, each bearing the R8 side chain at site 131. Sites D72 and S44 were selected for biotinylation as both of them are in the opposite surface to residue 131R8. As shown in Figure 7.9, the EPR spectrum of 131R8 in buffer showed the smaller 2Azz' values reflecting the highest contribution of τ_r. Interestingly the hyperfine splitting values in buffer for 131R8 is ~ 1.1 G smaller than that for 72R8. Since τ_r and τ_i are expected to have similar contributions to the EPR spectra at both sites, the differences in 2Azz' betweeen the two sites likely arise from higher contribution from fast backbone fluctuations in helix H as previously proposed by Columbus et al., 2001. Similarly to site 72R8, the 2Azz' values of 131R8 in sucrose are higher than that in buffer, but ~ 2.1 G smaller than in CNBr. The EPR spectra of 131R8 site-specifically attached to streptavidin beads *via* either site 44B or 72B are the same, within the experimental error, both of which are the same as the doubly biotinylated protein (44B/72B/131R8). In contrast to mutants 72R8 and 115-119RX, the 2Azz' values of the 131R8 mutants immobilized site-specifically are ~ 0.8 G smaller than that in CNBr. However it should be emphasized that two of the possible attachment points in the CNBr-sepharose resin are the lysine residues at sites 135 and 147, which are in close proximity to residue V131R8. Is is anticipated that attachment of the protein near the label can have an effect on the mobility of the nitroxide side chain and the

backbone motions of the helix, so it is possible that the larger 2Azz' values observed for residue 131R8 in CNBr arise from smaller contribution from τ_{bb} or from a more constrained motion of the R8 side chain. Additional evidence that supports this asumtion is shown below.

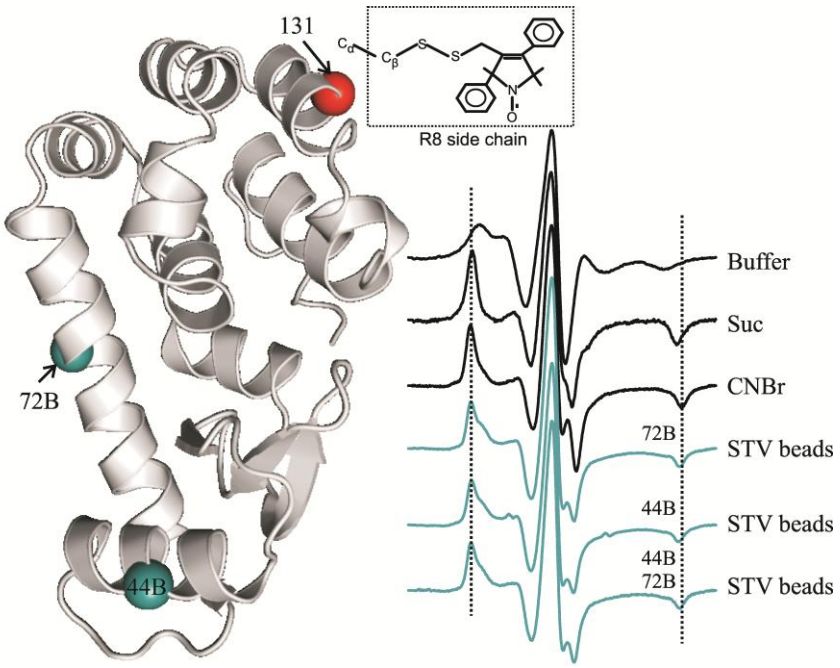


Figure 7.9: Site-specific coupling of biotinylated T4L to streptavidin beads *via* site 44 or 72 effectively reduces the rotational diffusion of the protein. EPR spectra of protein bearing the R8 side chain at site 131 under the indicated conditions. The black traces are used to represent T4L V131R8 in the WT background. The teal traces indicate spin-labeled biotinylated protein. The biotinylation sites are indicated by teal spheres at the Cα. The dotted lines (2Azz') are used to help guide the eye.

These results combined revealed that, independent of the site of attachment, single site-specific coupling of biotinylated T4L to streptavidin beads immobilizes the protein to a similar extent to that in CNBr on the X-band EPR time scale. It should be noted that the three sites selected for attachment 44B, 72B, and 131B are in relatively in ordered regions of T4L, thus it would be worth exploring in future studies whether the extent of immobilization is as efficient even if the biotinylation sites are placed in disordered protein regions.

7.3.3 *EPR spectra of R1-labeled proteins site-specifically immobilized on streptavidin beads:* In order to compare the spectra of R1 label protein immobilized site-specifically to those

randomly oriented and in viscous solution (30 % w/w sucrose), the spectra of residues 72R1 and 131R1 were recorded in the WT background (CNBr and sucrose) and in the singly and doubly biotinylated protein. For each R1 mutant, the EPR spectra of the protein in the WT and biotinylated background proteins in sucrose were the same (*see* Appendix), which shows that biotinylation does not affect the conformational properties of the protein. The spectra of 72R1 and 131R1 in buffer, sucrose, and CNBr are similar to those previously published [28-30]. As shown in Figure 7.10, the EPR spectra of residues 72R1 and 131R1 in sucrose exhibit a single dynamic component and clearly reflect a reduction in the mobility of the nitroxide side chain compared to that recorded in buffer due to the attenuation of the contribution from the overall rotary diffusion of the protein. As discussed in chapter 3 of this dissertation, the EPR spectra of residue 72R1 and 131R1 immobilized in the CNBr-activated sepharose reflect two components.

In each case, the more mobile component is similar to that observed in sucrose, while the small spectral intensity corresponding to a more immobilized component likely arises due to a fraction of the nitroxide being immobilized by direct interaction with the matrix near the site of attachment. It should be noted that T4L contains 13 native lysine residues randomly distributed on its protein surface, some of which are in the same surface as residues 72R1 and 131R1 (*see* Appendix), thus it is anticipated that attachment to the matrix *via* those lysine residues near the R1 side chain could give rise to the relatively immobile component due to tertiary interaction between the R1 side chain and the matrix. To test this idea, lysine-to-arginine substitutions were made near site 131R1 and the EPR spectrum of the variant was recorded in CNBr. Indeed, the EPR spectrum revealed a reduction in the intensity of the more immobilized component for the protein immobilized in CNBr-sepharose (*see* Appendix).

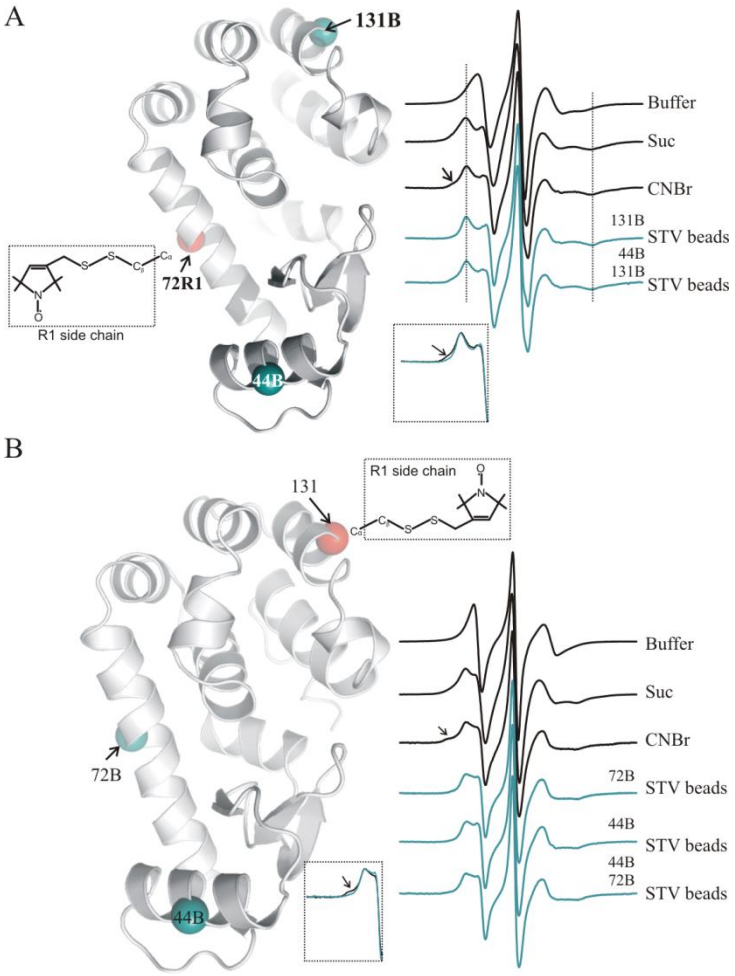


Figure 7.10: Site-specific attachment of spin-labeled protein eliminates tertiary interaction of the spin label side chain with the matrix. A) EPR spectra of protein bearing the R1 side chain at site 72 at the indicated conditions. B) EPR spectra of protein bearing the R1 side chain at site 131 at the indicated conditions. The black traces are used to represent spin-labeled protein in the WT background. The teal traces indicate spin-labeled biotinylated protein. The biotinylation sites are indicated by teal spheres at the Cα. The arrow in the spectrum in CNBr indicates a relatively immobile component. *Insets*: Low field spectra of immobilized proteins randomly (black) and site-specifically.

As discussed above, the biotinylation sites with site-specific coupling of R1 labeled proteins to streptavidin beads were selected so that the nitroxide side chain is oriented away from the matrix. As shown in Figure 7.10, the EPR spectra of 72R1 and 131R1 site-specifically attached to the streptavidin beads did not reflect the relatively immobile component observed in the CNBr-sepharose (*see* insets in panels A and B). These results combined support the contention that random immobilization of spin-labeled proteins can modulate the EPR spectra due to attachment near the R1 side chain and that such phenomenon can be effectively

eliminated by site-specific attachment of proteins. It should be noted that the spectral component observed for 72R1 and 131R1 immobilized site-specifically is qualitatively similar to those in sucrose and to the more mobile component in the CNBr-sepharose, which suggests that site-specific attachment of the spin-labeled protein to the streptavidin beads did not affect the conformational properties of the protein, at least, in the regions in which the nitroxide side chain was attached.

7.3.4 *Effect of site-specifically immobilization on streptavidin beads on the secondary structure of T4L.* Characterizing the structural properties of proteins immobilized in solid supports by biophysical methods such as x-ray crystallography and CD spectroscopy is impractical. As discussed in chapter 2 (*see* section 2.8.2), the topographical-dependent mobility of the R1 side chain provides a means for mapping secondary and tertiary folds using SDSL. To demonstrate the utility of SDSL for mapping the secondary structure of spin-labeled proteins immobilized site-specifically on solid supports, R1 side chain was introduced, one at a time, at sites 128 to 135, corresponding to residues within helix H. The EPR spectra were recorded for the protein immobilized site-specifically to streptavidin beads *via* residues S44B/D72B and compared to those obtained in sucrose. The EPR spectra of residues 128-135 in sucrose were published previously [29,31] and were reproduced here for comparison. As depicted in Figure 7.11A, the spectra of R1 mutants immobilized on streptavidin beads show the same periodic dependence in the mobility of the R1 side chain compared to that observed in solution. Most importantly, save for residue E128R1, the EPR spectra are qualitatively similar in solution and on streptavidin beads. For example, the EPR spectra of residues 131R1 and 132R1 in solution and in the streptavidin beads exhibit a single dynamic component of the R1 side chain reflecting anisotropic motion of the nitroxide, while residues A129R1, A130R1 and A134R1 reflect tertiary

interaction with the protein environment. In addition, the EPR spectrum of the buried residue L133R1 reflects a single dynamic component characteristic of a nitroxide with highly restricted motion. The small differences in the mobility of the nitroxide for E128R1 in sucrose and on streptavidin beads are not entirely understood, however in both cases the spectra reflect relatively high mobility of the R1 side chain. Quantitatively, solvent exposed residues (*ie.*, 128, 131, 132, and 135) reflect higher mobility than R1 at tertiary contact and buried sites in solution and in streptavidin beads as judged by the inverse central linewidth (ΔH^{-1}), which is a measure of mobility of the R1 side chain. As depicted in Figure 7.11C, the periodicity in R1 mobility observed for R1 in solution is maintained on streptavidin beads, which shows that site-specific attachment of T4L to the solid support does not affect the secondary structure of the protein. These findings are in agreement with previous EPR studies in T4L attached to a solid support *via* lysine residues, which suggest that immobilization of the protein in solid surfaces has little effect on the secondary structure of the protein [9]. However, as discussed above, nonspecific attachment near the R1 side chain can affect the mobility of the R1 side chain and hence the spectra of spin-labeled proteins, thus site-specific attachment provides a more effective means to immobilize the protein without the possibility of affecting the environment around the R1 side chain by the matrix.

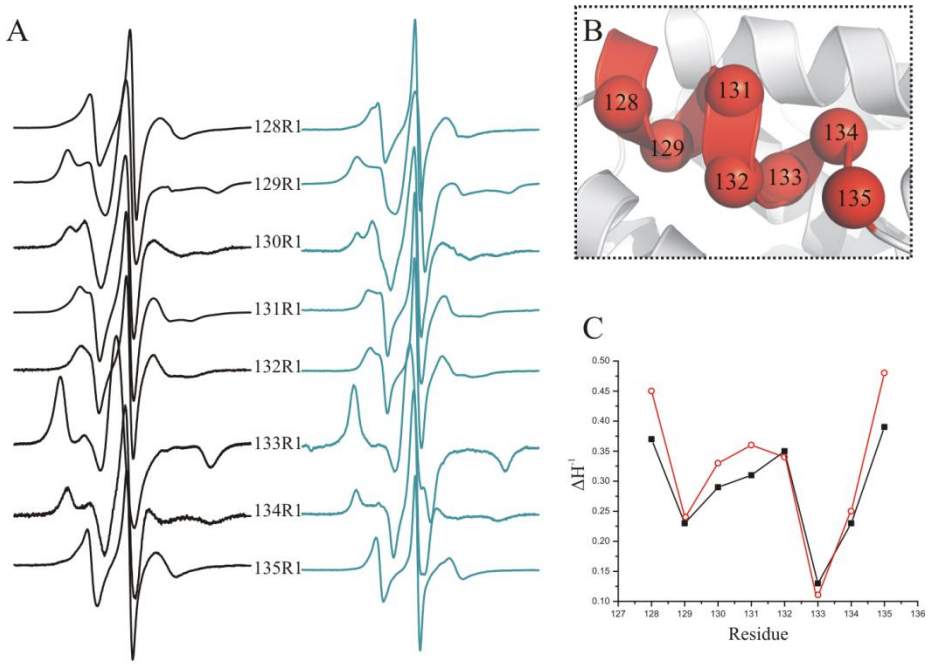


Figure 7.11: The secondary structure of T4L is not affected by site-specific attachment to streptavidin beads. A) EPR spectra of R1 residues introduced in helix H of T4L in solution (black) and immobilized in streptavidin beads (teal). B) Sites where R1 was introduced. B) Plot showing the inverse of the central linewidth per residue in solution (black) and in streptavidin beads (red).

7.4 Discussion

The primary goals of this project were to (1) develop a system that enables site-specific immobilization of spin-labeled proteins such that the nitroxide side chain and functionally important regions are oriented away from the molecular surface of the matrix and (2) study the structural and conformational properties of spin-labeled protein coupled site-specifically to the solid support. To this end, a genetically encoded unique bioorthogonal group (p-AcPhe) was introduced in T4 lysozyme (T4L) site-specifically, by using a genetically modified suppressor tRNA/aminoacyl-tRNA synthetase pair, for further modification and subsequent immobilization of the spin-labeled protein onto functionalized beads.

To selectively immobilize the protein on a solid surface in a site-specific manner, two strategies were employed based on the reactivity of hydrazides and hydroxylamines to the unique

keto group of the p-AcPhe side chain. The first involved a direct attachment of the p-AcPhe mutants to commercially available hydrazide beads and the second involved noncovalent immobilization of T4L based biotin/streptavidin interaction. Although the reaction between ketone to hydrazides has been shown to occur readily under mild buffer conditions (*i.e.*, PBS, pH 7), the coupling efficiency between the protein and hydrazide resin was inefficient (≤ 11 %). On the other hand, biotinylation and subsequent immobilization of protein bearing the p-AcPhe on streptavidin beads was highly efficient and specific as shown in Figure 7.7.

The extent of immobilization of the biotinylated protein to streptavidin beads was investigated by comparing the spectra of R8- and RX-labeled T4L mutants to those proteins immobilized nonspecifically in CNBr-activated sepharose and with those in viscous solutions (30 % w/w sucrose). As shown in Figure 7.8 and 7.9, coupling of biotinylated T4L to streptavidin beads reduced the contributions from τ_r more effectively than viscous sucrose solutions and to a similar extent compared to CNBr beads, which has been shown to completely immobilize the protein on the X-band EPR time scale [7]. It is remarkable that a single attachment point was sufficient to effectively immobilize the protein on the EPR time scale, thus future studies will not require the design multiple biotinylation sites for attachment points. The advantage of site-specific attachment over CNBr-sepharose and other nonspecific immobilization strategies is the ability to control the orientation of the protein in the matrix so that functionally important regions (*i.e.*, active sites, protein-protein interaction interfaces) and spin-label sites can be positioned away from the matrix. In principle, this should eliminate perturbation of the mobility of the R1 side chain from the matrix and maintain a *solution-like* activity of the protein. Indeed, our results demonstrate that judicious selection of sites for immobilization can effectively eliminate any perturbation of the matrix to the mobility of the R1

side chain. The important role of site-specific immobilization on the activity of protein has been demonstrated by several recent studies [12-14].

To investigate the structural and conformational properties of spin-labeled protein coupled site-specifically to the solid support, the EPR spectra of R1 mutants immobilized on streptavidin beads were compared to those in solution. The data shown in Figures 7.10 and 7.11 revealed that the EPR spectra of immobilized proteins were qualitatively similar to those obtained in solution. The periodicity observed for R1 at sites of helix H observed in solution is also maintained on the beads supporting the contention that site-specific immobilization on streptavidin beads does not affect the secondary structure of the protein. Future studies involving DEER spectroscopy will be done to probe the tertiary structure of the protein immobilized on beads. It should be emphasized that site-specific immobilization on solid supports is not expected to have the osmotic and volume effects that commonly used viscogens (*i.e.*, sucrose and Ficoll 70) have on regions undergoing conformational exchange (*see* Chapter 3).

The main advantages of the strategy used in this work for site-specific immobilization are the specificity, stability, and simplicity of the coupling; a spin-labeled protein bearing a unique keto group amino acid can be biotinylated with a commercially available biotin derivative and subsequently immobilized using commercially available streptavidin beads. In addition, the capture protein in the beads used here (streptavidin) has no native cysteine residues, which eliminates the possibility of background spin-labeling, thus enabling a system whereby the specific spin labels sites on the immobilized protein can be effectively reduced and re-labeled with different nitroxide side chains.

The main disadvantages of the strategy employed here are the need to introduce an unnatural amino acid to the protein of interest and the slow kinetics of the oxime-ligation

reaction near neutral pH [32]. An alternative strategy for site-specific immobilization to streptavidin beads that does not require the need for genetic introduction of p-AcPhe is based on the oxidation of the N-terminal amino group to ketone by pyridoxal 5'-phosphate [33]. Although this reaction can be used to potentially add a keto group site-specifically at the N-terminal residue of the protein of interest, the reaction yields can vary depending on the identity of the N-terminal residue and, apparently, the reaction is incompatible with proteins possessing N-terminal serine, threonine, thryptophan, and proline [33]. In addition, this strategy would not be useful in cases where the N-terminal region is functionally important. However, for proteins expressed in systems for which genetic incorporation of unnatural amino acids is not practical, the N-terminal oxidation of amino groups with pyridoxal 5'-phosphate can be an alternative approach worth exploring.

In this work, the oxyme ligation reaction was done at pH 4.5 due to the intrinsically faster reaction kinetics at mildly acidic pH. It should be noted that for some proteins, especially for highly unstable proteins, the oxyme ligation reaction at pH 4.5 is impractical due to the possibility of aggregation. For such cases, the oxyme ligation can be accelerated at neutral pH by using aniline, which has been shown to serve as a organocatalyst under anaerobic conditions [34].

It is realized that alternative strategies for site-specific immobilization using other bioorthogonal groups can be used for EPR spectroscopy. One of the strategies that will be used in future studies is based on the "click" chemistry between azide and alkyne chemical functionalities [35] given that unnatural amino acids containing the bioorthogonal azide side chain have been successfully incorporated in proteins expressed in *E. coli* and yeast [36-37]. The classic "click" chemistry reaction between azide and alkyne chemical functionalities involves a

copper (I) catalyzed cycloaddition forming 1,3-triazole product[3]. Due to the bioincompatibility of the copper (I), Bertozzi and coworkers synthesized a series of strained cyclooctyne molecules that reacted readily with azide groups without the need of copper [38-39]. Future studies involving the reaction between azide-containing spin-labeled proteins with strained cyclooctyne-functionalized resins (Figure 7.12) will be done and compared to the studies presented here in terms of binding efficiencies and extent of immobilization. The advantages of the azide-strained cyclooctyne reaction over the strategy employed in this work is the ability to do the reaction at neutral pH and the possibility for direct immobilization to the matrix without the need of a intermediate capture protein.

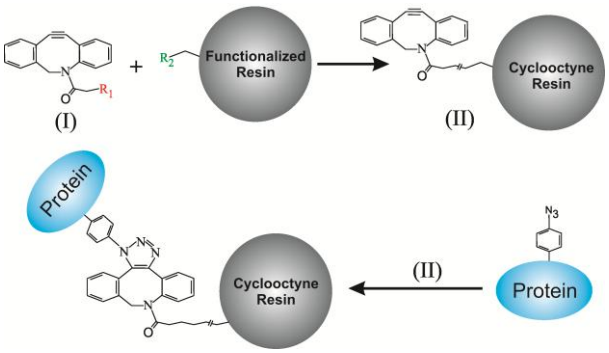


Figure 7.12: General strategy for site-specific immobilization of spin-labeled protein bearing pAcAz unnatural amino acid. Dibenzylcyclooctyne derivative (I) can be attached to a functionalized resin to make a cyclooctyne resin (II), which will be useful for *direct* attachment of proteins bearing pAcAz at pH 7.

These new methodologies will not only eliminate the contribution from τ_r to the X-band EPR spectra of small proteins without having the potential effects of commonly used viscogens (*i.e.*, sucrose or Ficoll; *see* Chapter 3), but ultimately enable a flow system whereby the structural and dynamical effects of other molecules (*e.g.*, ligands, substrates, proteins, osmolytes) can be assayed sequentially on the same spin-labeled protein.

[3] The reaction can occur in the absence of copper (I), but at a low efficiency (Rostovtsev et al., 2002; Tornoe et al., 2002).

7.5 Appendix

7.5.1 Binding efficiency of T4L bearing p-AcPhe at site 131 to hydrazide beads.

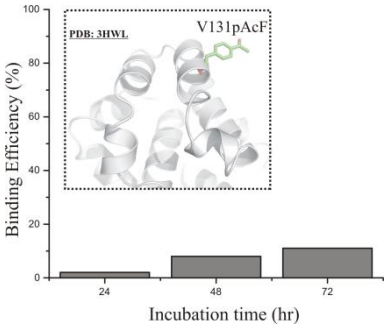


Figure 7.A.1: Binding efficiency of T4L V131p-AcPhe to Affi-gel hydrazide beads. 0.5 mg of protein was incubated with hydrazide beads in PBS buffer at pH 4.5. The binding efficiency was evaluated by UV_{280}.

Table 7.A.1: Hyperfine splitting (2Azz') of spin-labeled T4L mutants under the indicated conditions.

Mutant	Media	2Azz' (Gauss)	Error
D72R8	Buffer	60.52	0.02
	Sucrose	65.84	0.04
	CNBr	68.52	0.06
	STV 131B	68.32	0.06
	STV 44B/131B	68.20	0.06
	Frozen	69.62	0.06
115-119RX	CNBr	70.15	0.04
	STV 72B	70.26	0.10
	Frozen	71.00	0.10
V131R8	Buffer	59.38	0.10
	Sucrose	66.56	0.06
	CNBr	68.68	0.04
	STV 44B	67.80	0.04
	STV 72B	67.86	0.06
	STV 44B/72B	67.76	0.06
	Frozen	71.00	0.40
44R8	Suc	64.76	0.04
	STV 72B	66.26	0.06

7.5.2 Effect of biotinylation on the spectra of residues 72R1 and 131R1 in solution.

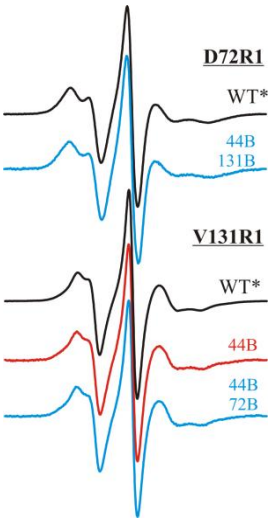


Figure 7.A.2: Biotinylation of T4L has no effect on the EPR spectra of residues D72R1 and V131R1 in solution. *Top panel*: EPR spectra of 72R1 in the WT background (WT*) and the doubly biotinylated protein in solution. *Bottom panel*: EPR spectra of 131R1 in the WT background (WT*) and the singly and doubly biotinylated proteins in solution.

7.5.3 Distribution of lysine residues in T4 lysozyme.

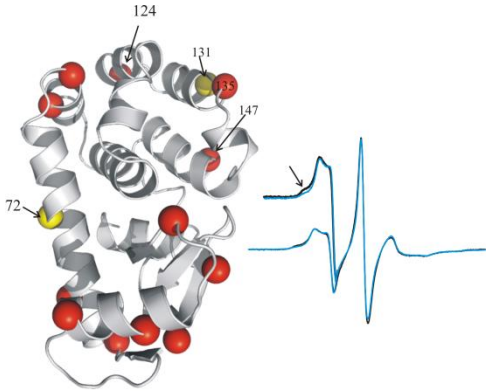


Figure 7.A.3: Random orientation to the solid support affect the EPR spectra of spin-labeled proteins. *Left panel*: Structure of T4L showing the sites for attachments in CNBr (red spheres). Residues 72 and 131 are indicated in yellow. *Right panel*: EPR spectra of 131R1 in the WT background (black) and K124R/K135R/K147R mutant (blue) in CNBr-beads. The low field regions of the spectra are amplified for clarity. The arrow highlights a reduction of the spectral intensity corresponding to a relatively immobile component arising due to attachment to the matrix near the R1 side chain.

7.6 References

1. Seong, S., and Choi, C. 2003. Fabrication and application of protein microarray. *Proteomics*, 3:2176-2189.
2. Bilitewski, U. 2006. Protein-sensing assays formats and devices. *Anal. Chim. Acta*, 568:232-247.
3. English, B. P., Min, W., van Oijen, A.M., Lee, K.T., Luo, G., Sun, H., Cherayil, B.J., Kou, S.C., and Xie, X.S. 2005. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat. Chem. Biol.* 2:87-94.
4. Monsan, P., and Combes, D. 1988. Enzyme stabilization by immobilization. *Meth. Enzym*. 137:584-598.
5. Horowitz, P., and Bowman, S. 1987. Reversible thermal denaturation of immobilized rhodanase. *J. Biol. Chem*. 262:5587-5591.
6. Kurimoto, E., Harada, T., Akiyama, A., Sakai, T., and Kato, K. 2001. *In-vitro* refolding of porcine pepsin immobilized on agarose beads. *J. Biochem*. 130:295-297.
7. Fleissner, M.R., Bridges, M.D., Brooks, E.K., Cascio, D., Kálai, T., Hideg, K., and Hubbell, W.L. 2011. Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* 108:16241-16246.
8. Berliner, E.J., Miller, S.T., Rosa, U.Y., and Royer, G.P. 1973. An ESR study of the active-site conformations of free and immobilized trypsin. *Biochim. Biophys. Acta* 315:195-199.

9. Guo, Z. 2003. Correlation of spin label side-chain dynamics with protein structure: Studies of T4 lysozyme with site-directed mutagenesis and X-ray crystallography. PhD thesis, University of California, Los Angeles
10. Butler, J.E., Ni, L., Brown, W.R., Joshi, K.S., Chang, J., Rosenberg, B., and Voss, E.W. 1993. The immunochemistry of sandwich ELISAs-VI. More than 90 % of monoclonal and 75 % polyclonal anti-fluorescyl capture antibodies (CAbs) are denatured by passive adsorption. *Mol. Immunol*. 30:1165-1175.
11. Lu, J.R., Su, T.J., Thirtle, P.M., Thomas, R.K., Rennie, A.R., and Cubitt, R. 1998. The denaturation of lysozyme layers adsorbed at the hydrophobic solid/liquid surface studied by neutron reflection. *J. Colloid Inter. Sci*., 206:212-223.
12. Peluso, P., Wilson, D.S., Do, D., Tran, H., Venkatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, N., Phelan, M., Witte, K., Jung, L.S., Wagner, P., and Nock, S. 2003. Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal. Biochem*. 312:113-124.
13. Cha, T., Guo, A., and Zhu, X.Y. 2005. Enzymatic activity on a chip: The critical role of protein orientation. *Proteomics*, 5:416-419.
14. Wong, L.S., Thirlway, J., and Micklefield, J. 2008. Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase. *J. Am. Chem. Soc*. 130:12456-12464.
15. Nieba, L., Nieba-Axmanm, S.E., Persson, A., Hämäläinen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A.F., and Plückthun, A. 1997. BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Anal. Biochem*. 252:217-228.

16. Lesaicherre, M.L., Lue, R.Y.P., Chen, G.Y.J., Zhu, Q., and Yao, S.Q. 2002. Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Am. Chem. Soc.* 124:8768-8769.

17. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. 2001. Expanding the genetic code of Escherichia coli. *Science* 292:498-500.

18. Xie, J., and Schultz, P.E. 2005. Adding amino acids to the genetic repertoire. *Curr. Opin. Chem. Biol*. 9:548-554.

19. Fleissner, M.R., Brustad, E.M., Kálai, T., Altenbach, C., Cascio, D., Peters, F.B., Hideg, K., Peuker, S., Schultz, P.G., Hubbell. W.L. 2009. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci.* 106:21637-21642.

20. Zeng, H., Jianming, X., and Schultz, P.G. 2006. Genetic introduction of a diketone-containing amino acids into proteins. *Bioorg. Med. Chem. Lett.* 16:5356-5359.

21. Brustad, E.M., Lemke, E.A., Schultz, P.G., Deniz, A.A. 2008. A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *J. Am. Chem. Soc.* 130:17664–17665.

22. Hornsey, V.S., Prowse, C.V., and Pepper, D.S.1986. Reductive amination for solid-phase coupling of protein. A practical alternative to cyanogen bromide. *J. Imm. Methods* 93:83-88.

23. Domen, P., Nevens, J.R., Mallia, K., Hermanson, G., Klenk, D.C. (1990). Site directed immobilization of proteins. *J. Chromatogr.* 510:293-302.

24. Green, N.M. 1975. Avidin. *Adv. Prot. Chem.* 29:85-133.

25. González, M., Argaraña, C.E., and Fidelio, G.D. 1999. Extremely high thermal stability of streptavidin and avidin upon biotin binding. *Biomol. Eng*. 16:67-72.

26. Sale, K., Sár, C., Sharp, K.A., Hideg, K., and Fajer, P.G. 2002. Structural determinants of spin label immobilization and orientation: a Monte Carlo minimization approach. *J. Magn. Res*. 156:104-112.

27. Columbus, L., Kálai, T., Jeko, J., Hideg, K., and Hubbell, W.L. 2001. Molecular motion of spin labeled side chains in alpha-helices:analysis by variation of side chain structure. *Biochemistry* 40:3828-3846.

28. Liang, Z., Lou, Y., Freed, J., Columbus, L., and Hubbell, W.L. 2004. A multifrequency electron spin resonance study of T4 lysozyme dynamics using the slowly relaxing local structure model. *J. Phys. Chem. B.* 108:17649-17659.

29. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35:7692-7704.

30. López, C.J., Fleissner, M.R., Guo, Z., Kusnetzow, A.K., and Hubbell, W.L. 2009. Osmolyte-perturbation reveals conformational equilibria in spin-labeled proteins. *Protein Sci.*, 18:1637-1652.

31. Pyka, J., Ilnicki, J., Altenbach, C., Hubbell, W.L., and Froncisz, W. 2005. Accessibility and dynamics of nitroxide side chains in t4 lysozyme measured by saturation recovery EPR. *Biophys. J*. 89:2059-2068.

32. Jencks, W.P. 1959. Studies on the mechanism of oxime abd semicarbazone formation. *J. Am. Chem. Soc.* 81: 475-481.

33. Gilmore, J.M., Scheck, R.A., Esser-Kahn, A.P., Joshi, N.S., Francis, M.B. 2006. N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed*. 45:5307-5311.

34. Dirksen, A., Hackeng, T.M., Dawson, P.E. 2006. Nucleophilic catalysis of oxime ligation. *Angew Chem-Int Edit.* 45:7581–7584.

35. Best, M.D. 2009. Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling biological molecules. *Biochemistry* 48:6571-6584.

36. Deiters, A., Cropp, T.A. Summerer, D., Mukherji, M., and Schultz, P.G. 2004. Site-specific PEGylation of proteins containing unnatural amino acids. *Bioorg. Med. Chem. Lett.* 14:5743-5745.

37. Tsao, M., Tian, F., and Schiltz, P.G. 2005. The Staudinger modification of proteins containing azido amino acids. *ChemBioChem* 6:2147-2149.

38. Agard, N.J., Prescher, J.A., and Bertozzi, C.R. 2004. A strain promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* 126:15046-15047.

39. Sletten, E.M., and Bertozzi, C.R. 2011. From mechanism to mouse: a tale of two bioorthogonal reactions. *Acc. Chem. Res*. 44:666-676.

Chapter 8: Materials and Methods

8.1 Construction and expression of T4 lysozyme mutants

The construction of the cysteine mutants A41C, K48C, D72C, R76C, R80C, N81C, A82C, T109C, T115C, V131C, A134C, and T151C has been described elsewhere [1-4]. New mutants (Tables 8-1 and 8-2) were generated using the overlap extension method [5] or QuickChange site-directed mutagenesis method (Stratagene) in the pseudo-wild type T4 lysozyme construct containing the substitutions C54T and C97A. All mutations were confirmed *in-house* using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, CA) or by Genewiz Inc. (South Plainfield, NJ). The mutants in the pHSe5 vector and pET11a vector were transformed into *E. coli* BL21 and BL21(DE3) expression strains (Stratagene), respectively and expressed as described previously [3].

Table 8.1: T4 lysozyme mutants in pHSe5 vector

1. L84A/T109C	6. V131C/L133G
2. G107A/T109C	7. M102K/V131C/L133D‡
3. T109C/G110A	8. V131C/L133D
4. T109C/G110A/G113A	9. L99A/N140C
5. V131C/L133A	11. N140C

Table 8.2: T4 lysozyme mutants in pET11a vector

1. L46A/K48C	14. L99A/V131C/F153A‡	27. W138A/N140C/T151C‡
2. D72C/A74M	15. M102A/M106A/V131C	28. N140C/T151C
3. D72C/I100A	16. L121A/V131C	29. A98V/T151C
4. D72C/F104A	17. L121A/V131C/L133A‡	30. L99A/T151C/F153A‡
5. D72C/L133G‡	18. M102K/V131C‡	31. L133G/T151C‡
6. F104C	19. A129V/V131C	32. I150A/T151C
7. M102K/T109C‡	20. A129M/V131C	33. T151C/F153A
8. M102A/M106A/T109C	21. A129V/V131C/L133F	34. N68C/L121A/V131C/L133A‡
9. F104A/T109C	22. V131C/L133A	35. D89C/L121A/V131C/L133A‡
10. T109C/V111A	23. V131C/L133F	36. T109C/L121A/V131C/L133A‡
11. T109C/L133G‡	24. V131C/I150A	37. N68C/W138A/N140C‡
12. T109C/L121A	25. V131C/F153A	38. D89C/N140C
13. T109C/L121A/L133A‡	26. W138A/N140C‡	39. D89C/W138A/N140C‡

‡ Mutants were expressed and purified from inclusion bodies (*see text below*)

8.2 Purification of T4 Lysozyme mutants

Soluble T4L mutants were purified as described by Mchaourab et al. [3]. The overall yield for some of the mutants (*see* * on Table 8-2) was very low (< 1 mg). Interestingly, for those mutants a significant amount of overexpressed protein was observed in the insoluble fraction of the crude extract (inclusion bodies). Isolation of T4 lysozyme from inclusion bodies was done as described by Liu et al. [6] with some modifications. After sonication of the crude extract and subsequent centrifugation step, the pellet containing the inclusion bodies was washed with 20 ml of lysis buffer containing 2 % of Triton X-100. The mixture was sonicated (on ice) for 5 minutes. The resuspended pellet was incubated for 2 hours at 4 ºC followed by centrifugation for 15 minutes at 15,000 rpm. The pellet was collected and resuspended in 25 mL of 50 mM Tris-Cl containing 2.5 % of n-octyl-β-D-glucosepyranoside (wash #2). The mixture was sonicated for 5 minutes followed by incubation for, at least, 3 hours at 4 ºC. After another centrifugation step, the washed inclusion body pellet was resuspended in 20 mL of 10 mM glycine and the solution was sonicated for 5 minutes. Solid urea was added under intense stirring to a final concentration of 4 M, and the solution was quickly titrated to pH 3.0 with phosphoric acid to resolubilize the inclusion body. The resolubilized inclusion body was then subjected to a final centrifugation step to remove any remaining insoluble material and the supernatant was dialyzed overnight against 2 L of a buffer consisting of 10 mM sodium citrate, 1 mM β-cyclodextrin, 15 % glycerol, 5 mM DTT pH 3.0. Another one-day dialysis was done against 2 L of the same buffer, except the pH was adjusted to 5. After removal of precipitate, the refolded T4 lysozyme protein was loaded into a Hi Trap CM FF ion exchange column (GE Healthcare) equilibrated with a solution consisting of 10 mM sodium citrate, 5 mM DTT pH 5.0. The T4 lysozyme fraction was eluted with a linear gradient (0 – 400 mM) of NaCl. The T4 lysozyme fraction eluted at 240 mM NaCl (conductivity

~ 25 mS/cm). Sample purity was greater than 95 % as judge by SDS-PAGE electrophoresis. To ensure that there were no misfolded soluble-oligomeric species in the sample, the T4 lysozyme fraction was injected into a Superdex 75 (GE Healthcare) gel filtration column. In all cases there was a single peak with a retention volume of ~ 15.1 mL corresponding to monomeric T4 lysozyme. Figure 8.1 shows a representative SDS-PAGE gel of the fractions and chromatographic profiles from purification of T4 lysozyme mutants from inclusion bodies.

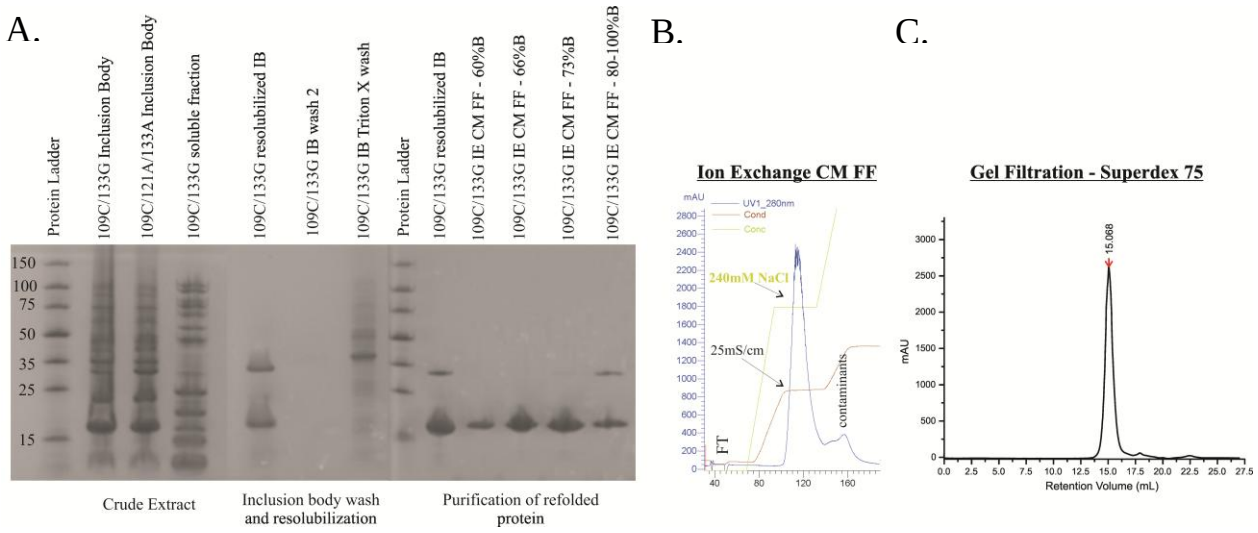


Figure 8.1: Isolation of T4 lysozyme mutants from inclusion bodies. A) SDS-PAGE of fractions collected during the purification of T4 lysozyme. B) Representative ion exchange chromatogram. The NaCl concentration at which the T4 lysozyme fraction eluted is indicated. C) Gel filtration profile of a mutant injected into a Superdex 75 gel filtration column. The retention volume is indicated at the top of the peak.

8.3 Spin-labeling of T4 lysozyme mutants

DTT was removed by using a Hi Trap Desalting column (GE Healthcare) equilibrated with buffer consisting of 50 mM MOPS, 25 mM NaCl pH 6.8. The eluted protein was incubated immediately with 10-fold molar excess of methanethiosulfonate nitroxide reagent. The sample was incubated overnight at 4 ºC. Excess spin label sample was removed with the same desalting column. The spin-labeled protein was concentrated (200 μM - 1 mM) with Amicon Ultra 10 kDa MWCO (Millipore).

8.4 Construction and expression of sperm whale myoglobin mutants

The plasmid pET17b (Novagen, Madison WI) carrying the WT gene of sperm whale myoglobin was kindly provided by Steven Boxer (Stanford University, Ca). The individual cysteine substitutions were introduced by using the QuickChange site-directed mutagenesis method (Stratagene). All mutations were confirmed by DNA sequencing. Mutant plasmid were transformed into the expression cells *E. coli* BL21(DE3). Cell cultures were grown to a OD_{600nm} of 1.0 AU and protein expression was induced by addition of 1 mM IPTG. For expression of apomyoglobin, the protein was forced to overexpress in inclusion bodies by incubating the culture at 42 ºC for 3-4 hours after addition of IPTG. Expression of holomyoglobin was done at 37 ºC.

8.5 Purification of holo- and apo-myoglobin mutants

Holomyoglobin mutants were purified as described by Springer and Sligar [7]. Myoglobin in the **apo** form was purified as previously described [8] with some modifications. Briefly, the washed inclusion body pellet was resolubilized in a solution containing 40%-50%

acetonitrile/0.1% trifluoroacetic acid followed by sonication. The resolubilized inclusion bodies were subjected to centrifugation at 15,000 rpm for 15 minutes to remove any insoluble material. The protein was purified by reverse-phase HPLC (Shimadzu Biotech) using a linear gradient (20 - 100 %) of an acetonitrile solution containing 0.1% trifluoroacetic acid as a mobile phase (see Figure 8.2). Following HPLC purification, the protein fractions containing myoglobin (as judged by SDS-PAGE) were lyophilized and stored at 4 °C until needed.

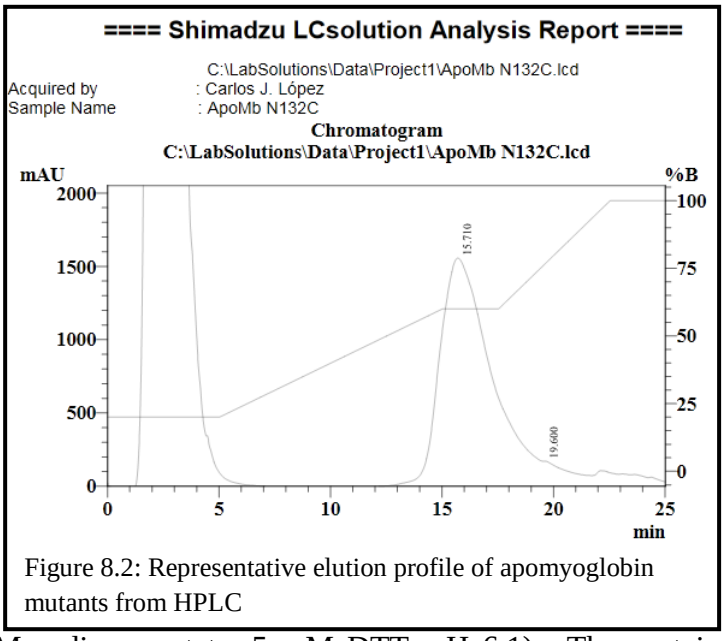


Figure 8.2: Representative elution profile of apomyoglobin mutants from HPLC

5 – 10 mg of lyophilized protein was resuspended in a freshly made resolubilization solution (10 mM sodium acetate, 6 M Urea, 5 mM DTT, pH 6.1) and the protein solution was allowed to equilibrate at 4 °C for at least 30 minutes. *In vitro* refolding of myoglobin was achieved by a rapid 10-fold dilution of the urea-denatured protein into the refolding buffer (10 mM sodium acetate, 5 mM DTT, pH 6.1). The protein concentration during refolding was maintained at or below 1 mg/ml to minimized aggregation problems. For some mutants, higher yields from *in vitro* refolding were achieved by overnight dialysis of the urea-denatured protein into the refolding buffer instead of rapid-dilution.

8.6 Spin-labeling of myoglobin mutants

DTT was removed using a Hi-Trap desalting column (GE Healthcare) and the eluted protein was immediately incubated with 10-fold molar excess of methanetiosulfonate reagent and allowed to react overnight at 4 ºC. Excess spin label was removed and the buffer was exchanged to gel filtration buffer (10 mM sodium acetate, 100 mM NaCl, pH 6.1) on the Hi-Trap desalting column. To separate the correctly folded monomeric species from any soluble missfolded aggregates, the spin-labeled protein was injected into a Superdex 75 column (GE Healthcare) equilibrated with 10 mM sodium acetate, 100 mM NaCl, pH 6.1. The missfolded and correctly folded fractions eluted at V_e of ~ 7.6 ml and ~ 12.6 ml (± 0.1 ml), respectively (Figure 8.3). The monomeric protein at this point was > 95% pure (as judged by SDS-PAGE). For EPR studies of myoglobin in the $\mathbf{N_{Apo}}$ state, the buffer was exchange to 10 mM sodium acetate, pH 6.1 and the proteins were concentrated to ~ 400 M with Amicon Ultra concentrator (10 kDa MWCO: Millipore) using an extinction coefficient of 15,400 $M^{-1}cm^{-1}$.

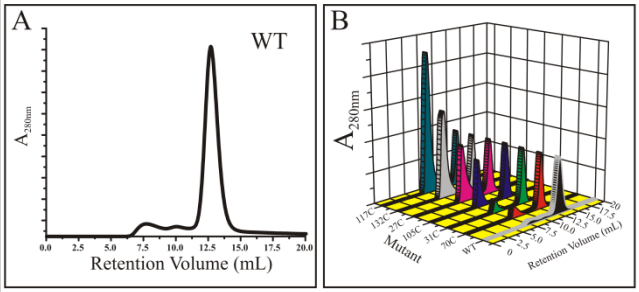


Figure 8.3: Representative elution profiles of wild-type and cysteine mutants of apomyoglobin from gel filtration chromatography.

For studies of myoglobin in the $\mathbf{I_{MG}}$ state, the buffer solution was exchange by repeated washes (5 * 15 ml each) with buffer containing 10 mM sodium acetate, 10% ethanol, pH 4.0 using the amicon concentrator. To populate the $\mathbf{U_{Acid}}$ state, the protein solution was exchange to a 10 mM acetic acid/HCl, pH 2.3.

8.7 Reconstitution of apomyoglobin mutants with bovine hemin

For EPR studies of myoglobin in the $\mathbf{N_{Holo}}$ state, the apo protein was incubated overnight at 4 ºC with a 1.2-fold molar excess of bovine hemin (Sigma) in 10 mM NaOH. Excess of heme was removed using the Amicon ultra concentrator and the binding of heme was assessed via a UV-Vis scan. The holo-protein reconstituted in this manner auto-oxidizes into the aquomet form (Fe^{+3}–H_2O) as judged by the absorption maxima at 409 nm [9]. The A_{409}/A_{280} ratio was measured to ensure that the protein was fully reconstituted (A_{409}/A_{280} ~ 5.16). The aquomet myoglobin samples were concentrated to ~ 400 μM by measuring the absorbance at 409 nm using an extinction coefficient of 157,000 $M^{-1}cm^{-1}$. Reconstitution of apo myoglobin with the diamagnetic zinc(II)-porphyrin (Sigma-Aldrich) was done as previously described [10]. The absorption maxima of the soret band for the myoglobin samples reconstituted with zinc(II)-porphyrin was 427 nm.

8.8 Construction and expression of rat intestinal fatty acid-binding protein

The plasmid pET11d (Novagen, Madison, WI) carrying the wild-type rat intestinal fatty acid binding protein was a generous gift from Alan K. Kleinfeld (Torrey Pines Institute for Molecular Studies, San Diego, CA). The cysteine substitutions at consecutive sites 14-35 and 55-56 were generated by using the QuickChange site-directed mutagenesis method (Stratagene); cysteine mutations for sites 73-75 were generated using the overlap extension PCR method [5]. All mutations were confirmed by DNA sequencing. Mutant plasmids for rIFABP were transformed into the expression strain *E. coli* BL21(DE3). *E. coli* cells containing the rIFABP mutant constructs were inoculated in 20 ml of LB containing 100 μg/mL of ampicillin and the cells were allowed to grow overnight at 37 ˚C. The overnight cultures were transferred to 1L of

fresh LB containing ampicilli and incubated at 37 ˚C until OD_{600} = 1, at which point the expression was induced with 1 mM of IPTG, and incubation continued at 37 ˚C for 4 hours. The cells were harvested by centrifugation at 6,000 rpm and the cell pellets were resuspended in lysis buffer (50 mM Tris-Cl, 1 mM EDTA, 0.05 % NaN_3 at pH 8.0). The crude extract was then frozen at -20 ˚C.

8.9 Purification of rat intestinal fatty acid-binding protein

Purification of rIFABP mutants was done as described by Lowe et al. [11] with some modifications. Following the ammonium fractionation step, the protein solution was loaded into HiTrap SP HP column and Hi Trap Q FF columns connected in series (GE Healthcare) that had been pre-equilibrated with low salt buffer (20 mM phosphate buffer, 1 mM EDTA, 5 mM DTT at pH 7.3). The rIFABP fraction was collected under isocratic conditions since the protein does not bind to any of the two columns used in this step. The rIFABP fractions were pooled and injected into a Superdex 75 gel filtration column (GE Healthcare). Sample purity was greater than 95 % as judged by SDS-PAGE electrophoresis.

8.10 Spin-labeling of rat intestinal fatty acid-binding protein

For spin labeling, DTT was removed by using a Hi Trap desalting column (GE Healthcare) equilibrated with spin labeling buffer (20 mM phosphate buffer, 1 mM EDTA at pH 7.3). The eluted protein was incubated immediately with 10-fold molar excess of methanethiosulfonate reagent and allowed to react overnight at 4 ˚C. Excess spin labeling reagent was removed with the same desalting column. The spin label protein was concentrated with Amicon Ultra 10,000 MWCO. At this point, the protein contains endogenous fatty acids

[11]. In order to generate the apo form of rIFABP, the protein was delipidated with Lipidex 1000 as described by Glatz and Veerkamp [12].

8.11 Protein immobilization on CNBr sepharose 4B beads

CNBr-acivated Sepharose 4B was purchased from GE Healthcare. The lyophilized media was suspended and washed several times with 1 mM HCl pH 2.0 following manufacturer instructions. HCl was removed with several medium volumes of coupling buffer (0.1 M $NaHCO_3$, 0.5 M NaCl pH 8.3) except for apo-Myb, where 10mM sodium acetate pH 6.1 was used instead. The spin labeled proteins (2-4 mg) were diluted (V_f ~ 400 μl) in the appropriate coupling buffer and added to the beads in a 2:1 protein solution / medium ratio. The coupling reaction was incubated overnight at 4 ˚C. Any uncoupled protein was eluted with coupling buffer and the coupling efficiency was assessed by UV_{280} absorption. The coupling reaction typically yielded an estimated adsorption of 10-20 mg of protein / ml of medium. The immobilized protein was equilibrated with the appropriate buffer selected for EPR studies.

8.12 Protein immobilization using glutaraldehyde crosslinking

400 μM of spin label protein was incubated with 25-fold molar excess of ice-cold glutaraldehyde solution (Electron Microscopy Sciences, Hatfield, PA) at 4 ˚C. The mixture was immediately loaded into a sealed 0.6 I.D. X 0.84 O.D. glass capillary tube (VitroCom Inc, N.J.) and the crosslinking reaction was allowed to proceed at room temperature. The crosslinked protein formed a gel with a yellowish color after ~15 minutes of incubation time.

8.13 Construction, expression, and biotinylation of T4 lysozyme mutants bearing p-AcPhe unnatural amino acids and immobilization onto Streptavidin beads

Construction, expression, and purification on T4 lysozyme mutants bearing p-AcPhe unnatural amino acid was done as previously described [13]. Biotinylated hydroxylamine derivative shown in Figure 8.4A was purchased from Invitrogen (Catalog #: A10550). To biotinylate T4 lysozyme variants bearing p-AcPhe unnatural amino acid, a small amount (~ 1 - 2 mg) of lyophilized biotinylated hydroxylamine derivative was dissolved in biotinylation buffer (20 mM phosphate buffer, 150 mM NaCl, pH 4.5) and added in a 20-fold molar excess to ~ 50 – 100 μM of protein bearing p-AzPhe. The solution mixture was incubated overnight at 37℃ in biotinylation buffer. A small amount of precipitate was observed after overnight incubation, which was removed *via* centrifugation or by using a 0.2 μm membrane filter. Excess biotinylated hydroxylamine was removed with a desalting column or *via* repeated washing cycles with coupling buffer (100 mM phosphate buffer, 150 mM NaCl pH 7.2).

High-capacity Streptavidin resin was purchased from Thermo Scientific (catalog #: 20361). About 200 μl of 50% slurry of resin per sample was equilibrated with 5 - 10 medium volumes of binding buffer. 1 – 2 mg of biotinylated protein samples were incubated in coupling buffer with the streptavidin resin for 16 hours at room temperature in a nutator mixer. Any uncoupled protein was eluted with coupling buffer and the coupling efficiency was assessed by UV_{280} absorption. The coupling efficiency was typically 65 – 100% as shown in Figure 8.4B. The immobilized protein was equilibrated with the appropriate buffer selected for EPR studies.

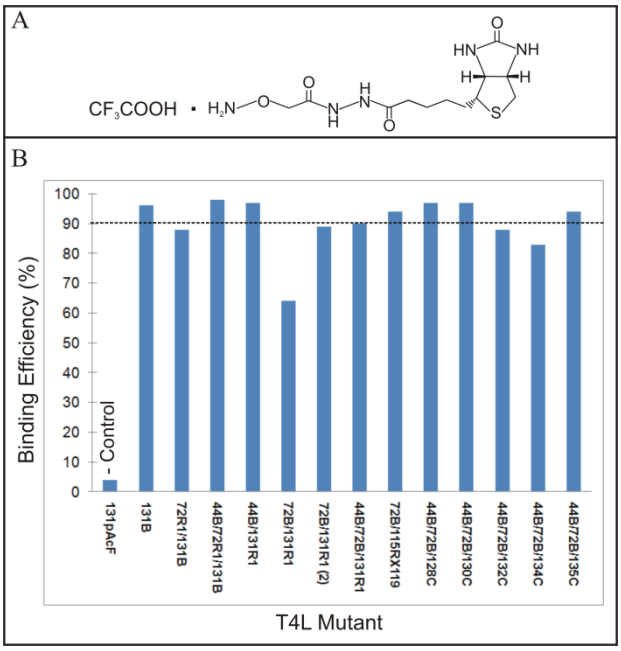


Figure 8.4: Biotinylation and coupling of T4 lysozyme mutants bearing p-AzPhe unnatural amino acid. A) Structure of N-(Aminooxyacetyl)-N'-(D Biotinoyl) hydrazyne, trifluoroacetic acid salt. B) Binding efficiencies of a set of biotinylated T4 lysozyme mutants to streptavidin beads. The biotinylated sites are denoted with a B.

8.14 EPR Spectroscopy

The EPR spectra were recorded in a Bruker ELEXSYS 580 fitted with a high sensitivity resonator at 298 K using an incident microwave power of 20 mW and modulation amplitude of 1 Gauss. Samples of at least 6 µl were contained in sealed glass capillary tubes (0.6 I.D. X 0.84 O.D; VitroCom Inc, N.J.). Spectra were recorded at X-band frequency with a scan width of 100 G. Solution spectra were measured either in 30 % w/w sucrose or 25 % w/w Ficoll 70 with a final protein concentration of 200 - 400 µM. To obtain EPR spectra of spin labeled proteins immobilized on sepharose, ~ 50 µl of the gel slurry containing the coupled protein was loaded into glass capillary tubes (1.5 I.D. X 1.8 I.D. ; VitroCom Inc, N.J.).

8.15 DEER spectroscopy

The four-pulse DEER experiments for the double spin-labeled proteins were conducted at 80K on the Bruker ELEXSYS 580 fitted with a 2-mm split ring resonator. The protein concentration during the DEER experiments was maintained at or below 200 μM. 20 μL of samples at the appropriate buffer conditions containing 20 % v/v glycerol as cryoprotectant were loaded into a quartz capillary tube (1.5 ID X 1.8 OD; VitroCom Inc., NJ) and then flash-frozen in liquid nitrogen. A 16 ns π–pump pulse was set at the maximum of the center-line absorption spectra and the observer π/2 (8 ns) and π (16 ns) pulses were positioned 70 MHz (25 Gauss) less than the pulse frequency, which corresponds to the absorption maxima of the low-field line. Additional parameters were set to the following values: gate = 68 ns, d0 = 432 ns, d1 = 200 ns, d2 = variable (1,500 – 3,000 ns), d3 = 100 ns, phase cycling = 2-step. Distance distributions were obtained from the raw dipolar evolution time raw data using the DEER Analysis 2009 program [14] (available at http://www.epr.ethz.ch/software/index) and with a custom program written in LabView software package (National Instruments, Austin, Tx) by Christian Altenbach (Jules Stein Eye Institute, UCLA). The primary data in a DEER experiment is the amplitude of the electron spin echo as a function of time, which has contributions from intra- and random inter-molecular dipolar interactions. A dipolar evolution function was obtained after an exponentially decaying background function was applied to correct for random inter-molecular dipolar interactions. Fourier transformation of the dipolar evolution function generated a dipolar spectrum, which was fit using Tikhonov regularization [15] techniques to get the interspin distance distribution.

8.16 Matching the effective viscosities of Ficoll 70 and sucrose solutions

The spin label 1 – oxyl -2,5,5 - trimethyl - 2,4 diphenyl - 2,5 - dihydro- 1*H* – pyrrol – 3 - ylmethyl methanethiosulfonate (a gift of Kalman Hideg, University of Pecs, Hungary) has been shown to give a nitroxide side chain (Figure 8.5) that is fully immobilized with respect to the protein when reacted with a surface cysteine residue [16]. This side chain was introduced at site 72 in T4L where there are minimal contributions from backbone fluctuations [17]. To a good approximation, the motion of the nitroxide for this label then measures the rotational motion of the protein as a whole. For EPR spectra of a nitroxide whose motion is close to the rigid limit, the rotational correlation time (τ_R) can be estimated as

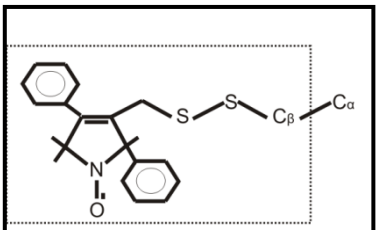


Figure 8.5: Spin-label used to match effective viscosities of sucrose and Ficoll 70 solutions.

$$\tau_R = a\left(1 - \frac{A_{zz}'}{A_{zz}^{R}}\right)^{b} \qquad 8.1$$

where A_{zz}' is the observed outer hyperfine splitting, A_{zz}^{R} is the value of hyperfine splitting in the absence of protein rotational motion, and *a* and *b* are constants [18]. Thus, τ_R is measured by A_{zz}' and the concentration of Ficoll that gives the same A_{zz}' as a 30% w/w sucrose solution at 295 K has an equal effective viscosity with respect to protein rotary diffusion. EPR spectra were recorded in the presence of 30% w/w sucrose (~0.988 M) and as a function of Ficoll 70 in the range of 0 – 30 % w/w. The hyperfine splittings were determined by individually fitting the low field and the high field peaks to a mixture of gaussian and lorentzian lineshapes using the Xepr program (Bruker, Germany) and by measuring the magnetic field separation between the low field and high field peaks to get $2A_{zz}'$. The hyperfine splitting A_{zz}' in 30% w/w sucrose was 32.92 ± 0.02 G. A plot of A_{zz}' vs Ficoll concentration showed that a solution containing 25%

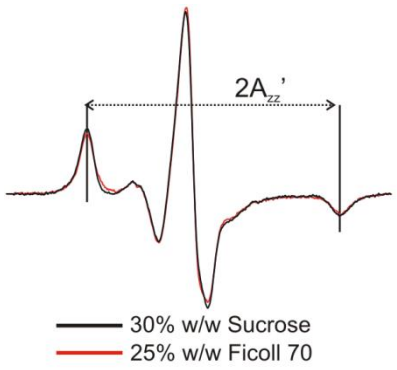


Figure 8.6: EPR spectra of D72R8 in sucrose and Ficoll 70.

w/w Ficoll has a hyperfine splitting value of 32.94 ± 0.02 G, identical to that of sucrose within the experimental error. An overlay of the EPR spectra of R8 at site 72 in 30% w/w sucrose and 25% w/w Ficoll shows them to be essentially identical (Figure 8.6).

8.17 Determination of osmolarity of sucrose and Ficoll 70 solutions

The osmolarity of 30% w/w sucrose and 25% w/w Ficoll 70 solutions were measured using a Wescor 5500 vapor pressure osmometer (Wescor, UT) according to manufacturer instructions.

8.18 Simulation of CW EPR spectra

EPR spectra were fit with the MOMD model of Freed and coworkers [19] using the program NLSL as previously described. The software is available at ftp://ftp.ccmr.cornell.edu/pub.freed. For all simulations, the starting values for the A and g magnetic tensor were taken as g_{xx} = 2.0078, g_{yy} = 2.0060, g_{zz} = 2.0022, and A_{xx} = 6.2, A_{yy} = 5.9, A_{zz} = 37, which correspond to a nitroxide side chain in an aqueous environment [20]. The $\log_{10}$ of the three components of the rotational diffusion tensor are given in a modified spherical form as $< R >= \frac{1}{3}(Rx + Ry + Rz)$, $N = Rz - \frac{1}{2}(Rx + Ry)$, and $Nxy = (Rx - Ry)$ where *Rx*, *Ry*, and *Rz* are the principal components in Cartesian form (1/s). The mobile states in the single and multicomponent spectra are taken to have axially symmetric (*Nxy* = 0) z-axis anisotropic motion with tilt angles of the diffusion tensor α_D = 0 °, β_D = 36 °, γ_D = 0 °, and subject to an ordering potential described by the coefficient C_{20} from which the order parameter S is computed [19]. To

constrain the numbers of fitting parameters for multicomponent spectra, the relatively immobile component were taken to have isotropic motion where $N = Nxy = 0$. Least-square fits were obtained by variation of only $<R>$, N, and C_{20} only for the mobile components and only $<R>$ for the immobile states. After the above parameters were optimized, the principal values of the A and g magnetic tensor were allowed to vary slightly in order to get the final best fits. From the final fits, the effective correlation time (τ_e) was calculated as $\tau_e = \frac{1}{6*10^{<R>}}$ and the order parameter was computed directly from C_{20} [19].

8.19 CD spectroscopy

8.19.1 Thermal and chemical denaturation of spin-labeled proteins: CD measurements were conducted on a Jasco-810 spectropolarimeter equipped with a peltier-temperature controlled sample holder using a cuvette with 1-mm path length. For T4L and apo-Myb, the ellipticity was monitored at 223 nm and the temperature was increased at a rate of 2 ˚C per minute from 20-80 ˚C. The T4L sample consisted of 1 mg/ml protein in 25 mM MOPS, 12.5 mM NaCl, pH 6.8. At this pH unfolding was irreversible, thus it was used only to give information on relative stabilities. In order to obtain accurate thermodynamic parameters, the denaturation of T4L was performed at a pH 2.95 (20 mM potassium phosphate, 25 mM potassium chloride), at which the T4L unfolding was reversible. For apo-Myb the concentration was adjusted to 1 mg/ml in buffer containing 10 mM sodium acetate, pH 6.1. For apo-rIFABP, denaturation was carried out at 25 ˚C by monitoring the CD signal at 216 nm as a function of guanidinium HCl concentration. Protein concentration was adjusted to 1 mg/ml with a solution containing 20 mM phosphate buffer, 1 mM EDTA, pH 7.3. The raw ellipticity (θ) were converted to mean residue ellipticity [θ], using the following equation.

$$[\theta]\ deg * cm^2/dmol = \frac{\left[\left(\theta/_{1000}\right)*(1000cm^3)*\left(\mathrm{Mw}/_{\#\ aa-1}\right)\right]}{[0.1cm*(C)*10dmol]} = \frac{\left(\frac{Mw}{\#\ aa-1}\right)}{C}, \quad 8.2$$

where *Mw* is the molecular weight of the protein in daltons, *aa* is the number of amino acids, and *C* is the protein concentration in mg/ml.

The chemical and thermal denaturation data were fit using a linear extrapolation method of Santoro and Bolen [21] as shown below.

8.19.2 Monitoring secondary structures of spin-labeled proteins: The far-UV spectra (190-260 nm) of T4L, apo-Myb, and apo-rIFABP were recorded at 25 ˚C using a 1-mm cell with a protein concentration of 0.2 - 0.5 mg/ml in the appropriate buffers. Spectra were recorded from 190 - 260 nm and were an average of 4 scans.

8.19.3 Equations used for thermal and chemical denaturation of spin-labeled proteins:

A) <u>The thermal denaturation data was fit using Origin 6 graphic software with equation 3 using the following set of definitions:</u>

Parameters: h,tm,u,l,p (h = starting enthalpy in cal/mol; tm = starting Tm in ºC; u = mean residue ellipticity of 100% folded protein; l = mean residue ellipticity of fully unfolded protein; p = ΔCp in cal/mol). For T4 lysozyme mutants ΔCp was set to 2,500 cal/mol*deg as described in Baldwin et al. [22].

```
m=tm+273.15;
t=x+273.15;
g=h*(1-t/m)-p*((m-t)+t* ln(t/m));
k=exp(-g/(1.987* t));
z=k /(1+k);
f=((u-1)*z)+1;
y=z*(u-l)+l
```

B) <u>For the chemical denaturation fit, ΔG(H_2O) and *m* values were calculated in Origin 6 using the following set of definitions:</u>

Parameters: P_1, P_2, *x*, m_1, m_2, m, c (**P_1** = mean residue ellipticity of 100% folded protein; **P_2** = mean residue ellipticity of fully unfolded protein; ***x*** = [denaturant]; **m_1** = pre-transition slope ; **m_2** = post-transition slope , **m** = *m*; **c** = $[Urea]_{1/2}$). The data was fit to the equation #4 to get **m** and **c:**

```
a=P1+m1*x;
b=(P2+m2*x)*exp(-1.688)*m*(c-x));
d=1+exp(-(1.689)*m*(c-x));
y=(a+b)/d
```

To calculate $\Delta G(H_2O)$, the data was fit to the following equation using the parameters: P_1, P_2, *x*, m_1, m_2, m, G ($G = \Delta G(H_2O)$) and m value obtained from the previous fit:

```
a=P1+m1*x;
b=(P2+m2*x)*exp(-(1.689)*(G-m*x));
d=1+exp(-(1.688)*(G-m*x));
y=(a+b)/d
```

8.20 Calculation of fraction of surface-buried

Fraction of solvent-accessibilities for myoglobin was calculated on a per residue basis with the program Getarea [23] (server can be directed accessed at: http://curie.utmb.edu/getarea.html) using the atomic coordinates of the high-resolution structures of myoglobin and T4 lysozyme (PDB: 2mbw, 3lzm, respectively) and a probe radius of 1.4 Å. In this program, the output of the calculation is the ratio of the solvent-exposed area of every side chain in the structure, which is defined as the solvent-exposed area of residue X in the context of the structure normalized by the solvent-exposed area of the same residue in the tripeptide Gly-X-Gly in a random conformation. Residues are considered to be solvent exposed if the ratio > 0.5 and to be buried if ratio < 0.2 [23]. To calculate fraction of surface buried, the solvent *inaccessible* area per residue was obtained using the output from the program Getarea. In order to obtain the *segmental* fraction of surface buried near sites where the nitroxide side chain was introduced, the values for the solvent

inaccessible surface area over a sliding window of desired length (6-9 residues around R1) were averaged. The short windows (6 residues) were used for cases in which the nitroxide was introduced near a helix terminus and for short helices (helices C and D), whereas the 9-residue (from *i - 4* to *i + 4*) window was used for R1 in the center of helices.

8.21 Modeling of the R1 side chain on the structures of T4L and Myoglobin

The R1 side chain was modeled at the indicated sites in the high-resolution x-ray structure (PDB ID: 2MBW for myoglobin and 3LZM for T4L) using DS Viewer Pro 5.0 software (Accelrys, Inc). The *X1 – X3* dihedral angles modeled were those previously observed at helical sites in crystals structures of T4 lysozyme [1,2,24,25]; either {-60, -60, *X3*}, {180, -60, *X3*}, or {180, 60, *X3*}, where *X3* is ± 90°. In each case, the *X4* and *X5* rotamers modeled were those reported by Guo et al. [2]. For each site, a particular R1 rotamer was selected such that there are no steric clashes or overlap on the nitroxide side chain with the rest of the molecule.

8.22 References

1. Guo, Z., Cascio, D., Hideg, K., Kálái, T., and Hubbell, W.L. 2007. Structural determinants of nitroxide motion in spin-labeled proteins: Tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. *Protein Sci.* 16: 1069-1086.
2. Guo, Z., Cascio, D., Hideg, K., and Hubbell, W.L. 2008. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci.* 17: 228-239.
3. Mchaourab, H.S., Lietzow, M.A., Hideg, K., and Hubbell, W.L. 1996. Motion of spin-labeled side chain in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35: 7692-7704.

4. Sathish, H.A., Stein R.A., Yang, G., Mchaourab, H.S. 2003. Mechanism of chanperone function in small heat-shock proteins. Fluorescence studies of the conformations of T4 lysozyme bound to alpha B-crystallin. *J Biol. Chem*. 278: 44214-44221.
5. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77: 51-59.
6. Liu, L., Baase, W.A., Michael, M.M., Matthews, B.W. 2009. Use of stabilizing mutations to engineer a charge group within a ligand-binding hydrophobic cavity in T4 lysozyme. *Biochemistry* 48: 8842-4451.
7. Springer, B., and Sligar, S.G. 1987. High-level expression of sperm whale myoglobin in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 84: 8961-8965.
8. Jennings, P.A. Stone, M.J., Wright, P.E. 1995. Overexpression of myoglobin and assignments of it amide, C^{α} and C^{β} resonances. *J. Biomol. NMR*. 6: 271-276.
9. Hapner, K.D., Bradshaw, R.A., Hartzell, C.R., Gurd, R.N. 1968. Comparison of myoglobins from harbor seal, porpoise, and sperm whale: I. Preparation and characterization. *J. Biol. Chem*. 243: 683-689.
10. Ulyanov, D., Bowler, B.E., Eaton, G. R., Eaton, S.S. 2008. Electron-electron distances in spin-labeled low-spin metmyoglobin variants by relaxation enhancement. *Biophysical J*. 95: 5306-5316.
11. Lowe, J. B., Sacchettini, J.C., Laposata, M., McQuillan, J.J., and Gordon, J.I. 1987. Expression of rat intestinal fatty acid-binding protein in *Escherichia coli. J. Biol. Chem*. 262: 5931-5937.
12. Glatz, J.F.C., and Veerkamp, J.H. 1983. Removal of fatty acids from serum albumin by Lipidex 1000 chromatography. *J. Biochem. Biophys. Methods* 8: 57-61.

13. Fleissner, M.R., Brustad, E.M., Kálai, T., Altenbach, C., Schultz, P.G., and Hubbell, W.L. 2009. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci.* 106:21637-21642.

14. Jeschke, G. 2006. DeerAnalysis2006: A comprehensive software package for analyzing pulse ELDOR. *Appl. Magn. Reson*. 191:473-498.

15. Jeschke, G., and Polyhach, Y. 2007. Distance measurements on spin-labeled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* 9: 1895-1910.

16. Sale, K., Sár, C., Sharp, K.A., Hideg, K., and Fajer, P.G. 2002. Structural determination of spin label immobilization and orientation: A Monte Carlo minimization approach. *J. Mag. Res*. 156: 104-112.

17. Columbus, L., Kálai, T., Jekö, J., Hideg, K., and Hubbell, W.L. 2001. Molecular Motion of Spin Labeled Side Chains in α-Helices: Analysis by Variation of Side Chain Structure. *Biochemistry* 40: 3828-3846.

18. Freed, J.H. *Theory of slow tumbling ESR spectra for nitroxides. In Spin Labeling*: Theory and applications, L.J. Berliner, ed. New York, Academic Press, pp. 53-132.

19. Budil, DE., Lee, S., Saxena. S., and Freed, J.H. (1996) Nonlinear least-squares analysis of slow-motional EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm. *J. Magn. Res. A.* 120: 155-189.

20. Kusnetzow A, Altenbach C, and Hubbell WL (2006) Conformational states and dynamics of rhodopsin in micelles and bilayers. *Biochemistry* 45:5538-5550.

21. Santoro M.M, and Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. Biochemistry 27:8063-8068.

22. Baldwin, E., Baase, W.A., Zhang, X., Feher, V., Matthews, B.W. 1998. Generation of ligand-binding sites in T4 lysozyme by deficiency-creating mutations. *J. Mol. Biol.* 277; 467-485.

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

24. Langen, R., Oh, K.J., Cascio, D., and Hubbell, W.L. 2000. Crystal structures of spin labeled T4 lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry* 39:8396-8405.

25. Fleissner, M.R., Cascio, D., and Hubbell, W.L. 2009. Structural origins of weakly ordered motion in spin-labeled proteins. *Protein Sci.* 18:893-908.