Title
Exploring the conformational space of cavity-containing proteins with pH, high pressure, and site-directed spin labeling electron paramagnetic resonance

Author
Kreitman, Margaux

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Exploring the conformational space of cavity-containing proteins with pH, high pressure, and site-directed spin labeling electron paramagnetic resonance

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemistry

by

Margaux Jane Kreitman

2019

ABSTRACT OF THE DISSERTATION

Exploring the conformational space of cavity-containing proteins with
pH, high pressure, and site-directed spin labeling electron paramagnetic resonance

by

Margaux Jane Kreitman
Doctor of Philosophy in Chemistry
University of California, Los Angeles, 2019
Professor Wayne L. Hubbell, Chair

Proteins in pre-existing conformational equilibria sample different conformational states, some of which have important functional roles. Among intermediate conformational states, some proteins adopt a molten globule (MG) structure, which is compact and contains a relatively high content of native-like secondary structure, but has fewer tertiary contacts. The MG is a dynamic, flexible intermediate that adapts to a variety of conformations. Growing evidence that the MG is a hub in conformational equilibria between folded states emphasizes its putative functional importance. While the MG states of many α-helical proteins have been thoroughly studied, our understanding of β-sheet protein dynamics is incomplete. The plasticity of MG-like states of intestinal fatty acid binding protein (I-FABP), a β-barrel protein, and the T4 lysozyme (T4L) L99A mutant, a predominately α-helical protein, are thought to facilitate binding of a variety of ligands to their large cavities (230 and 150 $Å^3$ for I-FABP and T4L L99A, respectively). X-ray

crystal and nuclear magnetic resonance structures of I-FABP and T4L L99A do not identify an open conformation that permits ligand entry, suggesting that a rare MG may be facilitative.

Site-directed spin labeling and EPR spectroscopy (SDSL EPR) is a sensitive tool for identifying backbone dynamics, conformational exchange, and ligand binding of proteins. In this dissertation, a panel of EPR experiments (continuous wave ((CW)) EPR, saturation recovery, and double electron-electron resonance) provides information on nanosecond-microsecond timescale motions, their amplitudes, and the thermodynamic equilibrium between these states to lay a basis for the structure and dynamics of the I-FABP and T4L L99A MGs. These results are supported by other forms of optical spectroscopy (intrinsic fluorescence and dynamic light scattering).

As a means to populate rare intermediates, hydrostatic pressure is known to shift conformational equilibria to populate intermediate states through cavity hydration or structure-relaxation mechanisms. Internal hydration of proteins has structural parallels with the MG state. A comparison of the acid pH-stabilized MG of I-FABP and its high pressure states reveals structural similarity, in support of a model for protein conformational equilibria that involves a limited number of discrete conformational states rather than a continuum of intermediate conformations.

The MG-like character of the high pressure state of the T4 lysozyme cavity mutant L99A is also investigated using pressure-resolved CW EPR. In contrast with wild-type T4L, the L99A mutant shows site-specific smooth sigmoidal transitions with pressure that indicate a two-state equilibrium between the native and intermediate states. The goal of this work is to determine how cavity hydration affects the protein structure allosterically.

The main goal of this dissertation is to characterize the extent of structural heterogeneity

of the MG from the SDSL EPR perspective, which encompasses fluctuation amplitudes, motional timescale, and the thermodynamic equilibrium state of the MG.

This dissertation of Margaux Jane Kreitman is approved.

James U. Bowie

Gabriel Travis

Wayne L. Hubbell, Committee Chair

University of California, Los Angeles

2019

For my family

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Professor Wayne Hubbell for his mentorship and guidance throughout the course of my graduate work. The opportunity to freely explore science in his laboratory has been an incredible boon to my development as a biochemist. I would also like to thank Cherie Hubbell who, during her time in the lab, was a source of support and encouragement.

I would like to also thank the many members of the Hubbell lab, past and present, who have shaped my research and overall graduate experience. In particular, Christian Altenbach has been a source of wisdom and logic. Carlos López, Ned van Eps, Michael Lerch, Kevin Eden, Matthias Elgeti, Michael Bridges, and Mengzhen Chen have all introduced me to practical laboratory techniques and models of analysis and have treated me to illuminating conversations of all kinds.

I am also grateful to the faculty and fellow students of the Jules Stein Eye Institute from which I have been acquainted with the field of vision science—which, although not reflected in this dissertation, has been a significant component of my graduate studies.

Lastly, I appreciate the time and effort of the members of my thesis committee.

This work was supported by NIH grant 5R01 EY005216 and the Jules Stein Professor Endowment.

VITA

Education

B.S. in Biochemistry and Molecular Biology, University of California, Davis (2011)

Publications

Lerch MT, López CJ, Yang Z, **Kreitman MJ**, Horwitz J, & Hubbell WL (2015) Structure-relaxation mechanism for the response of T4 lysozyme cavity mutants to hydrostatic pressure. Proc Natl Acad Sci USA 112(19):E2437-46.

Awards

Best oral presentation, UCLA Vision Science Retreat, October 2016

"Exploring the conformational space of Arrestin-1 with high pressure and electron paramagnetic resonance"

Teaching experience

Teaching Assistant for CHEM 153A "Introduction to structure, enzymes, and metabolism"

Winter 2012 with Dr. Heather Tienson,

Summer 2012 with Dr. Jim Gober

TA for CHEM 153C "Biosynthetic and energy metabolism and regulation"

Spring 2012 with Dr. Steven Clarke

Chapter 1. Protein dynamics and conformational equilibria

Connecting protein structure and dynamics has long been a major focus in protein science. X-ray crystallography has long been considered the "gold standard" of structural biology[1,2], providing all-atom protein structures that are high-resolution but static descriptions of a protein conformational state. Conformational motion and protein dynamics can be inferred from comparisons between crystal forms of a protein. Another indication of dynamics via interpretation of thermal B-factors reflects flexibility of protein structure within the crystal lattice. In contrast, solution-phase NMR has directly revealed protein dynamics in small proteins (typically <50 kDa), but suffers from poor sensitivity for structures that are highly heterogeneous. Electron paramagnetic resonance (EPR) of spin labeled proteins a powerful and sensitive monitor of protein motion and dynamics in solution, not limited by the lattice or structural complexity. Appreciation of the flexible nature of proteins has been a cornerstone of protein science and the development of models to describe protein structure and function have increasingly emphasized the motional characteristics of proteins.

1.1 A historical perspective on protein dynamics

Early studies focused on catalysis by enzymes to elucidate how these biomolecules operate. Emil Fischer, from a lifetime of studying how asymmetric forces on the molecular level drive reactions, proposed a model for efficient catalysis by enzymes in 1894[3] (and 1898[4]). The "lock-and-key" model explains enzyme specificity towards particular substrates due to the asymmetry or steric specificity of the ligand and protein active site. However, modeling rigid substrate and enzyme structures proved inconsistent with new experimental observations, opening the field to a number of new developments to the lock-and-key model.[5,6] Even as early

as 1938, Felix Haurowitz demonstrated that deoxyhemoglobin crystals shatter when introduced to oxygen due to a conformational change in the protein large enough to break crystal packing interactions.[7] In 1958, the same year the first all-atom X-ray crystal structure of a protein was published[8], Daniel E. Koshland Jr. proposed a new model for protein function by articulating how protein motion plays a role in ligand binding.[6] His "induced-fit" model provides a simple mechanism for molecular recognition that requires a protein conformation to restructure its active site in the presence of a specific bound substrate. In order for catalysis, direct evidence for a conformational change in the enzyme was shown by ligand binding in a kinase that resulted in exposure of a catalytic side chain.[9] The induced-fit model ties enzyme flexibility to its ligand specificity in a way that the lock-and-key model cannot.

However, enzymes can have specificity towards a broad range of ligands beyond the "one enzyme—one substrate—one reaction" paradigm and have been known to exhibit ligand binding promiscuity[1]. Experimental evidence of enzyme promiscuity in the 1960s following the introduction of Koshland's model includes activity of enzymes such as carbonic anhydrase[10], pepsin[11], and chymotrypsin[12], which were shown to be capable of catalyzing multiple related reactions. Implied from the broad functionality of promiscuous enzymes is a flexible protein conformation that stabilizes the catalytic transition state is and accommodates different ligand structures. A limitation of the induced-fit model is that it assumes the active site of the protein is both highly structured and ligand-accessible in the native state, an assertion that is not always supported by the structures revealed by X-ray crystallography. For example, in the crystal lattice, many apoproteins have an occluded active site for which the ligand point of entry is not obvious. Growing evidence for protein dynamics in the solution state from a variety of techniques

[1] A semantic aside: "enzyme promiscuity" is a term sometimes only applied to nonbiological enzyme functions, such as enzyme activity towards a substrate not found in a particular cell type. Enzymes that catalyze a broad range of substrates can be instead considered "multispecific" or "broad-specificity enzymes."[68]

including nuclear magnetic resonance[13,14] and intrinsic fluorescence[15] show that even in the native conformation, solvent penetration into the protein interior is permitted. The role of transient fluctuations of protein structure was outlined by Alan Cooper in 1976, who described protein conformational states as time-averaged in the ensemble[16], introducing a picture of dynamic conformational states of proteins. This new dynamics perspective is in contrast to the compact and static conformations observed in X-ray crystal structures and the discrete, rigid states described in the lock-and-key and induced-fit models.

Finally, the connection between thermal equilibrium protein fluctuations and function was made clear by extensive studies of ligand binding in holo myoglobin (holo Mb)[17], in which multiple binding events were revealed by temperature dependence on rebinding rates of CO and O_2 after laser photodissociation. Analysis of binding rates yielded activation energies for four separate barriers that separate discrete conformational states, each with a different mode of ligand binding. Furthermore, it was shown that these states are preferentially populated in separate temperature regimes. Multi-step reaction kinetics directly implicated the presence of multiple conformational states (at least in the vicinity of the CO- and O_2-binding cavities) of holo Mb in equilibrium. That the flexibility of a folded protein can be connected to an ensemble of conformations led to a description of the energy landscape of protein states and motions by Frauenfelder *et al.*[18] in a scheme that separates conformational exchange from flexibility by the magnitude and timescale of different kinds of motions.

1.2 The energy hierarchy of protein states

Protein conformational states can be organized into a conceptual energy landscape that maps all conformational states to their relative free energies.[19] The native state has the lowest

relative free energy and forms the lowest point of this surface, from which an increasing number of higher-energy conformations emerge. Unfolded states populate the top of the energy landscape. This map is in the shape of a hollow funnel (Fig. 1.1a). Each state is populated according to its relative free energy such that a preponderance of the native state is in equilibrium with minor populations of higher-energy conformations according to the pre-equilibrium model. Because protein conformations are state-dependent and are therefore coupled to conditions such as temperature, pressure, and solvent properties, changes in the environment

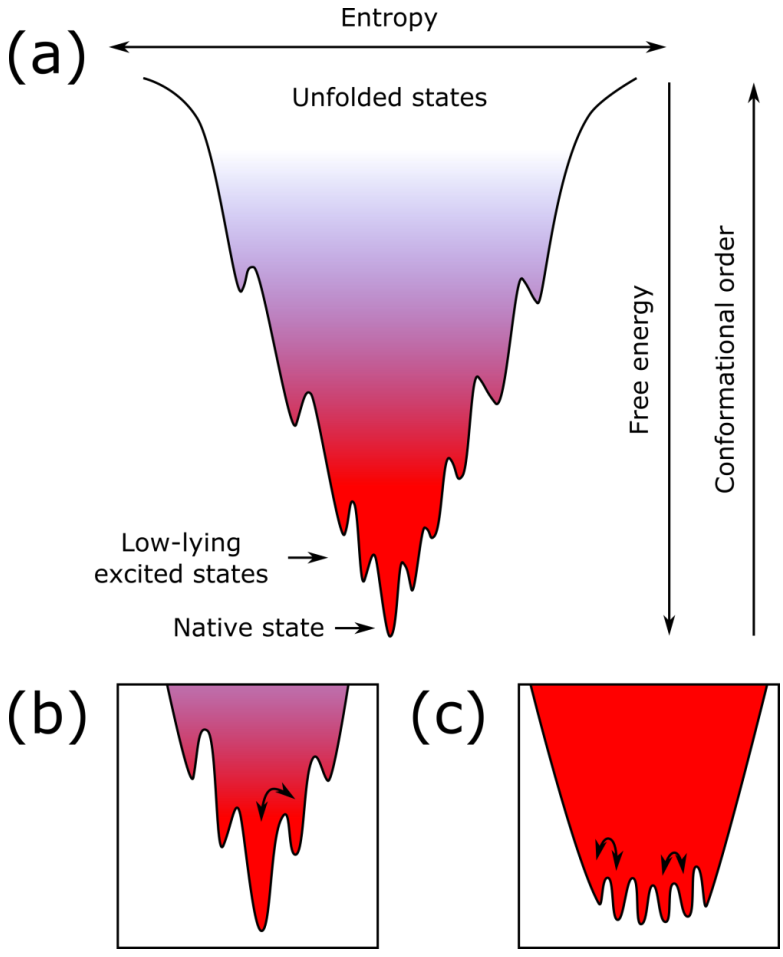

Figure 1.1 The energy hierarchy of protein states. In (a), each conformational state is represented by a local well on an energy surface that desribes all possible structures of a protein. In (b), the states with similar energy to the native state are shown, and transitions between these conformational states occur on the μs-ms time scale. In (c), transitions between statistical substates within an energy well of a conformational state occur on the ps-ns time scale.

alter the shape of the protein energy landscape.

Proteins explore other low-lying conformational states under physiological conditions (Fig. 1.1b). Conformational states are separated by energy barriers >1 k_BT, the reference for thermal energy (the product of k_B, the Boltzmann constant and T, temperature).[20] Exchange between conformational states involves structural changes that are large in amplitude and low in frequency with exchange lifetimes on the μs-ms time scale. Such large-amplitude motions have been categorized by Chothia *et al*. as hinge and shear movements.[21] Hinge motions that involve rigid-body movement of secondary structure elements are independent of side chain interactions, while shear movements are large-amplitude sliding motions require side chain repacking. Other slow conformational motions on this time scale include order-to-disorder transitions and fluctuations between poorly-ordered states which often fit induced-fit or "flycasting" models of protein function (proposed by Shoemaker *et al.* in 2000[22], described below).

Within the energy well of each conformational state exist a large number of statistical substates with lower energy barriers (<1 k_BT) that have short lifetimes on the ps-ns time scale (Fig. 1.1c). These states are often treated as a statistical sum of states best described by their fast exchange timescale. Transitions between statistical substates involve small-amplitude motions limited to distances predicted from Debye-Waller factors (or B factors)[23] that are due to thermal fluctuations.

It is evident that the multidimensionality of protein structure is essential to understand mechanisms of action for proteins, encompassing amplitude, time, and solvent conditions that all shape the energy landscape. In the "pre-equilibrium" model, all possible conformations of a protein exist at equilibrium with populations inversely proportional to their free energies. Exchange between conformational states requires reorganization of tertiary and secondary

structure elements. Protein science today is focused on identification of catalytically active or functionally important conformational states, which may be populated only to a few percent[24] as a means to regulate protein activity.[25]

1.3 The molten globule is a compact, non-native conformational state

Non-native or excited states of proteins are now recognized as essential to protein function, stability, and folding. The number of conformational states is finite and the primary sequence of a protein encodes nearly all the 3D structural information necessary to fold correctly (rather than reaching a desired fold by a random search, as argued against in Levinthal's paradox[26]). Folding transitions between states at the top and bottom of the energy landscape funnel (Fig. 1.1a) are often modeled as two-state. However, as discussed above, refolding kinetics suggest the two-state theory is insufficient due to observations of stepwise folding pathways with long-lived intermediates. Common features between numbers of intermediates have led to the identification of a protein state called the molten globule (MG). The MG possesses native-like secondary structure but unfolded tertiary structure which accommodates hydration of the hydrophobic core. This hydration in the MG is has been described as "trapped but exchangeable" water, based on ^{17}O magnetic relaxation dispersion NMR.[27–29] The hydration of the native, MG, and unfolded states are demonstrably different: while the native state has fewer tightly associated waters than the MG, the unfolded state has fewer yet.[30] Experimental conditions which are shown to populate MGs include low and high pH, presence of inorganic salt denaturants (esp. $LiClO_4$), mild concentrations of chemical denaturants (~2M urea, GdnHCl), reduction of disulfide bonds, and removal of structural divalent cations (Ca^{2+}). Acid pH populates a MG either by protonation of the surface side chains or by breaking salt bridges.

The first mechanism results in net positive charge distributed across the surface of a protein, which expands the protein via electrostatic repulsion. By protonating salt bridges, the second mechanism can also "unlock" interactions that stabilized the native state. In contrast, chemical denaturants populate MG states of proteins, even at neutral pH, by mechanisms that are not entirely clear, but either involve strong interactions of the denaturant with solvent that restructure the water molecules associated with the protein (hydrophobic solvation)[31] or by directly interacting with protein[32]. Lastly, pressure can also populate a MG by mechanisms described in Chapter 3.

A number of small, globular proteins have well-characterized MG states. Notable examples include apo myoglobin (apo Mb) and α-lactalbumin (α-LA), which have a high degree of native-like secondary structure and are weakly stabilized by tertiary interactions. The MG states of bovine pancreatic trypsin inhibitor (BPTI) and β-lactoglobulin (β-LG) have been less thoroughly investigated but also provide insight into how the MG is formed in mixed α-helical/β-sheet proteins. The native folds of these proteins contain different arrangements of secondary structure elements (Fig. 1.2) and the structural changes in the MG states of these proteins are diverse, as described below.

MG states of apo Mb, populated at low pH[33] and high pressure[34], show significant structural heterogeneity. At pH 4.1, the MG of apo Mb has stable native-like α-helical structure in helices A, B, G, and H[35], while helices C, D, E, and the N-terminal half of helix B have low amounts of secondary structure[35]. Pressure perturbation to 2000 bar of the pH 6.0 native state of apo Mb populates a similar MG in which α-helices B, C, D, and E retain native-like secondary structure and have large-amplitude rigid body motions[35]. Both low pH and high pressure conformations satisfy the conditions required of a MG.

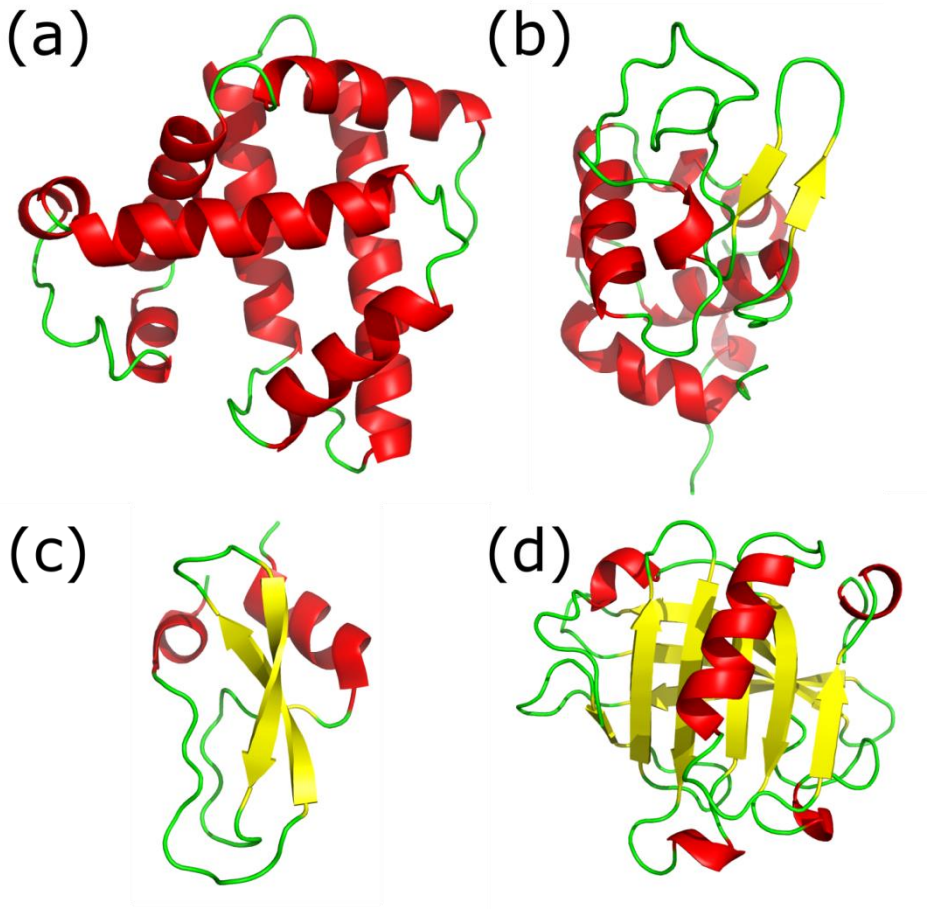

Figure 1.2 X-ray crystal structures of the native conformations of four proteins with well-characterized MGs. In order of decreasing α:β secondary structure content: (a), holo myoglobin (RCSB accession code: 2BMW), (b) α-lactalbumin (1A4V), (c) bovine pancreatic trypsin inhibitor (1BPI), and (d) β-lactoglobulin (1BEB).

Population of the MG can also involve domain-specific structural changes. The MG of α-LA is studied under a variety of conditions: low pH, mild denaturant conditions, and reduction of disulfide bonds all produce a similar conformational state[36], in which local unfolding in the β-sheet domain is accompanied by an α-helical domain that is in an expanded yet native-like conformation.[37]

Similarly, the low pH MG of BPTI has a persistent α-domain with a β-domain MG. The BPTI MG is noticeably less compact compared to its native state compared to other α-helical

MGs.[38] However, use of this model peptide for characterizing the nature of the β-sheet MG is complicated by its small size, necessitating investigation of other β-sheet MGs (as will be discussed in Chapter 4). The BPTI β-domain is made up of 2 edge strands, which are expected to have low stability compared to β-sheets with protected or hydrogen-bonded edge strands, even in the native state. Generally, exposed edges of β-sheets are associated with aggregation(55) and have high propensity for forming protein-protein interactions(56) and as such, BPTI may not be representative of most monomeric β-proteins.

A last example highlights the need to carefully consider which states are true MGs. β-LG MGs are populated at low pH[39] or high pressure[40]. Far-UV CD shows a transition in secondary structure from native β-sheets to non-native α-helices by forming a non-native disulfide network[40,41], populating not a MG but a MG-like state, as a canonical MG would be more closely related to the native state. While repacking of the primary sequence of β-LG into two stable conformations is remarkable, this unique transition indicates that the MG-like states of β-LG are populated by mechanisms not common to all MGs.

Overall, proponents for the significance of the MG emphasize that observation of similar MG states in a diverse group of globular proteins populated by a variety of conditions suggest that the MG is an excited state conformation common to proteins in general.

While global experimental techniques do not indicate localization of the MG within a protein fold, experiments with site-specific information best describe the conformation of the MG. For example, 1D ^{1}H NMR nuclear Overhauser effect measurements provide information about the hydrophobic clustering of aromatic residues that persist in the MG state[42,43], distinguishing the ordered core of a MG from other dynamic, fluctuating elements. Examples of how structural heterogeneity is linked to possible functions of the MG are described below.

1.4 Functional role of the molten globule

To date, the MG has been demonstrated in only a limited number of proteins, though they may be common to the folding pathways of all proteins. The MG is therefore implicated in misfolding and aggregation, with consequences related to amyloid diseases.[44,45] In addition to their roles in protein folding, some MGs play roles in cellular processes and retain the ligand binding function of the native fold[46]. Mechanisms for the ligand binding ability of a MG may involve the models discussed above for typical, compact proteins, but can be applied to the MG state:

1. The MG can bind ligand via an induced-fit model in the low ligand concentration limit[47], in which an encounter complex between a partially structured MG interacts weakly with a ligand and the interaction is strengthened via a conformational while bound to ligand.
2. Ligand binding via conformational selection in the high ligand concentration limit[47] involves a pre-existing high-affinity ligand binding conformation that is a MG[48].
3. Ligand binding via the fly casting model[22] entails an unstructured region of the MG that fluctuates rapidly between partially folded states, some of which form few, weak initial contacts with ligand, and are subsequently "reeled in" as folding and binding proceed simultaneously. Conceptually, this model is a multistate variation of the induced fit model.

MG states with ligand-binding activity have been described in a small handful of globular proteins. For example, ubiquitin is a highly plastic protein with a MG state that is populated either by hydrostatic pressure at 3000 bar or by the Q41N mutation[49]. Ubiquitin association with E2-ubiquitinating enzyme is enhanced by MG character[49], indicating that the functional role of ubiquitin proceeds through protein-protein interactions which directly involve the MG state via a

mechanism that can be described by either induced fit or conformational selection models.

Similarly, the MG state of *E. coli* periplasmic binding proteins, which transport alkyl amino acids and carbohydrates from the periplasmic space between the cell wall and the cell membrane into the cytoplasm[50], bind native ligands. PBPs transport alkyl amino acids and carbohydrates into the cytoplasm[50]. The low pH MG state of PBPs bind these ligands without conversion to the native state, though sometimes with reduced affinity.[51] Promiscuous yet specific binding of PBPs to a variety of carbohydrates[52] is consistent with the fly casting model, by which a variety of stable complexes are formed from a flexible apoprotein.

Other examples of MG states capable of binding ligands include the nuclear coactivator binding domain (NCBD) of CREB binding protein, which interacts with a host of transcription factor proteins, including p53[53], p73[54], steroid receptor coactivators[55], and interferon regulatory proteins[56]. The apo state of NCBD is a MG[48], with a compact core that resembles the steroid receptor coactivator-bound conformation.

In some cases, ligand binding to the MG results in ordering of the protein. The MG state of α-lactalbumin (α-LA) is a target of molecular chaperones such as GroEL which assist folding. Conflicting reports on the efficacy of this interaction suggest that the MG of α-LA, populated by removal of Ca^{2+}, only weakly associates with GroEL[37], indicating that the native state of α-LA and residual secondary structure elements within the α-LA MG are therefore not recognized by GroEL. Instead, the MG populated by either removal of Ca^{2+} and disulfide reduction[37] or low pH in the presence of cations at high ionic strength[57] are shown to strongly interact with GroEL. From these conditions it was shown the exposed hydrophobic surface in the MG state is a dominant force in recognition and internalization of the α-LA MG within the GroEL multimer.[58] GroEL specificity toward a common shared structural feature of all MGs supports the role of the

MG as a discrete conformational state.

A last example illustrates harnessing the unique characteristics of the MG to engineer a higher-efficiency enzyme: a nonnative loop introduced into chorismate mutase (CM) stabilizes the monomeric state of the enzyme (mCM), which has MG-like structure in the apo form.[59] Remarkably, binding and release of ligand are both more rapid in the mutant mCM than for dimeric CM[46], providing promise for design of enzymes based on dynamic properties, especially of the MG.

1.5 Thermodynamics of the molten globule

Thermodynamic characterization of the MG has revealed how stability is achieved in a flexible protein. Because the MG has a high amount of secondary structure, it has a similar enthalpy of unfolding compared to the native state. However, loss of close packing results in small unfavorable decrease in enthalpy[60], and the heat capacity of the MG is slightly higher than the native state because some of the solvent-inaccessible surface in the native state is made solvent-accessible.

Additionally, a dry MG (DMG, discussed more completely in Section 1.6) has been proposed which is stabilized by a large gain in conformational entropy.[61–63] That this increase in entropy stabilizes the MG is part of a new perspective on protein structure that close packing of globular proteins is not always the major mechanism by which protein structure is stabilized.[64]

1.6 Recent perspectives on the MG

A small assortment of MG-like states have been identified, creating the need for precisely defining the MG. Two such states are the “pre-molten globule” and the “dry molten globule,”

which share some but not all characteristics with the (implied "wet") MG. Differences between the MG and the intrinsically disordered protein (IDP) are clarified by more complete characterization, generally requiring multiple experimental techniques.

The pre-MG is an equilibrium conformation but not a true MG, but instead has poor preservation of secondary structure (down to ~50%), no rigid tertiary structure, compactness lower than the MG but higher than unfolded states, and intermediate accessibility of buried Trp residues compared to the native and unfolded states.[65] The pre-MG can be composed of non-native secondary structure, as well, and was first identified as a rapid burst phase folding intermediate.[66]

The dry MG (DMG) is a compact intermediate first defined in 1989, but was regarded as nonphysiological for soluble proteins and possible to observe only in nonaqueous environments (e.g. in membranes).[67] Renewed interest in the DMG in recent years has been accompanied by a redefinition of the state. Small-amplitude shifts in secondary structure elements loosen tight packing interactions, without hydration of the protein core.[63] This state would be indistinguishable from the native state by hydrogen exchange or intrinsic fluorescence approaches and the small magnitude of structural rearrangement presents challenges in measurement. Direct evidence of a stable DMG intermediate was identified in single-chain monellin, in which a combination of intramolecular distance measurements by fluorescence resonance energy transfer paired with GdnHCl-population of the DMG[62] led to reevaluation of the DMG as a discrete step in the folding or unfolding of proteins[61]. Subsequent work has even identified a modified form of the DMG: the transition state between the native and DMG states of the villin headpiece subdomain (HP35) is an expanded DMG.[63]

Lastly, intrinsically disordered proteins make up another class of highly flexible

conformational states of proteins. Comparison between disordered structures and intrinsic coils, both lacking secondary structure, may appear MG-like; however, the maintenance of secondary structure in the MG sets it apart.

1.7 Conclusions

Experimental approaches to fully characterize protein structures need to be sensitive to a large array of ordered and disordered states and are ideally site-specific to map conformational heterogeneity. Changes in secondary and tertiary structure, flexibility of intraprotein domains, and conformational exchange are all dynamic processes that confer function. Site-directed spin labeling (SDSL) in combination with electron paramagnetic resonance (EPR) provides information on protein motions occuring on the ps-ms time scale, which captures the mentioned diverse range of structural fluctuations. The EPR spectrum is sensitive to local structure about the spin label, and the origin of the absorbance line shape of the spectrum is discussed in Chapter 2.

1.8 Bibliography

1. Bond, A. D. Why do we trust X-ray crystallography? *Resonance* **19**, 1087–1092 (2015).
2. Zheng, H. *et al.* The future of crystallography in drug discovery. *Expert Opin. Drug Discov.* **9**, 125–137 (2015).
3. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch. Chem. Gesellschaft* **27**, 2985–2993 (1894).
4. Fischer, E. Bedeutung der Stereochemie für die Physiologie. in *Untersuchungen Über Kohlenhydrate und Fermente* 82–83 (Hoppe-Seyler's Z. Physio. Chem., 1898).

5. Lumry, R. Some aspects of the thermodynamics and mechanism of enzymic catalysis. in *The Enzymes* (eds. Boyer, P. D., Lardy, H. A. & Myrbäck, K.) 157 (Academic Press, 1959).

6. Koshland, D. E. J. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **44**, 98 (1958).

7. Haurowitz, F. Das Gleichgewicht zwischen Hämoglobin und Sauerstoff. *Hoppe. Seylers. Z. Physiol. Chem.* **254**, 266–274 (1938).

8. Kendrew, J. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666 (1958).

9. Yankeelov, J. A. J. & Koshland, D. E. J. Evidence for Conformation Changes Induced by Substrates of Phosphoglucomutase. *J. Biol. Chem.* **240**, 1593–1602 (1965).

10. Pocker, Y. & Stone, J. T. The Catalytic Versatility of Erythrocyte Carbonic Anhydrase. The Enzyme-Catalyzed Hydrolysis of p-Nitrophenyl Acetate. *J. Am. Chem. Soc.* **87**, 5497–5498 (1965).

11. Reid, T. W. & Fahrney, D. The Pepsin-Catalyzed Hydrolysis of Sulfite Esters. *J. Am. Chem. Soc.* **89**, 3941–3943 (1967).

12. Nakagawa, Y. & Bender, M. L. Modification of α-Chymotrypsin by Methyl p-Nitrobenzenesulfonate. *J. Am. Chem. Soc.* **91**, 1566–1567 (1969).

13. Allerhand, A. *et al.* Conformation and Segmental Motion of Native and Denatured Ribonuclease a in Solution. Application of Natural-Abundance Carbon-13 Partially Relaxed Fourier Transform Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **93**, 544–546 (1971).

14. Hvidt, A. & Nielsen, S. O. Hydrogen Exchange in Proteins. *Adv. Protein Chem.* **21**, 287–

386 (1966).

15. Lakowicz, J. R. & Weber, G. Quenching of Protein Fluorescence by Oxygen. Detection of Structural Fluctuations in Proteins on the Nanosecond Time Scale. *Biochemistry* **12**, 4171–4179 (1973).

16. Cooper, A. Thermodynamic fluctuations in protein molecules. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 2740–2741 (1976).

17. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. Dynamics of Ligand Binding to Myoglobin. *Biochemistry* **14**, 5355–5373 (1975).

18. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The Energy Landscapes and of Proteins Motions. *Science (80-.).* **254**, 1598–1603 (1991).

19. Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).

20. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).

21. Gerstein, M., Lesk, A. M. & Chothia, C. Structural Mechanisms for Domain Movements in Proteins. *Biochemistry* **33**, 6739–6749 (2005).

22. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci.* **97**, 8868–8873 (2000).

23. Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558–563 (1979).

24. Mulder, F. A. A., Mittermaier, A., Hon, B., Dahlquist, F. W. & Kay, L. E. Studying

excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 932–935 (2001).

25. Kalodimos, C. G. NMR reveals novel mechanisms of protein activity regulation. *Protein Sci.* **20**, 773–782 (2011).

26. Zwanzig, R., Szabo, A. & Bagchi, B. Levinthal's paradox. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 20–2 (1992).

27. Denisov, V. P., Jonsson, B. H. & Halle, B. Hydration of denatured and molten globule proteins. *Nat. Struct. Biol.* **6**, 253–260 (1999).

28. Farrell, H. M. *et al.* Molten Globule Structures in Milk Proteins: Implications for Potential New Structure-Function Relationships. *J. Dairy Sci.* **85**, 459–471 (2010).

29. Mora-Gutierrez, A., Kumosinski, T. F. & Farrell, H. M. Oxygen-17 Nuclear Magnetic Resonance Studies of Bovine and Caprine Casein Hydration and Activity in Deuterated Sugar Solutions. *J. Agric. Food Chem.* **45**, 4545–4553 (1997).

30. Modig, K., Kurian, E., Prendergast, F. G. & Halle, B. Water and urea interactions with the native and. *Protein Sci.* **12**, 2768–2781 (2003).

31. Zhou, R., Li, J., Hua, L., Yang, Z. & Berne, B. Comment on 'Urea-mediated protein denaturation: A consensus view'. *J. Phys. Chem. B* **113**, 12816–12824 (2009).

32. Das, A. & Mukhopadhyay, C. Urea-Mediated Protein Denaturation: A Consensus View. *J Phys Chem B* **113**, 12816–12824 (2009).

33. Hughson, F. M., Wright, P. E. & Baldwin, R. L. Structural characterization of a partly folded apomyoglobin intermediate. *Science (80-.).* **249**, 1544–1548 (1990).

34. Kitahara, R., Yamada, H., Akasaka, K. & Wright, P. E. High pressure NMR reveals that apomyoglobin is an equilibrium mixture from the native to the unfolded. *J. Mol. Biol.* **320**, 311–319 (2002).

35. Lerch, M. T., Yang, Z., Brooks, E. K. & Hubbell, W. L. Mapping protein conformational heterogeneity under pressure with site-directed spin labeling and double electron–electron resonance. *Proc. Natl. Acad. Sci.* **111**, E1201–E1210 (2014).

36. Kuwajima, K. The molten state of a-lactalhunun. *Faseb* **10**, 102–109 (1996).

37. Okazaki, A., Nikaido, K. & Kuwajima, K. The chaperonin GroEL does not recognize apo-α-lactalbumin in the molten globule state. *Nat. Struct. {&} Mol. Biol.* **1**, 439–446 (1994).

38. Ferrer, M., Barany, G. & Woodward, C. Partially folded, molten globule and molten coil states of bovine pancreatic trypsin inhibitor. *Nat. Struct. Biol.* **2**, 211–217 (1995).

39. Ikeguchi, M., Kato, S. I., Shimizu, A. & Sugai, S. Molten globule state of equine β-lactoglobulin. *Proteins Struct. Funct. Genet.* **27**, 567–575 (1997).

40. Yang, J., Dunker, A. K., Powers, J. R., Clark, S. & Swanson, B. G. β-Lactoglobulin molten globule induced by high pressure. *J. Agric. Food Chem.* **49**, 3236–3243 (2001).

41. Griko, Y. V. & Privalov, P. L. Calorimetric Study of the Heat and Cold Denaturation of β-Lactoglobulin. *Biochemistry* **31**, 8810–8815 (1992).

42. Flemming, P. M., Jeffrey, H. C. & Christopher, D. M. Structural Study of the Hydrophobic Box Region of Lysozyme in Solution Using Nuclear Overhauser Effects. *Biochemistry* **19**, 2597–2607 (1980).

43. Koga, K. & Berliner, L. J. Structural Elucidation of a Hydrophobic Box in Bovine α-Lactalbumin by NMR: Nuclear Overhauser Effects. *Biochemistry* **24**, 7257–7262 (1985).

44. Pedrote, M. M. *et al.* Aggregation-primed molten globule conformers of the p53 core domain provide potential tools for studying p53C aggregation in cancer. *J. Biol. Chem.* **293**, 11374–11387 (2018).

45. Skora, L., Becker, S. & Zweckstetter, M. Molten globule precursor states are

conformationally correlated to amyloid fibrils of human β-2-microglobulin. *J. Am. Chem. Soc.* **132**, 9223–9225 (2010).

46. Vamvaca, K., Jelesarov, I. & Hilvert, D. Kinetics and Thermodynamics of Ligand Binding to a Molten Globular Enzyme and Its Native Counterpart. *J. Mol. Biol.* **382**, 971–977 (2008).

47. Hammes, G. G., Chang, Y.-C. & Oas, T. G. Conformational selection or induced fit: A flux description of reaction mechanism. *Proc. Natl. Acad. Sci.* **106**, 13737–13741 (2009).

48. Kjaergaard, M., Teilum, K. & Poulsen, F. M. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci.* **107**, 12535–12540 (2010).

49. Kitazawa, S. *et al.* Solution structure of the Q41N variant of ubiquitin as a model for the alternatively folded N2 state of ubiquitin. *Biochemistry* **52**, 1874–1885 (2013).

50. Tam, R. & Saier, M. H. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* **57**, 320–46 (1993).

51. Prajapati, R. S., Indu, S. & Varadarajan, R. Identification and thermodynamic characterization of molten globule states of periplasmic binding proteins. *Biochemistry* **46**, 10339–10352 (2007).

52. Munshi, P. *et al.* Molecular details of ligand selectivity determinants in a promiscuous β-glucan periplasmic binding protein. *BMC Struct. Biol.* **13**, (2013).

53. Gu, W., Shi, X.-L. & Roeder, R. Synergistic Activation of Transcription by. *Lett. to Nat.* **387**, 819–823 (1997).

54. Burge, S. *et al.* Molecular basis of the interactions between the p73 N terminus and p300: Effects on transactivation and modulation by phosphorylation. *Proc. Natl. Acad. Sci.* **106**,

3142–3147 (2009).

55. Li, H. & Chen, J. D. The receptor-associated coactivator 3 activates transcription through CREB-binding protein recruitment and autoregulation. *J. Biol. Chem.* **273**, 5948–5954 (1998).

56. Lin, R., Heylbroeck, C., Pitha, P. M. & Hiscott, J. Virus-Dependent Phosphorylation of the IRF-3 Transcription Factor Regulates Nuclear Translocation, Transactivation Potential, and Proteasome-Mediated Degradation. *Mol. Cell. Biol.* **18**, 2986–2996 (2015).

57. Katsumata, K., Okazaki, A., Tsurupa, G. P. & Kuwajima, K. Dominant forces in the recognition of a transient folding intermediate of α-lactalbumin by GroEL. *J. Mol. Biol.* **264**, 643–649 (1996).

58. Okazaki, A., Katsumata, K. & Kuwajima, K. Hydrogen-Exchange Kinetics of Reduced α-Lactalbumin Bound to the Chaperonin GroEL. *J Biochem* **121**, 534–541 (1997).

59. Vamvaca, K., Vogeli, B., Kast, P., Pervushin, K. & Hilvert, D. An enzymatic molten globule: Efficient coupling of folding and catalysis. *Proc. Natl. Acad. Sci.* **101**, 12860–12864 (2004).

60. Lazaridis, T., Archontis, G. & Karplus, M. Enthalpic Contribution to Protein Stability: Insights from Atom-Based Calculations and Statistical Mechanics. *Adv. Protein Chem.* **47**, 231–306 (1995).

61. Baldwin, R. L., Frieden, C. & Rose, G. D. Dry molten globule intermediates and the mechanism of protein unfolding. *Proteins Struct. Funct. Bioinforma.* **78**, 2725–2737 (2010).

62. Jha, S. K. & Udgaonkar, J. B. Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Acad. Sci.* **106**, 12289–12294 (2009).

63. Neumaier, S. & Kiefhaber, T. Redefining the dry molten globule state of proteins. *J. Mol. Biol.* **426**, 2520–2528 (2014).

64. Baldwin, R. L. & Rose, G. D. Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* **23**, 4–10 (2013).

65. Chaffotte, A., Guillem, Y., Goldberg, M. E., Delepierre, M. & Hinz, H. J. The Isolated C-Terminal (F2) Fragment of the Escherichia coli Tryptophan Synthase β2-Subunit Folds into a Stable, Organized Nonnative Conformation. *Biochemistry* **30**, 8067–8074 (1991).

66. Chaffotte, A. F., Iñaki Guijarro, J., Guillou, Y., Delepierre, M. & Goldberg, M. E. The 'pre-molten globule,' a new intermediate in protein folding. *J. Protein Chem.* **16**, 433–439 (1997).

67. Shakhnovich, E. I. & Finkelstein, A. V. Theory of cooperative transitions in protein molecules. *Biopolymers* **28**, 1667–1680 (1989).

68. Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505 (2010).

Chapter 2. Electron paramagnetic resonance

This chapter discusses the electron paramagnetic resonance (EPR) spectroscopy technique and the origin of the EPR line shape. It begins with a description of the discovery of the EPR phenomenon and then discusses how protein structure and dynamics studies benefit from use of the combined techniques of EPR and site-directed spin labeling. The reader is provided a general description of magnetic resonance focused on electronic spin state transitions. This section draws from many texts wherein the same important equations are given: see Carrington and McLachlan[1], Marsh[2], and Poole[3]. This chapter continues with a description of the instrumental layout of a continuous-wave (CW) EPR spectrometer and outlines various experimental approaches that take advantage of the EPR spectral line shape (both CW and pulsed EPR) to measure protein structure and dynamics on various time and length scales.

2.1 Discovery of EPR

The EPR phenomenon was discovered by Yvengy K. Zavoisky in 1944.[4] In the pioneering work, he described the absorption of microwave radiation by manganese salts in a static applied magnetic field. By varying field strength, radiation absorbed by $MnSO_4$ was maximal at a particular value of ν/H_0 for the sample. This ratio of applied microwave radiation frequency to magnetic field strength indicates the requirements to induce magnetic dipole transitions between electronic spin states. EPR is a highly sensitive, non-destructive technique that measures the local electronic properties of unpaired spins. Analysis of magnetic resonance behavior for various paramagnetic ions and molecules is centrally important in a multitude of studies across the physical and biological sciences. Application of EPR to biological systems in particular lends a unique perspective to molecular structure and dynamics.

2.2 Origin of the EPR absorbance line

The following sections introduce basic magnetic resonance theory to describe the origin of the EPR absorption line due to microwave absorption by electrons transitioning between two nondegenerate spin states in an applied magnetic field. From this discussion, the motional and structural properties of R1 are shown to contribute to unique EPR line shapes.

2.2.1 The resonance condition

An electron is a fermion—an elementary particle with half-integral spin—that has a permanent magnetic moment due to contributions from its spin angular momentum S and orbital angular momentum L. The vector sum of these two momenta (the total angular momentum $J = S + L$) can be related to the electronic and magnetic dipole moment

$$\boldsymbol{\mu}_e = g\beta_e \boldsymbol{J} \tag{1}$$

where g is the unitless g-factor, β_e is the Bohr magneton, defined as $\beta_e = |z|\hbar/2m_e c^2$, where $z = -1$ and $m_e = 9.1\times10^{31}$ kg are the charge and mass of an electron, respectively, c is the speed of light in a vacuum, and $\hbar$ is Planck's constant divided by 2π. The orbital angular momentum L can be approximately defined as the classical motion of an electron in an atomic orbital on a nucleus. In the simplest case, an electron in the spherically symmetric *s*-orbital has an average orbital angular momentum of zero and the spin angular momentum is the dominant contributor to the magnetic moment ($J \approx S$). Spin angular momentum is defined as

$$|S| = \hbar\sqrt{s(s+1)} \tag{2}$$

An electron has a spin quantum number of $s = 1/2$ and, as such, can exist in $(2s + 1)$ two different spin states: $m_s = +1/2$ and -1/2. For an electron, Eq. 2 becomes

$$|S_e| = \frac{\sqrt{3}}{2}\hbar \tag{3}$$

and as a result, the magnetic dipole of an electron can be re-expressed as

$$\boldsymbol{\mu}_e = g\beta_e \boldsymbol{J} \approx g_e \beta_e \boldsymbol{S}_e = \frac{\sqrt{3}\hbar^2 g_e}{4m_e c^2} \tag{4}$$

where g_e is the electronic g-factor (2.00232 for a free electron). Note that the vectorial direction of the magnetic moment and the spin angular momentum are antiparallel for an electron due to its negative charge.

While degenerate in a zero-field environment, in the presence of an external magnetic field H_0, the energies of the m_s = +1/2 and -1/2 spin states diverge as defined by

$$E_- = -\frac{1}{2} g_e \beta_e H_0 \tag{5a}$$

$$E_+ = \frac{1}{2} g_e \beta_e H_0 \tag{5b}$$

where the lower-energy state E_- corresponds to the m_s = -1/2 spin state, the higher-energy state E_+ corresponds to the m_s = +1/2 spin state, and the energy difference between these two states is simply

$$\Delta E = E_- - E_+ = g_e \beta_e H_0 \tag{6}$$

The Zeeman effect describes this relationship between the energy difference and field strength, which is due to the interaction between the electron spin angular momentum and the external field (Fig. 2.1). Transitions between the m_s = +1/2 and -1/2 spin states can be induced by microwave radiation at a frequency specified by the Planck equation, $\Delta E = h\nu$.

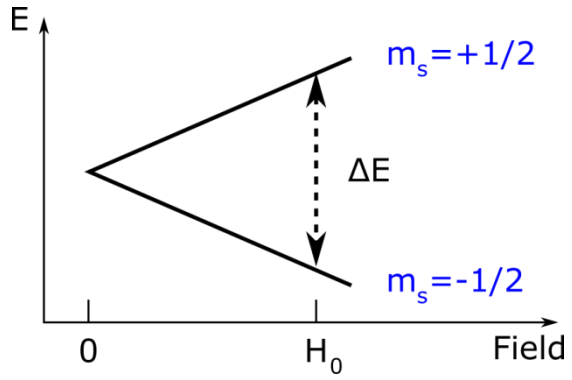

Figure 2.1 Diagrammatic representation of the Zeeman effect. Degenerate spin states in the absence of a magnetic field $\boldsymbol{H_0}$ are split as field strength increases. The spins directed with and against the field have an energy difference of $\boldsymbol{\Delta E}$.

In EPR spectroscopy, the external field H_0 is applied along the laboratory z-axis by convention. Excitation of electron spins occurs due to the weak oscillating field H_1 created in the xy-plane (perpendicular to H_0) from the incident microwave radiation. Transitions between the two spin states occur only when the frequency of electromagnetic radiation satisfies the resonance condition, a combination of Eq. 6 and the Planck equation, in the form

$$h\nu = g_e \beta_e H_0 \tag{7}$$

From this equation we see that either the magnetic field strength or frequency of radiation can be varied in order to find the resonance condition for an electron, shown in Fig. 2.1. In practice, CW EPR spectroscopy employs a magnetic field sweep in the presence of a constant excitation frequency (a "continuous wave") in order to locate the resonant field-frequency pair. Typically, EPR is performed at "X-band" microwave frequencies (~8-10 GHz), requiring field strengths—approximated by substituting the values of the free electron g-factor and Bohr magneton (9.285×10^{-28} J G^{-1}) into Eq. 7—on the order of 2850 to 3560 G. Other common EPR frequencies are shown in Table 2.1.

Table 2.1 Commonly used CW EPR frequencies and the field strengths required at resonance

Frequency	Microwave band	Field strength (G)	Field strength (T)
2.0-4.0 (commonly 4.0)	S	1400	0.140
8.0-10.0 (commonly 9.75)	X	3470	0.340
34.0	Q	12500	1.250
95.0	W	34000	3.400

2.2.2 Stimulated and spontaneous transitions

In all forms of spectroscopy, net absorption is observed only when the equilibrium populations of the two excitable states are not equal. In the presence of a static magnetic field, not only are the energies of the two electron spin states different, but the relative populations of

the electrons with spins parallel and antiparallel to the field are also different. At equilibrium, the relative populations of spins can be calculated by their respective energies and temperature according to the Boltzmann distribution

$$\frac{N_+}{N_-} \equiv e^{\frac{-\Delta E}{k_B T}} = e^{\frac{-g_e \beta_e H_0}{k_B T}} \quad (8)$$

Even in a practical experimental sense, because H_1 is weak, a system studied by CW EPR will be close to thermodynamic equilibrium, such that $\Delta N \neq 0$ and $N_+ \,/\, N_- \neq 1$. The signal strength of an EPR experiment depends on the size of the population split, and as such the signal strength is related to the energy difference (and field strength H_0). For X-band EPR at room temperature, $N_+/N_- = 0.9984$. The net difference amounts to only 0.08% of the total number of spins. Additionally, the sensitivity of CW EPR is highlighted when comparing the energy of microwave photons used in EPR experiments to that of thermal noise, which is on the order of ~1 $k_B T$, where $k_B = 1.38\times10^{-23}$ J K^{-1}.[5] At room temperature, this corresponds to roughly 2.5 kJ mol^{-1} of energy. In EPR, microwave photons at X-band are on the order of 4.0 J mol^{-1}, which is three orders of magnitude smaller than thermal noise. It is a testament to the sensitivity of the EPR technique that such measurements are even possible. EPR spectroscopy relies on application of the oscillating field H_1 to redistribute the relative N_- and N_+ populations, giving rise to the absorbance spectrum. To restore the equilibrium spin populations over time, a spontaneous transition process exists. Because the local thermal motions of the spin system depend on the energy of the lattice environment, the upward and downward spontaneous transition probabilities, W_{-+} and W_{+-}, are termed spin-lattice relaxation probabilities. We can express the rate of change in N_- due to spin-lattice relaxation as

$$\frac{dN_-}{dt} = N_- W_{-+} - N_+ W_{+-} \quad (9)$$

If one defines the population split as $\Delta N = N_- - N_+$, the spin state populations at any time can

be expressed as

$$N_- = N_+ + \Delta N = \frac{N_{tot}+\Delta N}{2} \tag{10a}$$

$$N_+ = N_- - \Delta N = \frac{N_{tot}-\Delta N}{2} \tag{10b}$$

By combining Eq. 9 with 10a and 10b the rate of change for the low energy state population is

$$\frac{dN_-}{dt} = \frac{1}{2}\frac{d\Delta N}{dt} = \frac{1}{2}[N_{tot}(W_{-+} - W_{+-}) - \Delta N(W_{-+} + W_{+-})] \tag{11}$$

and accordingly the rate of change for the population difference is

$$\frac{d\Delta N}{dt} = N_{tot}(W_{-+} - W_{+-}) - \Delta N(W_{-+} + W_{+-}) \tag{12}$$

which can be rewritten by substituting the population difference at thermal equilibrium $\Delta N_0 = N_{tot}[(W_{-+} - W_{+-})/(W_{-+} + W_{+-})]$ into Eq. 12 to give

$$\frac{d\Delta N}{dt} = -(\Delta N - \Delta N_0)(W_{-+} + W_{+-}) = -\frac{\Delta N - \Delta N_0}{T_1} \tag{13}$$

Combining the transition probabilities into a new time quantity defines the spin-lattice relaxation time, $T_1 = (W_{-+} + W_{+-})^{-1}$, which is of particular importance in EPR, as it describes the time required to establish thermal equilibrium of spin populations via processes that transfer the energy to the lattice (in some texts this is described as a small heat exchange), and which involes local thermal motions of the spin system. It is environment- and system-dependent and determines the lifetime of N_+.

In contrast to the equilibrating force due to spontaneous transitions, stimulated transitions are due to the H_1 excitation field. The rate of change of the N_- population, for example, depends on a new transition probability P. Note that P is equal in both directions ($P = P_{-+} = P_{+-}$); although this may seem intuitive, it is shown explicitly using expressions derived from time-dependent perturbation theory found in Carrington and McLachlan.[1] Fig. 2.2 shows a comparison of the P and W transitions between spin states.

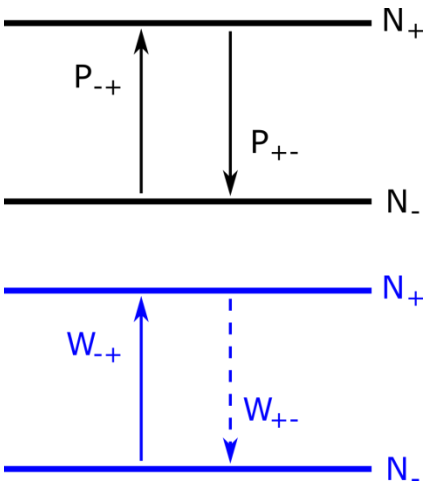

Figure 2.2 Top, stimulated transitions in the upward and downward directions are equally probable, due to an applied $\boldsymbol{H_1}$ in the presence of a static field $\boldsymbol{H_0}$. Bottom, spontaneous reequilibration of spins due to spin-lattice relaxation drives a nonequilibrium spin population back to the Boltzmann distribution due to asymmetric exchange frequencies between the two states.

In a similar treatment, the rate of change of the low energy state population is given by

$$\frac{dN_-}{dt} = \frac{1}{2}\frac{d\Delta N}{dt} = P(N_+ - N_-) \tag{14}$$

and a new expression for the rate of change in the population split is

$$\frac{d\Delta N}{dt} = -2P\Delta N \tag{15}$$

When solved, this differential equation yields an exponential decay behavior

$$\Delta N = \Delta N_0 e^{-2Pt} \tag{16}$$

An expression for absorbed energy over time is

$$\frac{d\Delta E}{dt} = N_- P(E_+ - E_-) + N_+ P(E_- - E_+) = \Delta N P \Delta E \tag{17}$$

This expression indicates no absorption will occur ($\Delta E = 0$) when there is no difference between spin state populations ($\Delta N = 0$), i.e., basic absorption spectroscopy is not possible when a resonance line is "saturated." In a hypothetical situation where only stimulated spin transitions are possible, an EPR signal would decay and disappear permanently in the presence of the oscillating field H_1. Clearly, this would not be a useful method for spectroscopic measurements! The combined action of stimulated and spontaneous transitions produces a continuously measurable absorption spectrum. A complete expression for the rate of absorption of energy due

to additive stimulated and spontaneous transition processes is given by combining Eq. 13 and 15

$$\frac{d\Delta N}{dt} = -2P\Delta N - \frac{\Delta N - \Delta N_0}{T_1} \tag{18}$$

such that at equilibrium ($d\Delta N/dt = 0$), the population difference in the presence of H_1 is

$$\Delta N = \frac{\Delta N_0}{1+2PT_1} \tag{19}$$

Combining Eq. 17 and 19 gives a final expression for the rate of absorption of energy

$$\frac{dE}{dt} = \Delta N P \Delta E = \Delta N_0 \frac{P}{1+2PT_1} \tag{20}$$

Note that when $2PT_1 \ll 1$, the system is not saturated. CW EPR generally requires nonsaturating conditions unless the saturation behavior is desired to obtain information about T_1; this experiment is called power saturation EPR.[3] Finally, CW (field-swept) EPR is only possible because both spontaneous and stimulated transitions are concerted and, in practice, competing; that is, non-radiative stimulated transitions can be induced between the $m_s = +1/2$ and -1/2 spin states, while spontaneous heat dissipation of the system into the lattice reequilibrates spin populations back to Boltzmann equilibrium over time.

2.2.3 Bulk magnetization

So far, the spin resonance experiment has been shown to produce a sustained absorption line over time at the resonance condition. A more complete picture of spin relaxation is necessary to account for the observed line broadening of resonances. A macroscopic or bulk description of spins in a magnetic field is represented by the net magnetization vector $\boldsymbol{M}$, which has M_x, M_y, and M_z components, and is the sum total of the magnetic dipole moments of a set of spins. In the absence of a magnetic field, the projection of $\boldsymbol{M}$ in the z-axis (which would be aligned with H_0 should it exist) is

$$M_z = \gamma_e \hbar (N_- - N_+) \tag{21}$$

where γ_e is the magnetogyric ration of the electron. In this case, $\Delta N_0 = 0$ and the z-axis net magnetization, M_z, is proportional to ΔN, such that its substitution into Eq. 13 gives

$$\frac{dM_z}{dt} = \frac{-M_z}{T_1} \tag{22}$$

In the presence of H_0, the equilibrium term of Eq. 13 must be reintroduced to include the magnitude of equilibrium magnetization $M_0 = \gamma_e \hbar \Delta N_0$, which is

$$\frac{dM_z}{dt} = \frac{M_0 - M_z}{T_1} \tag{23}$$

In contrast to the behavior of the z-axis magnetization, the orthogonal components of $\boldsymbol{M}$ (M_x and M_y) decay exponentially to zero with time because H_0 is perfectly aligned to the z-axis. A second relaxation time T_2 describes the decay due to M_x, M_y dependence on transverse relaxation processes with the following form:

$$\frac{dM_x}{dt} = \frac{-M_x}{T_2} \tag{24a}$$

$$\frac{dM_y}{dt} = \frac{-M_y}{T_2} \tag{24b}$$

While spin-lattice (T_1) relaxation stems from interactions between excited spins and the lattice, spin-spin (T_2) relaxation accounts for processes independent of energetic spin transitions where, as described below, spin dephasing occurs.

By virtue of its magnetic moment and spin angular momentum, an electron (and therefore the bulk magnetization) experiences a torque $\boldsymbol{M} \times H_0$ due to the static magnetic field. This torque imperfectly aligns the spin magnetization vector with the field, with non-zero spin projections in the xy-plane, causing the spin vector to "precess" about the z-axis in a conical shape at the Larmor frequency $\omega_0 = \gamma_e H_0$. The torque exerted on the spin angular momentum is

$$\frac{dM}{dt} = \gamma_e (\boldsymbol{M} \times H_0) \tag{25}$$

Due to their negative charge, electrons precess counterclockwise about an applied magnetic field (when viewed from above the xy-plane). This precession is illustrated in Fig. 2.3.

The time dependence of bulk magnetization in the presence of the external field is now a combination of spin-lattice and spin-spin relaxation (Eq.s 24, 25a, and 25b) and this new mechanism (Eq. 25) to yield the Bloch equation[6]

$$\frac{d\boldsymbol{M}}{dt} = \gamma_e(\boldsymbol{M} \times H_0) - i\frac{M_x}{T_2} - j\frac{M_y}{T_2} - k\frac{M_0 - M_z}{T_1} \tag{26}$$

where i, j, and k are the x, y, and z unit vectors in the laboratory frame. Returning to Eq. 24 and solving it for the precession of $\boldsymbol{M}$ about a static field H_0 in the z-axis yields three expressions:

$$\frac{dM_z}{dt} = 0 \tag{27a}$$

$$\frac{dM_x}{dt} = \gamma_e H_0 M_y = \omega_0 M_y \tag{27b}$$

$$\frac{dM_y}{dt} = -\gamma_e H_0 M_x = \omega_0 M_x \tag{27c}$$

which can be combined with the expressions for the rate of magnetization decay due to T_1 and T_2 relaxation to derive a less general form of the Bloch equation in three parts:

$$\frac{dM_x}{dt} = \omega_0 M_y - \frac{M_x}{T_2} \tag{28a}$$

$$\frac{dM_y}{dt} = -\omega_0 M_x - \frac{M_y}{T_2} \tag{28b}$$

$$\frac{dM_z}{dt} = 0 - \frac{M_0 - M_z}{T_1} \tag{28c}$$

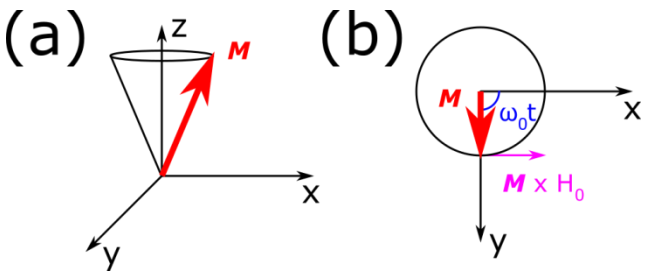

Figure 2.3 Precession of spins about the z-axis due to the static field $\boldsymbol{H_0}$. In (a), precession of spins produces a net magnetization predominately in the z-axis. In (b), the perpendicular component of $\boldsymbol{M}$ in the xy-plane precesses about the z-axis in the laboratory frame. Here, the instantaneous rotating frame is aligned with the laboratory frame.

It is useful to consider the transverse components of this system in the xy-plane together as the perpendicular magnetization $\boldsymbol{M}_{\perp}$

$$\boldsymbol{M}_{\perp} = \frac{iM_x + jM_y}{T_2} \tag{29}$$

while the longitudinal component in the z-axis, $M_{\parallel}$, parallel to H_0, retains the following magnetization

$$M_{\parallel} = \frac{k(M_0 - M_z)}{T_1} \tag{30}$$

which separates the T_1 and T_2 relaxation processes into two rather than three terms.

In summary, lifetime of transverse relaxation in the xy-plane, $d\boldsymbol{M}_{\perp}/dt$, is determined by the time T_2, which describes spin-spin relaxation of magnetic moments orthogonal to H_0. The net transverse magnetization $\boldsymbol{M}_{\perp}$ decays to zero via T_2 (spin-spin) relaxation while the magnetization in the z-axis (M_z or $M_{\parallel}$) approaches an equilibrium value M_0 via T_1 (spin-lattice) relaxation. The mechanism of T_2 relaxation has remained ambiguous in this discussion and has so far been described as the effect of dephasing in the xy-plane, for which the bulk $\boldsymbol{M}_{\perp}$ vanishes. These spins precess indefinitely, although the relative phases of individual spins change with time. T_2 relaxation is attributed to variation or fluctuation in local fields in the xy-plane, but is often considered a "catch-all" to describe any and all relaxation processes that result in loss of phase coherence. For solution-phase samples, motional contributions to T_2 relaxation are typically dominated by molecular tumbling. Additional mechanisms such as collisional spin-spin interactions via Heisenberg exchange, dipole-dipole interactions between the electron magnetic moment and another dipole[7], and orientation-dependence of anisotropic nitroxide motion also result in loss of phase coherence.[1]

2.2.4 Lorentz line shape

Larmor precession plays an additional role in the efficiency of the weak oscillating field H_1. When the H_1 field frequency matches the Larmor frequency, ω_0, a large oscillating magnetic moment is produced and the maximum of bulk spins meet the resonance condition. Just as in the case of H_0, H_1 produces a torque on the bulk magnetization. The Bloch equation (Eq. 25) now becomes

$$\frac{d\boldsymbol{M}}{dt} = \gamma_e(\boldsymbol{M} \times H_1) + \gamma_e(\boldsymbol{M} \times H_0) - i\frac{M_x}{T_2} - j\frac{M_y}{T_2} - k\frac{M_{eq} - M_z}{T_1} \tag{31}$$

the oscillation of H_1 in the rotating frame ($H_1{}'$) is given by

$$H_1{}' = H_1[\cos(\omega_H t) - j\sin(\omega_H t)] \tag{32}$$

where the field angular velocity ω_H is not necessarily the Larmor angular velocity ω_0. The slight deviation in angular velocities can be observed in a rotating frame that follows H_1, where H_1 is aligned with the +x′-axis in the x′,y′,z′ axis system. The xy- and x′y′-planes are coplanar and z- and z′-axes are parallel. In this system, the rate of change in $\boldsymbol{M'}$ can be expressed as

$$\frac{\delta \boldsymbol{M'}}{\delta t} = i'\frac{dM_{x'}}{dt} + j'\frac{dM_{y'}}{dt} + k'\frac{dM_z}{dt} \tag{33}$$

where i', j', and k' are the unit vectors in the rotating frame. This apparent rate of change of the bulk magnetization $\boldsymbol{M}$ is related to the oscillating magnetization $\boldsymbol{M'}$ in the rotating frame by

$$\frac{d\boldsymbol{M}}{dt} = \frac{\delta \boldsymbol{M'}}{\delta t} + \omega_H{}' \times \boldsymbol{M} \tag{34}$$

where the angular velocity $\omega_H{}'$ appears to change with time for an observer in the laboratory frame. The transverse components of $\boldsymbol{M}$ can be calculated from these reference values to give

$$\frac{dM_{x'}}{dt} = (\omega_0 - \omega_H)M_y{}' - \frac{M_x}{T_2} \tag{35a}$$

$$\frac{dM_{y'}}{dt} = -(\omega_0 - \omega_H)M_x{}' + \gamma_e H_1 M_z - \frac{M_{y'}}{T_2} \tag{35b}$$

$$\frac{dM_z}{dt} = -\gamma_e H_1 M_y' - \frac{M_{eq} - M_z}{T_1} \tag{35c}$$

which resemble Eqs. 28a, 28b, and 28c. The solutions to these equations at equilibrium is determined by $dM_x'/dt = dM_y'/dt = dM_z'/dt = 0$

$$M_x' = \frac{\gamma_e H_1 T_2^2 (\omega_0 - \omega_H) M_{eq}}{1 + T_2^2(\omega_0 - \omega_H)^2 + \gamma_e^2 H_1^2 T_1 T_2} \tag{36a}$$

$$M_y' = \frac{\gamma_e H_1 T_2 M_{eq}}{1 + T_2^2(\omega_0 - \omega_H)^2 + \gamma_e^2 H_1^2 T_1 T_2} \tag{36b}$$

$$M_z = \frac{1 + T_2^2(\omega_0 - \omega_H)^2}{1 + T_2^2(\omega_0 - \omega_H)^2 + \gamma_e^2 H_1^2 T_1 T_2} \tag{36c}$$

Re-expression of these quantities in the stationary laboratory frame show oscillatory behavior with a counterclockwise-rotating field. The transverse components of Eqs. 36b and 36c can be re-expressed as follows:

$$M_x = M_x'[i\cos(\omega_H t) - j\sin(\omega_H t)] \tag{37a}$$

$$M_y = M_y'[i\cos(\omega_H t + \pi) - j\sin(\omega_H t + \pi)] \tag{37b}$$

From these relationships it is possible to derive an expression for the magnetic susceptibility $X''(\omega_H)$ in the laboratory frame in order to write an expression for the average power absorbed. While this derivation is not covered here, it can be found in the text by Carrington and McLachlan[1] and is of the form

$$\frac{dE}{dt} = 2\omega_H H_1^2 \chi''(\omega_H) = \frac{H_1^2 \chi \omega_0 \omega_H (T_2)}{1 + T_2^2(\omega_0 - \omega_H) + \gamma_e^2 H_1^2 T_1 T_2} \tag{38}$$

In an unsaturated state, the quantity $\gamma_e^2 H_1^2 T_1 T_2$ is small such that the absorption is in the form of a well-known probability distribution function of Lorentzian form: $L = a(1 + bx^2)^{-1}$

$$g(\omega_H) = \frac{T_2}{\pi} \frac{1}{1 + T_2^2(\omega_H - \omega_0)^2} \tag{39}$$

The absorption full width at half maximum (FWHM) is $\Delta H = 2/\gamma_e T_2$. This characteristic line shape is shown in Fig. 2.4. The field-swept EPR line shape is acquired through detection of the

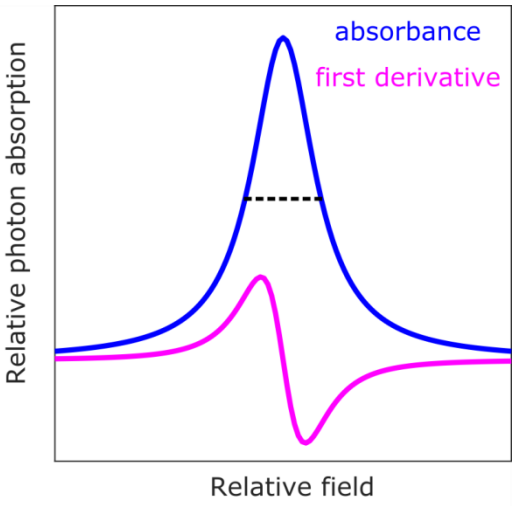

Figure 2.4 The Lorentz line shape (blue) and first derivative (magenta) are shown. The dashed line demonstrates the full width at half maximum, a function of $\boldsymbol{T_2}$.

transverse magnetization $M_y{}'$ in the presence of a small oscillating field $H_1{}'$ and a large stationary field H_0. Analogous to Eq. 7, the signal is maximal at the resonance condition. By re-expressing Eq. 36b in terms of field strength

$$M_y{}' = \frac{\gamma_e H_1 T_2 M_{eq}}{1+\gamma_e^2 T_2^2 (H_0 - H)^2 + \gamma_e^2 H_1^2 T_1 T_2} \tag{40}$$

it is clear that a maximum is found when $H = H_0$. Under non-saturating conditions, the absorption spectrum is primarily dependent on T_2—not T_1—because H_1 is weak, so the rightmost quantity in the denominator disappears.

Field modulation and phase-sensitive detection are common methods in EPR spectroscopy to enhance signal-to-noise and cause the signal to be measured as a first derivative of the absorption spectrum, as it will be discussed later. The modulated first-derivative signal has a peak to negative peak linewidth of $\Delta H_{pp} = 2/\sqrt{3s}\gamma_e T_2$, where the saturation factor, s, is generally 1 when the oscillating field H_1 is sufficiently weak. Stronger H_1 fields saturate spins, producing broadening effects due to the nonequilibrium distribution of spin states, because T_1 relaxation processes are too slow to re-equilibrate during the course of the continuous-wave (CW) EPR experiment at each field frequency. In fact, the microwave power dependence of

signal intensity can be exploited to determine T_1 and T_2 in a so-called power-saturation experiment[3].

2.2.5 The spin Hamiltonian

Resonant absorption of microwaves by an unpaired electron in a magnetic field produces a single absorption peak; the field position of this peak depends on the spectrometer frequency and electron g-factor as seen in Eq. 7. This dissertation focuses on EPR spectroscopy of nitroxide spin labels, a type of stable neutral free radical in which the unpaired electron is primarily localized in the $2p_z$ orbital of the nitrogen atom therein. The ^{14}N isotope of nitrogen has a nonzero nuclear spin ($I = 1$) with a natural abundance of 99.6%. Because the unpaired electron occupies a non-spherically symmetric orbital on the nitrogen atom, it experiences a spin-orbit coupling that yields a small, directionally-dependent orbital angular momentum (previously disregarded in the above derivations) that results in an anisotropic magnetic moment.

The energy of an electron is mathematically described by a two-part Hamiltonian, expressing the Zeeman and hyperfine energies. Additional terms come into play for interactions with additional paramagnetic centers via dipolar interactions or Heisenberg exchange and are not discussed here. As such, the following simplified total spin Hamiltonian defines the energy of the spin system adequately for our needs:

$$\widehat{H} = \widehat{H}_{Zeeman} + \widehat{H}_{hyperfine} \tag{41}$$

Note that the Zeeman interaction includes the effect of spin-orbit coupling introduced above.

2.2.6 The Zeeman interaction

The unpaired electron of a nitroxide interacts with both the static magnetic field and the

nuclear magnetic moment of the nitrogen in a direction-dependent (anisotropic) matter, modifying Eq. 4 such that the electronic g-factor is now a vector rather than a scalar. The magnetic dipoles of the electron and nitrogen nucleus are, respectively:

$$\boldsymbol{\mu}_e = \beta_e \boldsymbol{S}_e \cdot \boldsymbol{g}_e \tag{42a}$$

$$\boldsymbol{\mu}_N = \beta_N \boldsymbol{I}_N \cdot \boldsymbol{g}_N \tag{42b}$$

The complete Zeeman spin Hamiltonian is

$$\widehat{H}_{Zeeman} = \boldsymbol{\mu}_e \cdot H_0 - \boldsymbol{\mu}_N \cdot H_0 \tag{43}$$

Because the nuclear magneton (β_N) is three orders of magnitude smaller than the electronic magneton, Eq. 43 can be approximated as

$$\widehat{H}_{Zeeman} = \beta_e \boldsymbol{S}_e \cdot \boldsymbol{g}_e \cdot H_0 \tag{44}$$

The g-tensor is further simplified by re-expression with respect to the molecular axis system, including two spatial angles, θ and ϕ, as

$$\boldsymbol{g}_e = \sqrt{g_{xx}^2 \sin^2\theta \cos^2\phi + g_{yy}^2 \sin^2\theta \sin^2\phi + g_{zz}^2 \cos^2\theta} \tag{45}$$

thereby reducing it to a simplified diagonal tensor containing three principal values g_{xx}, g_{yy}, and g_{zz}. These diagonal elements can be substituted into Eq. 44

$$\widehat{H}_{Zeeman} = \beta_e \left(g_{xx} H_{0,x} S_{e,x} + g_{yy} H_{0,y} S_{e,y} + g_{zz} H_{0,z} S_{e,z} \right) \tag{46}$$

Although g-anisotropy predominates in inorganic radicals and transition metal complexes, it plays a small role for nitroxides, as well. Because the g-tensor determines the frequency of radiation that satisfies the resonance condition in a static magnetic field, small shifts in principal values of the g-tensor (g_{xx}, g_{yy}, g_{zz}) are indicative of anisotropy in the local field, which heightens the sensitivity of EPR spectra to local structure and environment.

2.2.7 The hyperfine interactions

An additional consequence of the interaction between the unpaired electron with the nuclear magnetic moment of the nitrogen atom is a multiplet splitting of the absorbance line, the so-called hyperfine structure. The nitroxide unpaired electron can interact with all nuclei of nonzero spin in addition to ^{14}N, namely the ^{1}H, ^{2}H, ^{13}C, ^{15}N, and ^{17}O atoms (which have natural abundances of 99.98%, 0.02%, 1.1%, 0.4%, and 0.04%, and spin states I = 1/2, 1, 1/2, 1/2, and 5/2, respectively). However, the strength of the interaction is dependent on the through-bond proximity of the electron to each atom and, as such, the interaction with the ^{14}N nucleus is dominant. A nucleus with a nonzero nuclear spin in a homogenous magnetic field has $2I + 1$ spin states; ^{14}N has 3 such states with spin quantum numbers m_I = -1, 0 and 1, each with a different interaction strength in an applied magnetic field. An electron can be found on a nitrogen nucleus in any of these three spin states, and so electron spin transitions can be induced and observed at three different fields. The low field line corresponds to the m_I = 1 state, the center line to the m_I = 0 state, and the high field line to the m_I = -1 nuclear spin state.

The hyperfine interaction between the electron and the ^{14}N nucleus is made up of two components: the Fermi contact and dipolar interactions

$$\hat{H}_{hyperfine} = \hat{H}_{Fermi} + \hat{H}_{dipolar} \tag{47}$$

For the Fermi contact interaction, the occupancy of the unpaired electron in a spherically symmetric state is quantified and describes the isotropic component of the hyperfine interaction

$$\hat{H}_{Fermi} = a\boldsymbol{I}_N \cdot \boldsymbol{S}_e \tag{48}$$

where a is the isotropic hyperfine coupling constant that depends on the probability density of the electron at the nuclear origin ($|\psi|^2$)

$$a = \frac{8\pi}{3} g_e \beta_e g_N \beta_N |\psi(\boldsymbol{r} = 0)|^2 \tag{49}$$

On the other hand, he dipolar interaction accounts for anisotropy in the interaction between the electronic and nuclear magnetic moments which arise, for example, from electrons in p, d, or f orbitals where $\boldsymbol{r} = 0$. In classical electrodynamics, the energy of the interaction is

$$E_{dipolar} = \frac{\mu_e \cdot \mu_N}{\boldsymbol{r}^3} - \frac{3(\mu_e \cdot \boldsymbol{r})(\mu_N \cdot \boldsymbol{r})}{\boldsymbol{r}^5} \tag{50}$$

Substituting Eqs. 41a and 41b (using scalar quantities of g_e and g_N as an approximation) into Eq. 50 gives the dipolar Hamiltonian

$$\widehat{H}_{dipolar} = -g_e \beta_e g_N \beta_N \left[\frac{\boldsymbol{S}_e \cdot \boldsymbol{I}_N}{r^3} - \frac{3(\boldsymbol{S}_e \cdot \boldsymbol{r})(\boldsymbol{I}_N \cdot \boldsymbol{r})}{r^5}\right] \tag{51}$$

where the three dot products are, of course, orientation-dependent. Just as the g-factor was expanded into a g-tensor, a second-rank tensor $\boldsymbol{A}'$ is constructed to describe the dipolar interaction

$$\boldsymbol{A}_{dipolar} = \boldsymbol{I}_N \cdot \boldsymbol{A}' \cdot \boldsymbol{S}_e \tag{52}$$

with elements that depend on the electron-nucleus distance vector $\boldsymbol{r}$

$$A'_{ij} = -g_e \beta_e g_N \beta_N < (r^2 \delta_{ij} - 3x_i x_j) r^5 > \tag{53}$$

where the Dirac delta function $\delta_{ij} = 1$ only for diagonal elements $i = j$ (A_{xx}', A_{yy}', and A_{zz}'); otherwise, for the off-diagonal elements, $\delta_{ij} = 0$.

As described, the dipolar interaction is unlike the Fermi contact interaction in that it is orientation-dependent, and the diagonal elements of $\boldsymbol{A}'$ can be resolved as peak splittings in the derivative EPR line shape. The magnitude of these splittings was investigated by Griffith *et al.*[8] by orienting crystals of nitroxides along different axes of the laboratory frame in a magnetic field. A maximum splitting of 32 G was observed for crystals oriented parallel to the z-axis ($A_{zz}' = A_{\parallel}$) and a minimum splitting of ~6 G for perpendicularly-oriented crystals in the xy-plane ($A_{xx}' = A_{yy}' = A_{\perp}$). All other non-orthogonal orientations have EPR spectral splittings

that are easily calculable from spatial projections onto the z-axis and into the xy-plane by the angle θ between H_0 and the principal axis of the nitroxide molecular frame by

$$A_\theta = \sqrt{A_\parallel \cos^2\theta + A_\perp \sin^2\theta} \tag{54}$$

Fig. 2.5 demonstrates the orientation-dependence on the line shape for rigid crystalline nitroxides, however, molecular motion also influences the observed anisotropy of the EPR spectrum. Spectral characteristics for rapidly tumbling nitroxides are typically motionally averaged and thus only exhibit an isotropic splitting constant

$$A_{isotropic} = \frac{1}{3}(A_{xx}' + A_{yy}' + A_{zz}') \tag{55}$$

This is typically observed for nitroxides with correlation times shorter than the maximum difference between diagonal elements of the A-tensor ($\Delta A = A_{zz}' - (A_{xx}' + A_{yy}')/2$)

$$\tau_R < \frac{h}{2\pi g_e \beta_e \Delta A} \tag{56}$$

on the order of short nanoseconds and which defines the boundary of the *very fast motional regime*. For a typical nitroxide spin label attached to a surface helix site on a protein, the typical A_{zz}' splitting is 31 G, which corresponds to a correlation time of 1.8 ns. The time window of the *very fast regime* is therefore between 0.1 – 1.8 ns. The *fast* to *slow* tumbling limits for ordered

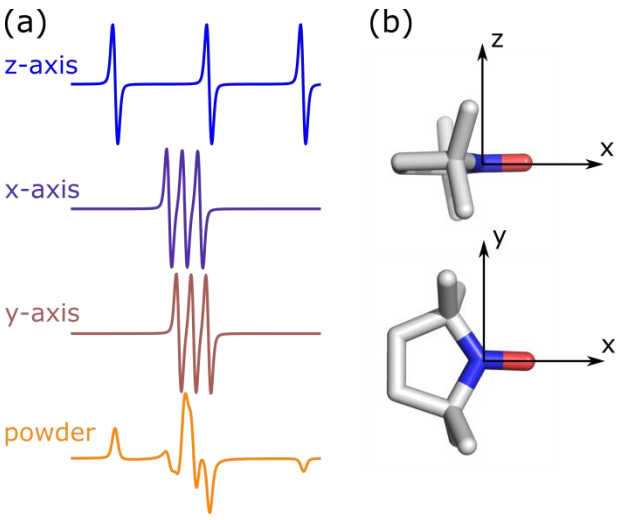

Figure 2.5 (a) Spectral anisotropy of the EPR line shape for a nitroxide with molecular axes oriented along the axes in the laboratory frame, where the direction of $\boldsymbol{H_0}$ defines the z-axis. (b) In the molecular frame, the nitrogen $\boldsymbol{2p_z}$ corresponds to the z-axis.

proteins resolve A- and g-anisotropy, which factor into the total spin Hamiltonian from Eq. 47 such that the resonance condition from Eq. 7 can now be re-expressed as

$$\Delta E = h\nu = \boldsymbol{g}\beta_e H_0 + \boldsymbol{A}' m_I \tag{57}$$

As described above, the Zeeman and hyperfine interactions both have the potential to be strongly anisotropic and so produce spectral anisotropy that is highly sensitive to local fields. Thus, the origin of EPR spectral sensitivity is a direct measure of the motion and local electronic environment of the nitroxide.

2.3 Nitroxide motion

Spectral line shapes obtained for proteins labeled with R1 are heavily dependent on the structure local to and internal motion of the side chain. These label-specific motions can be identified and isolated from those attributable to the protein in order to quantify structure and dynamics of proteins. At X-band, a CW EPR spectrum is able to distinguish nitroxide motions ranging from long-picoseconds to nanoseconds (0.1 – 100 ns).[9] The observed correlation time (τ_R) can be attributed to the combination of three different isotropic motional contributions of the nitroxide: internal or rotameric side chain motions (with correlation time τ_{int}), fast backbone fluctuations (τ_{bb}), and overall protein rotational diffusion (τ_r). Since these motions are expected to be independent, they are additive in inverse

$$\frac{1}{\tau_R} = \frac{1}{\tau_r} + \frac{1}{\tau_{bb}} + \frac{1}{\tau_{int}} \tag{58}$$

and the fastest process dominates. If the protein diffusion lifetime is longer than the EPR timescale, τ_r has a negligible effect on the observed correlation time. Line broadening typically associated with τ_r is minimized for values above 20 ns, which corresponds to proteins ~65 kDa or larger according to the Stokes-Einstein equation for spherical globular proteins.[10] For smaller

proteins, molecular tumbling is so fast that the EPR spectrum is motionally narrowed and no dynamic or structural information can be extracted. Instead, τ_r is increased by the presence of viscogens (typically either 25% w/w Ficoll-70[11] or 30% w/w sucrose[12]) or when attached to solid support beads[13]. When a spin labeled protein tumbles slowly or is completely immobilized, τ_{int} and τ_{bb} exclusively influence the observed nitroxide correlation time. Understanding the physical origins of and separating the effects of these two correlation times has been a major experimental effort in the Hubbell Lab over the last 20 years in order to identify and measure protein-specific motions.

2.3.1 Site-directed spin labeling

Application of EPR to proteins first relied on intrinsic EPR signals. Metalloproteins are intrinsically "EPR-active" by virtue of paramagnetic metal centers, particularly for the low-spin electrons of Fe(III), and these systems were the focus of early EPR studies.[14] Application of the EPR method to diamagnetic biomolecules requires chemical modification to introduce a paramagnetic center. So-called "spin labeling" was developed by McConnell and others in the mid-1960s[15,16] and opened the field of EPR to a large variety of biological systems. Further developments in genetic modification and recombinant protein expression added site-specificity to the technique such that spin labels could be controllably inserted into a biomolecule. This technique, pioneered by Hubbell[17] and known as site-directed spin labeling (SDSL), is a robust method by which to investigate complex biological systems and to characterize the structure and dynamics of proteins.

A common method in modern SDSL employs a cysteine residue engineered at a particular protein site of interest (often after removing all other reactive cysteines in the native

protein) using site-directed mutagenesis. This residue is modified with a sulfhydryl-reactive reagent to yield a new covalently-attached stable nitroxide free radical side chain. The most common spin label is 1-oxy-2,2,5,5,-tetramethylpyrroline-3-methylmethanethiosulfonate[18], which reacts to form a disulfide bond with the cysteine sulfhydryl moiety. The reaction is very specific, rapid, and stoichiometric[12], and the resulting spin label side chain is designated "R1," shown in Fig. 2.6. It has been demonstrated that modification of a protein by SDSL at solvent-exposed sites only minimally perturbs the protein structure and any destabilizing effect from SDSL is generally due to cysteine mutagenesis rather than incorporation of R.[12]

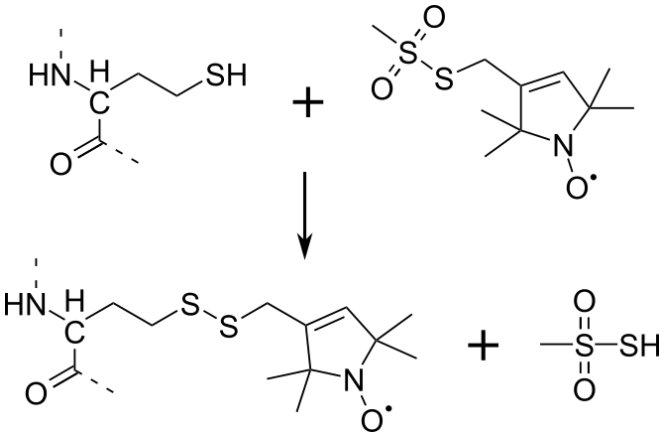

Figure 2.6 Site-directed spin labeling is the reaction of a protein cysteine sulfhydryl with the spin label reagent MTSL. A cysteine is engineered into the sequence of a protein. Reaction with MTSL produces a new side chain R1, which introduces a paramagnetic sensor compatible with EPR studies.

EPR spectroscopy of spin labeled proteins prepared in this way produces nuanced EPR spectra with characteristic line shapes, sensitive to local protein structure. Because the EPR line shape depends on the orientation and motion of a nitroxide, information about the chemistry of the nitroxide, the local protein geometry, and the rotational tumbling of the protein can be determined from the EPR spectrum.

2.3.2 Line shape analysis

The structure and dynamics of the R1 side chain on solvent-facing surfaces of α-helices and ordered loops have been extensively characterized by X-ray crystallography[19–23],

mutagenesis of local side chains[12], and quantum mechanical calculations[24] to develop a model for R1 motion. From these efforts, the "X_4/X_5 model[25]" was devised and shown to robustly predict the internal motion of R1 tethered to solvent-facing helices (Fig. 2.7). The model defines the rotameric motion of the nitroxide side chain as the independent rotational isomerization of each dihedral ("X") angle in the R1 tether, and describes how they vary by the atoms involved and the types of internal interactions that exist. X-ray structures indicate that a hydrogen bond between $S_\delta \cdots HC_\alpha$ restricts the X_1 and X_2 angles each to three preferred rotamers (-60°, +60°, and 180° orientations). The disulfide dihedral X_3 isomerizes very slowly on the EPR timescale and is limited to two rotamers (-90° and +90° orientations)[26]. Therefore, the observed R1 motion is primarily a result of X_4 and X_5 rotameric exchange and, as such, the motion is anisotropic due to these motional restrictions. Using the X_4/X_5 model to account for internal R1 motion, one can usually separate the effects of fast protein backbone motion to investigate site-specific protein structure and dynamics. However, Chapter 4 provides commentary on cases where this model does not apply well, such as for rotamers of R1 in β-sheets.

A- and *g*-tensor anisotropy reveals polarity and electronic structure information. Resolving *A*- and *g*-tensor anisotropy requires motion be slow enough to avoid motional averaging but fast enough such that nitroxide motion is actually sampled by the EPR technique. These two temporal boundaries separate three motional regimes: fast, slow, and very slow or rigid limit (100 ps – 2 ns, 2 ns – 100 ns, and >100 ns, respectively). Motions on the timescale of

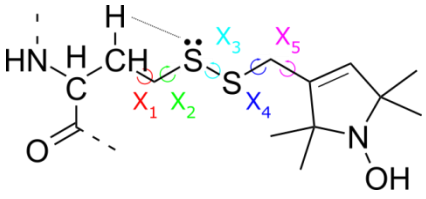

Figure 2.7 Dihedral angles ($\mathbf{X_1 - X_5}$) are labeled to discuss segmental motion of the R1 side chain. The "$\mathbf{X_4/X_5}$" model predicts that fast isomerization of the final two dihedral angles contribute to nanosecond motion of the nitroxide in the context of a solvent-facing helix or loop.

T_2 relaxation are resolved in EPR spectra in the slow time regime, producing line shapes that are particularly informative: the spectral diversity between 2 – 100 ns is of particular interest because small differences in motion are easily resolved. In contrast, "fast" motions are highly averaged such that the line shape invariably resembles three sharp lines; all motions slower than ~100 ns appear immobile, yielding a so-called "rigid limit spectrum." Fig. 2.8 demonstrates the variation in EPR line shape for a nitroxide with motions in the fast, slow, and rigid limit motional regimes.

Relative mobility can be assessed semiquantitatively in the fast and very slow motional regimes. Line shape analysis of the peak-to-peak distance of the first derivative line shape in

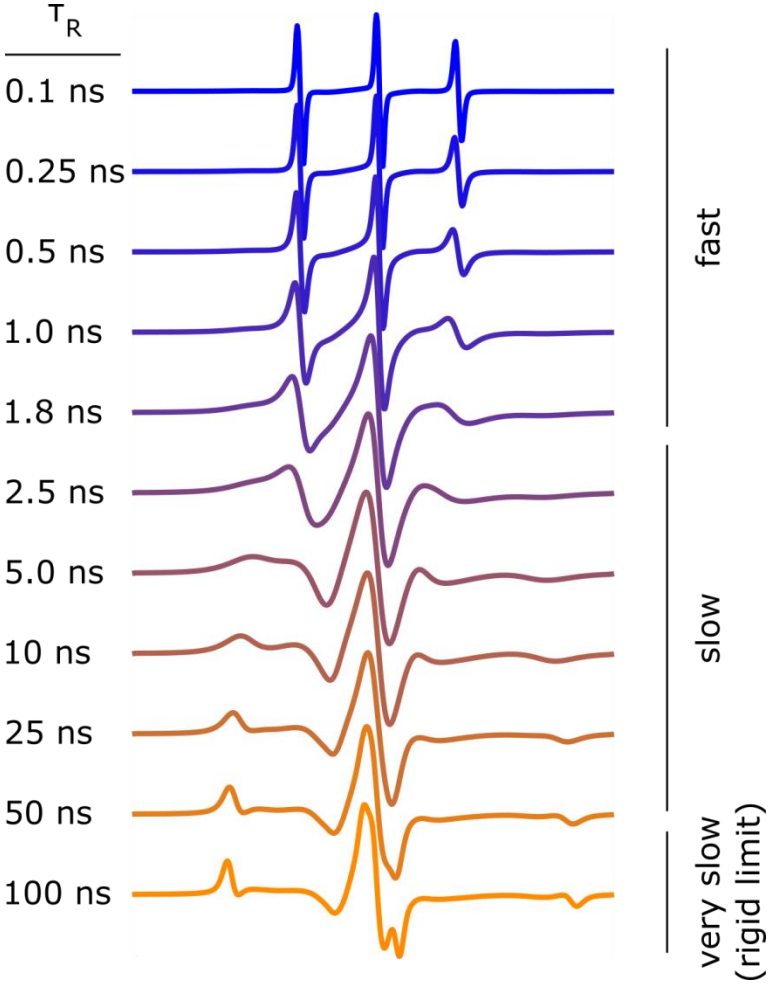

Figure 2.8 The nitroxide correlation time has a significant influence on the EPR line shape, here shown for nitroxides which have isotropic motion (the value of $\boldsymbol{\tau_R}$ is noted on the left). Nitroxides tumbling in the fast motional regime show three sharp lines; as the correlation time slows, the high-field line ($\boldsymbol{m_I} = -1$) broadens compared to the center- and high-field lines. In the slow motional regime, these lines are broadened by various mechanisms to produce a large variety of line shapes. Rigid limit motion produces line shapes with large splittings between the high and low field lines ($\boldsymbol{m_I} = -1$ and 1). The effect of $\boldsymbol{A}$-tensor averaging is visible by comparing the peak-to-peak splitting of the most strongly immobilized spectrum ($\boldsymbol{\tau_R} = 100$ ns) to those with faster motion, which decreases the observed splitting width.

highly mobile spectra (fast regime) provides an inverse central resonance linewidth (ΔH_{pp}^{-1}) that is proportional to the overall τ_R^{-1} for simple anisotropic motion of R1. Often the R1 side chain makes local contacts that strongly dampen its motion. The value of ΔH_{pp}^{-1} in these cases is no longer related to τ_R only, but is still a useful measure of relative mobility. Additionally, in the very slow motional regime, the relative mobility can be described by the peak-to-peak distance of the high and low field lines ($2A_{zz}'$).

Quantitative line shape analysis of spectra in the slow motional regime by spectral simulation has been particularly informative, used for revealing site-specific differences arising from protein backbone motion (τ_{bb}).[9,27] The method and parameters for spectral simulation are discussed in the following section.

2.3.3 Spectral simulation

Simulations of spectra in the slow motional regime are more complicated than those for fast motion and rigid limit lineshapes, where the magnetic tensors and rotational dynamics are more easily resolved. Parameterization of these spectra and developing a 'simplest-interpretable' model to describe the lineshape sufficiently is the goal of spectral simulation and fitting. A useful quantitative method of simulating R1 motion employs the theoretical model of Freed, wherein the entire spectral line shape can be simulated using a large number of physically relevant parameters[28]. The Microscopic Order Macroscopic Disorder model (MOMD) uses the stochastic Liouville equation to semiclassically model motion of the nitroxide when tethered to a macromolecule[29]. The model is semiclassical in that a set of partial differential equations used to simulate experimental spectra treats electron spins quantum mechanically and atomic/molecular motions classically. The quantum mechanical object, the unpaired electron, is found on a spin

label tethered to a protein experiencing restricted classical motion (microscopic order), while the protein classically and isotropically tumbles, as long as it is globular and spherical (macroscopic disorder). The motion of the electron, which primarily populates the nitrogen $2p_z$ orbital of R1, is restricted to a small cone projecting above and below the plane of the nitroxide ring, coupled to the motion afforded the side chain by the physical constraints described in the X_4/X_5 model. The internal correlation time τ_{int} and the order parameter S ($0 \leq S \leq 1$, where $S = 0$ is isotropic and $S = 1$ is perfect order) are highly discriminating spectral parameters with which to describe EPR line shapes.

Below is a summary of important parameters that factor into the basis of MOMD; a more thorough explanation of all fitting parameters is given elsewhere.[28] Optimization of the following fit parameters is the primary focus of spectral simulation:

(1) The R diffusion tensor represents motion of the nitroxide by a Brownian diffusion model. Its components (R_x, R_y, R_z) are the rate of rotation about each of the diffusion frame axes (x_R, y_R, z_R). Nitroxide motion is generally at least axially symmetric, that is, $R_x = R_y$, while isotropic motion is the condition in which $R_x = R_y = R_z$.

(2) Small-amplitude motion of the internal nitroxide side chain is defined by an ordering potential, which is limited by the torsions allowed in the R1 side chain and is anisotropic. The potential function is given by

$$U(\Omega) = -\mathrm{k_B}T \sum_{L,K} c_L^K D_{0K}^L(\Omega) \qquad (59)$$

where the inner terms c_L^K and $D_{0K}^L(\Omega)$ are weighting coefficients and spherical harmonics functions, respectively, and $\Omega = (\alpha, \beta, \gamma)$ are the angles that relate the diffusion tensor R axes to a director frame (ranging from 0 - 90°).

(3) A-tensor values determined for R1 by multifrequency EPR are fixed to $A_{xx} = 5.7$, $A_{yy} =$

6.0, and $A_{zz} = 37.5$ G.[30]

(4) g-tensor values, also determined by multifrequency EPR, are fixed to $g_{xx} = 2.00780$, $g_{yy} = 2.00580$, and $g_{zz} = 2.00230$.[30]

(5) The diffusion tilt angles $(\alpha_D, \beta_D, \gamma_D)$ relate a new rotational frame (x_R, y_R, z_R) to the principal axis system of A- and g-tensors. The angles are obtained by first, counterclockwise rotation about z_R through the angle γ_D, second, counterclockwise rotation about y_R through the angle β_D, and third, counterclockwise rotation about z_R through the angle α_D. Spectra are generally invariant over the angle α_D (in the case of axial symmetry); thus, β_D and γ_D are varied in spectral simulations.

The director frame is a vector z_D, related to $\boldsymbol{H}_0$ by the angle ψ, and is the axis of symmetry. In mobile, disordered samples such as proteins in solution, z_D about $\boldsymbol{H}_0$ is isotropically distributed such that ψ is averaged over all orientations. In this situation, a simplified expression for $U(\Omega)$ is

$$U(\theta) = -k_B T \left[\frac{1}{2} c_0^2 (3\cos^2\theta - 1)\right] \tag{60}$$

where c_0^2 is the first weighting coefficient and θ is the instantaneous angle between z_R and z_D, e.g. for one protein orientation at the time of measurement. Higher order harmonic terms can be incorporated, but are generally not necessary in spectral fitting. Using the first weighting coefficient c_0^2 and the first term $D_{0,2}^0$ from Eq. 59, the amplitude of motion is

$$S_0^2 = \frac{1}{2}[\![3\cos^2\theta - 1]\!] \tag{61}$$

For each simulated spectrum, a set of (A_{xx}, A_{yy}, A_{zz}), (g_{xx}, g_{yy}, g_{zz}), $(\alpha_D, \beta_D, \gamma_D)$, (R_x, R_y, R_z), and S_0^2 parameters are necessary to describe most solvent-facing nitroxide line shapes. The diffusion tensor components can be reexpressed in a modified spherical notation

$$\bar{R} = \frac{1}{3}\left(R_{xx} + R_{yy} + R_{zz}\right) \tag{62a}$$

$$N = R_{zz} - \frac{R_{xx}+R_{yy}}{2} \quad (62b)$$

$$N_{xy} = R_{xx} - R_{yy} \quad (62c)$$

where isotropic spectra are characterized by $N = 0$ and $N_{xy} = 0$. The isotropic correlation time for a spherical particle is given by

$$\tau_R = \frac{1}{6\times10^{\bar{R}}} \quad (63)$$

in units of seconds.

A LabVIEW program by Christian Altenbach (available for download at https://sites.google.com/site/altenbach/) using EPR spectral simulation code written in Fortran by Budil *et al*.[28] provides a graphical interface by which these calculations are performed. However, because the fitting algorithm cannot identify which parameter values are physically reasonable (e.g., $A_{zz} = 100$ G would be an impossible result), it is up to the user to determine whether the simulation makes sense. The LabVIEW program provides the user with a table of correlation values for the various fitting parameters used to aid the user in identifying which parameters should not be varied simultaneously in a given fit. In this dissertation, fitting parameters reported are those that simultaneously provide adequate (albeit rarely perfect) fits to data obtained, and reasonably fit the simplest structural, motional, or thermodynamic model presented.

In an ideal case, SDSL of a solvent-facing, noninteracting site in a rigid helix (τ_{bb} = 0) is described well by a simulated spectrum with $S = 0$ and $\tau_{int} = 2$ ns. Deviations from these values are due to unique local protein structure. More mobile spectra have smaller values of S and/or τ that reflect backbone fluctuations on the nanosecond timescale ($S < 0.5$ and/or $\tau_{bb} < \tau_{int}$).[21] Less mobile spectra have larger values of S and/or τ_R due to interactions of the side chain with local protein structure in addition to the interactions predicted by the X_4/X_5 model. By separating order and correlation time parameters, the motional characteristics of a nitroxide can

be described in great detail. For example, highly ordered yet mobile spectra can be identified despite appearing as simply 'immobile' at a first glance. The high sensitivity of EPR spectra to variations in S for a fixed value of τ_R is demonstrated in Fig. 2.9.

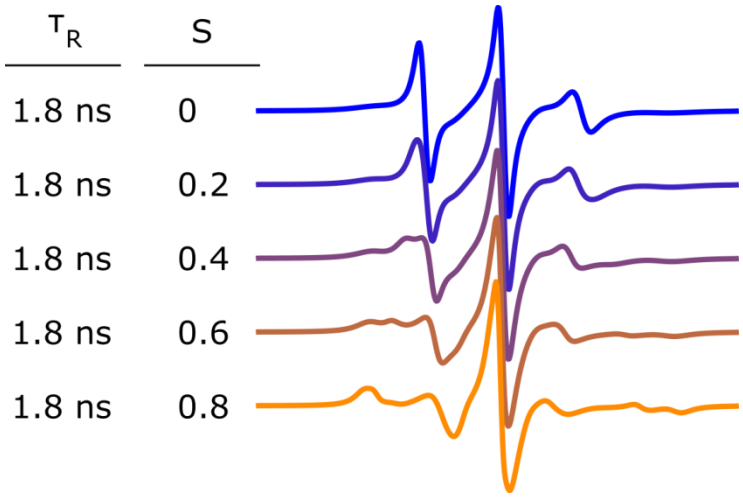

Figure 2.9 The X-band EPR spectrum depends on the correlation time (here, all spectra are simulated with $\boldsymbol{\tau_R} =$ 1.8 ns) and order, represented by the order parameter $\boldsymbol{S}$.

2.3.4 Multicomponent spectra

In many cases, the EPR spectrum for a spin labeled protein is "multicomponent," where the observed spectrum is the linear combination of two or more spectra; these complex spectra result from nitroxide exchange between multiple distinct states. If the exchange lifetimes connecting these states are longer than the EPR timescale at X-band (> 100 ns), a perfectly separable multicomponent spectrum will be observed. A multicomponent spectrum is the population-weighted sum of its underlying components when they are in slow exchange on the intrinsic EPR timescale. Differences in these spectral components depend on the difference in the local structure and dynamics of the states. Exchange on or faster than the EPR timescale results in averaging into a single spectral component rather than discrete components. Additionally, because nitroxides with motion in the fast and slow limits (τ_R < 1 ns and > 20 ns) have indistinguishable line shapes, two spectral components with similar fast (or slow) motion can be difficult to resolve. Despite this potential difficulty, line shape analysis is a powerful method by

which to obtain structural and dynamic information about the sample being studied.

In general, multicomponent spectra can arise from one of two types of exchange: rotameric or protein conformational exchange. Rotameric exchange of a nitroxide between an interacting and a noninteracting state, for example, can produce a complex lineshape resulting from the linear combination of with mobile and immobile lineshapes. This type of exchange is generally faster than protein conformational exchange, which is typically on the microsecond to millisecond timescale. A multicomponent EPR spectrum contains no information on the exchange rate, other than that it exceeds the intrinsic limit for spectral averaging. Distinguishing rotameric from conformational exchange requires either osmolyte perturbation paired with CW EPR[11] (a stabilizing osmolyte shifts the conformational equilibrium towards the state with lower solvent-exposed surface area; rotameric states have near-equivalent surface areas and their populations are unaffected by the osmolyte), saturation recovery (a pulsed EPR technique sensitive to slower motions on the microsecond timescale), or, as will be shown, high hydrostatic pressure paired with CW EPR.

Quantification of the states resulting in a two-component lineshape is most easily handled by spectral simulation. The individual lineshapes can be co-simulated via the MOMD model, from which the corresponding populations, S, and τ_R parameters can be extracted. The fractional populations of multicomponent spectra are of particular interest for determining equilibrium constants and the underlying relative energies and energy difference for the transition between the two states, which can be surveyed with varying experimental conditions such as pH and hydrostatic pressure.[31]

2.4 EPR instrumentation

The EPR spectrometer is a tool that measures the absorption of microwave electromagnetic radiation by a sample of interest under controlled field conditions This technique involves detecting changes in reflected energy due to resonant absorption, which is highly sensitive to the small absorption signals associated with the relative energies and population differences of Zeeman spin states. There are a number of practical difficulties that make EPR more technically challenging than other forms of spectroscopy (in particular, it's magnetic resonance cousin, NMR), but field-sweeps and modulation at fixed frequency (in contrast to frequency sweeps in a static field, as is done in NMR) has long been possible with microwave technology (only recently have NMR-analogous pulsed methods become accessible to EPR spectroscopists). An essential part of the method involves precise static control of microwave frequency, which must be "tuned" appropriately to detect resonant absorption in a reproducible fashion.

Fig. 2.10 is a schematic picture of a CW EPR spectrometer. The microwave bridge houses a klystron vacuum tube or Gunn diode, known as the "microwave source," which supplies the radiation used in the experiment. Coaxial cables and rectangular waveguides with dimensions that match the microwave wavelength direct the microwaves to the sample chamber. The sample chamber is a resonator cavity capable of internally reflecting the microwave energy is critically coupled so that no signal is reflected back to the detector when tuned. A number of different coupling mechanisms have been devised to change the shape of the resonator slightly in order to match this condition. A variety of resonator designs accommodate different sample sizes, concentrations, and experimental conditions (for example, temperature). The quality factor of a resonator is its radiation "storage efficiency" and reflects its sensitivity, given by

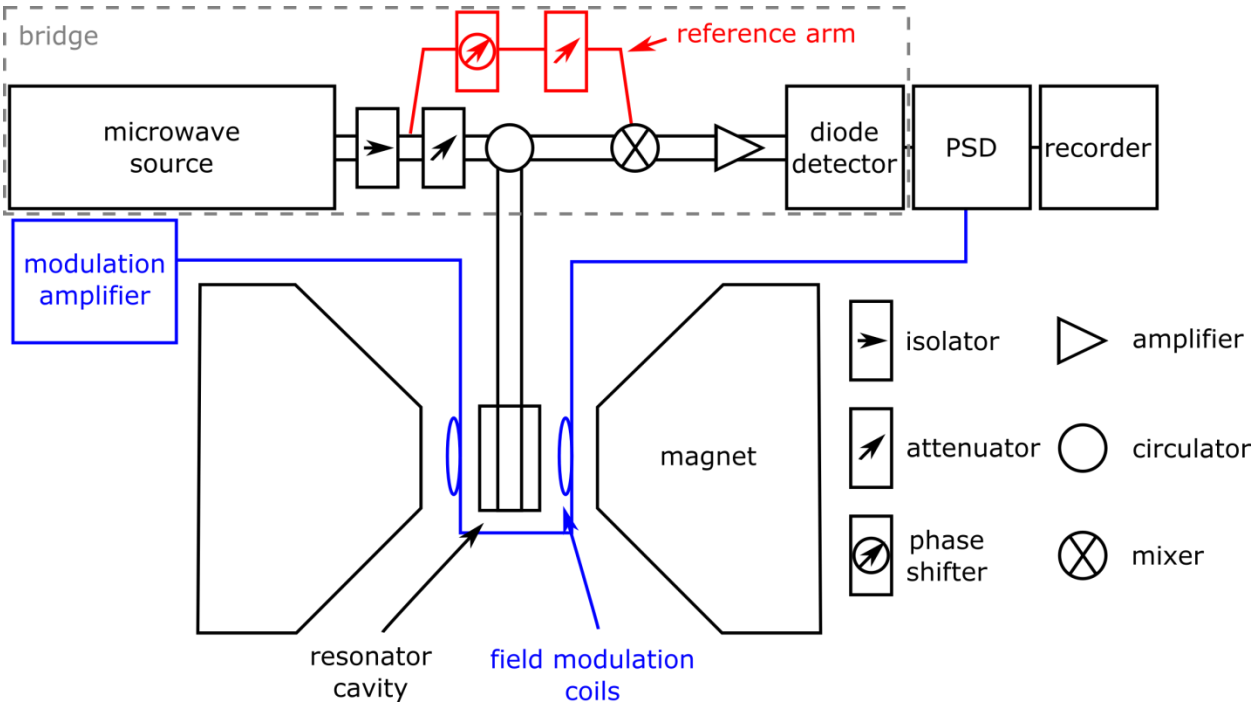

Figure 2.10 Organization of the microwave bridge, magnet, resonator cavity, and recorder of a CW EPR spectrometer. The microwave bridge components are shown schematically in order to demonstrate how the parallel reference arm integrates into the signal channel before reaching the detector.

$$Q = 2\pi \frac{Energy\ stored}{Energy\ dissipated\ into\ the\ resonator\ body} \qquad (64)$$

which is selected for particular EPR experiments.

A closer look at the microwave bridge shows how the reflected signal is obtained (refer to the dashed box in Fig. 2.10). Microwave radiation produced by the source passes through an isolator, which prevents backflow of radiation into the source components. An attenuator follows, which precisely controls the microwave power. The radiation is directed via a circulator towards the resonator cavity, which is tuned such that any reflected radiation is indicative of absorptions due to electronic transitions in the sample. The resonant absorptions manifested as reflected radiation constitutes the EPR signal. Parallel to the circulator path is a reference arm path, which provides a reference signal from the source to the detector directly. The microwaves traveling through this reference arm channel are phase-adjusted to match the signal phase and are

added to the reflected signal radiation via a mixer. This adds a bias to the EPR signal in order to bring the total energy into the linear detection range of the detector. The detector is a Schottky diode, which converts the microwave photons to a voltage that is digitally recorded by a computer or similar acquisition system.

Signal sensitivity is significantly increased by phase-sensitive detection of the field-modulated EPR signal; in fact, without the advent of phase-sensitive detection and field modulation, CW EPR of radicals at micromolar concentration would be impossible. 100 kHz field modulation is applied to the resonator cavity through coils mounted within the magnetic field. As the field strength oscillates sinusoidally, so does the detected signal amplitude at this same frequency. Phase-sensitive detection selects only those 100 kHz signals with matching phase to the reference signal. As long as the modulation amplitude is small compared to the intrinsic EPR linewidth (e.g., 0.5 – 2.0 G for typical nitroxide-labeled proteins), the change in the detected signal is approximately linear. The rate of change of the modulated signal is recorded, which results in a derivative spectrum and so it is typical for CW EPR spectra to be presented in this form, rather than a post-acquisition integration to be performed to yield an absorption curve.

2.5 EPR experiments

In addition to CW EPR, a number of experiments have been designed to expand the experimental range of the EPR spectrometer. This section describes three such EPR experiments.

2.5.1 Dipolar broadening

Dipolar couplings between two nitroxides within 8 to 20 Å of each other (for example, on a doubly-labeled protein) broaden the CW EPR line shape in a distance-dependent manner. The

magnetic dipole of one spin induces a change in the local field near the second spin, which contributes to the hyperfine interaction and changes the line splitting. In the rigid limit where molecular tumbling is slow, the dipolar coupling produces a splitting dependent on the interspin distance r and the orientation of the interspin vector in the magnetic field θ

$$\Delta H_{dd} = \pm\frac{3}{4} g\beta \frac{(3\cos^2\theta - 1)}{r^3} \quad (65)$$

this relationship holds as long as the nitroxides have correlation times that are slower than ~20 nanoseconds.[32]

In practice, the dipolar-broadened line shape is compared to a noninteracting reference spectrum containing the sum of single-nitroxide line shapes, often obtained by limited spin labeling to produce a protein with a statistical distribution of single and doubly-labeled cysteines. Often the spin label is diluted with a diamagnetic analogue (an acetylated nitroxide closely resembles the structure of R1 but has no paramagnetic electron). A 1:3 molar ratio of R1 to acetylated R1 produces a combinatorial mixture of labeled proteins: 9/16 contain doubly-labeled acetyl-R1 (which is EPR silent), 6/16 contain one R1 and one acetyl-R1 label attached to each protein (which provides a spectrum containing the two noninteracting line shapes), and 1/16 contain doubly-labeled R1 (a small fraction yielding a dipolar-broadened EPR spectrum that is subtracted from the mixed-label line shape). The resulting "sum of singles" line shape is then convoluted with a broadening function to reproduce the dipolar-broadened spectrum. This broadening function is a Pake pattern, which results from dipolar-coupled spin pairs that are randomly oriented with respect to H_0. The ΔH_{dd} splitting can be measured between the peaks of a Pake pattern for an interspin vector perpendicular to H_0, e.g. $\theta = 90°$. Determination of r is accomplished by substitution of these values into Eq. 65. A more complete solution to Eq. 65 by integration over all orientations of the interspin vector ($0 < \theta < 180°$) provides an interspin

distance distribution.

2.5.2 Saturation recovery

In contrast contrast to the above CW method, pulse EPR experiments probe additional properties of nitroxides that are not revealed directly in the line shape. Saturation recovery (SR) measures the experimental/effective T_1 ($T_{1,eff}$) relaxation time of a nitroxide, which differs from its intrinsic T_1 due to contributions from additional relaxation mechanisms. These mechanisms are described below but a more rigorous discussion is found elsewhere.[33,34]

A saturating pulse with a fixed power and duration perturbs the Boltzmann distribution of spin populations to a situation of zero net magnetization ($N_\alpha = N_\beta$). Recovery of magnetization (M_z) is achieved through spin-lattice relaxation (with a characteristic T_1 time constant) as an exponential function

$$M_z(t) = M_{z,0} + \left(M_{z,SR} - M_{z,0}\right)e^{-\frac{t}{T_{1,e}}} \tag{67}$$

where $M_{z,0}$ is the equilibrium magnetization before the pulse, $M_{z,SR}$ is the magnetization immediately after the pulse at $t = 0$ (which should equal zero for a completely saturating pulse), and $T_{1,e}$ is the intrinsic spin-lattice relaxation time of the electron belonging to the nitroxide spin label. Spectral diffusion contaminates the measured $T_{1,e}$ via contributions from molecular tumbling, nuclear spin relaxation, "cross relaxation", or fast multi-state exchange[33]; hence the experimentally determined exponential recovery is "effective." In addition, the low-power CW detection method used to obtain SR data contributes to the relaxation time constant (increasing power decreases the effective $T_{1,eff}$), so care must be taken when reporting "intrinsic" relaxation time constants to extrapolate this power-dependent behavior to zero. Heisenberg exchange from collisions with other paramagnetic species (i.e. molecular oxygen) may also accelerate the

observed $T_{1,eff}$ according to the frequency of collisions, which is dependent on the exchange reagent concentration and accessibility of the nitroxide to the exchange reagent. Lastly, an additional paramagnetic species with sufficiently small $T_{1,e}$ (e.g. a metal ion) can provide a distance-dependent shortening of $T_{1,eff}$ through the dipolar effect.[33] For nitroxides attached to a site in a protein undergoing conformational exchange, the two field environments experienced by the electron give rise to a second $T_{1,eff}$ weighted by its fractional population. Each of these exchange processes contribute to the SR signal as a sum of exponentials, which for the mechanisms delineated above amount to a complex, additive form of Eq. 67

$$M_z(t) = A_1 e^{-(2W_e+2W_{ex})t} + A_2 e^{-(2W_e+2W_{ex}+3W_n)t} + A_3 e^{-(2W_e+2W_{ex}+3W_R)t} + A_4 e^{-(2W_e+2W_{ex}+3W_n+3W_R)t} + \cdots \quad (68)$$

where each A is a pre-exponential amplitude and each W is an exchange frequency for one of many various relaxation processes. In the above example equation, $W_e = 2T_{1,e}^{-1}$ is due to intrinsic electron $T_{1,e}$ relaxation; W_{ex} is the Heisenberg exchange frequency; $W_n = 2T_{1,n}^{-1}$ is the nuclear relaxation rate that describes "spectral diffusion" between adjacent nuclear manifolds (i.e. nuclear spin flips); and $W_R = \tau_R^{-1}$ is the spectral diffusion of the nitroxide due to rotational diffusion. These exchange frequencies occur on different timescales ($W_e \approx 0.1 – 1.0$ MHz, $W_n \approx$ 10 MHz, and $W_R \approx$ kHz – MHz). Spectral diffusion processes faster than W_e, for a sufficiently long saturating pulse, play minor roles in SR and so can be completely ignored from Eq. 68, such that higher order terms are omitted because their contributions are small enough to be negligible.

Ignoring the effects of spectral diffusion by performing long pulse experiments, an experimental SR curve according to Percival and Hyde[35] is proportional to

$$S(t) \propto M_{y,SR} e^{-\frac{t}{T_{2,e}}} + \gamma_e H_1 T_{2,e} \left[\frac{M_{z,SR}-M_{z,\infty}}{1+\gamma_e^2 H_1^2 T_{1,e} T_{2,e}}\right] e^{\left[-\frac{t}{T_{1e}}-\gamma_e^2 H_1^2 T_{2,e} t\right]} +$$

$$\frac{M_{z,\infty}\gamma_e H_1 T_{2,e}}{1+\gamma_e^2 H_1^2 T_{1,e} T_{2,e}} - M_{z,SR}\gamma_e H_1 T_{2,e} e^{-\frac{t}{T_{2,e}}} \quad (69)$$

where $M_{y,SR}$ and $M_{z,SR}$ are the magnetization vectors along the y- and z-axes immediately following the pump pulse, $M_{z,\infty}$ is the equilibrium magnetization along in the z-axis, γ_e is the gyromagnetic ratio of an electron, and H_1 is the observe field strength. This expression can be decomposed term-by-term. The first quantity, a free induction decay (FID), can be removed by phase cycling data acquisition in the y-direction. The second term makes up the quantity that contains the $-t/T_{1,e}$ exponential factor essential to the SR experiment. Also within the exponent is a term that describes the effect of high observe power on exhibited T_1: high values of H_1 result in a smaller measured effective $T_{1,eff}$; if detection were possible at zero observe power, the measured T_1 would represent the intrinsic quantity. The third term is a time-independent baseline value indicating the final intensity of the SR curve after complete recovery following the pulse ($t \approx \infty$). The fourth term vanishes for complete saturation ($M_{z,SR} = 0$).

The pulse sequence is straightforward. A high power pulse at a fixed frequency typically saturates the center line ($m_I = 0$ manifold) of spins for a specified pulse length. A defense pulse overlaps the pump pulse to protect the bridge detector. Recovery of the EPR signal at this frequency is monitored by CW with low observe power; that is, application of an observe field H_1 in the minimally disrupting low power case. Correction for artifacts due to heat generated by the saturating pulse and instrumental or electronic noise is made by subtracting an off-resonance SR trace obtained during the experiment. Optimization of the observe power requires that it is sufficiently low to avoid distortion of the SR signal but high enough to allow sufficient signal to noise; experimentally it is typically ~100 μW.

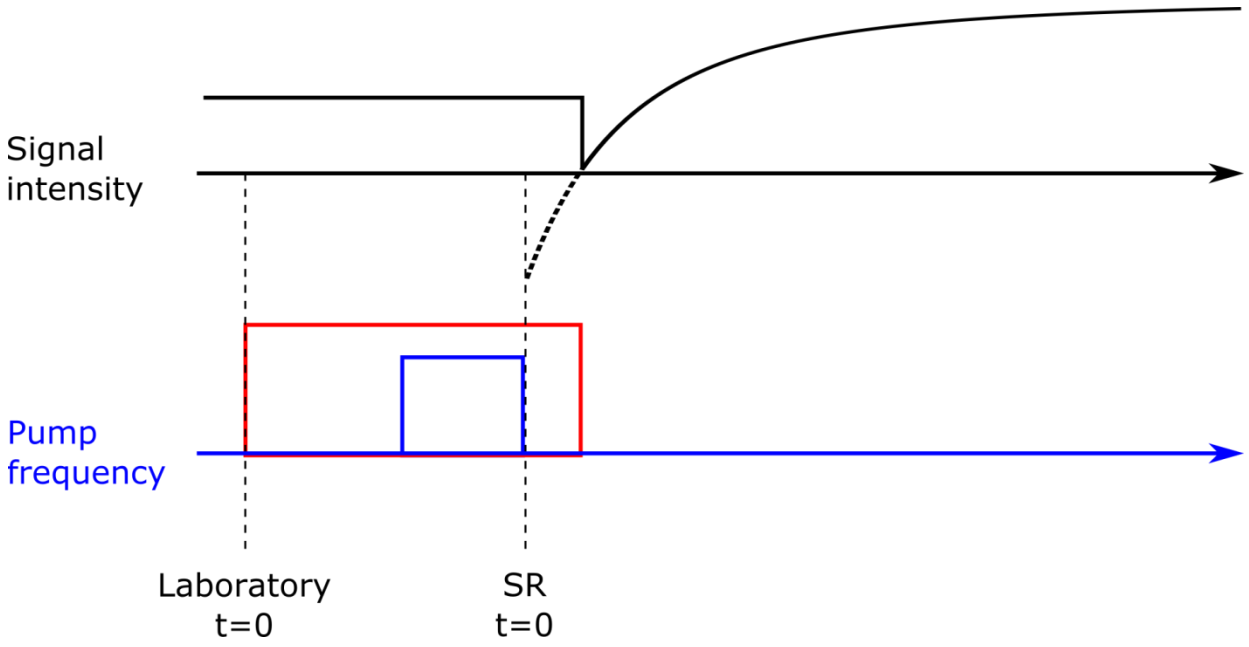

Figure 2.11 Schematic of the SR pulse sequence and the signal amplitude measured as a function of time. The recorded signal amplitude (black) contains the defense pulse (red) and the SR signal from which the exponential is extrapolated to the experimental t=0 at the end of the saturation pulse (blue).

2.5.3 Double electron-electron resonance

While dipolar-broadened CW spectra resolve distances between nitroxide spin labels in the 8 – 20 Å range, pulsed EPR is capable of determining distances between two nitroxides up to 80 Å. This experiment is called double electron-electron resonance (DEER); the 4-pulse DEER technique is discussed herein.[36]

DEER EPR monitors the dipolar coupling between spins in the time domain, which begins with an electron spin-echo (ESE) pulse sequence to manipulate the direction of spin magnetization, shown in Fig. 2.1. A 90° pulse tips the spin magnetization, which was aligned mostly with the z-axis as described in Section 2.2.3, into the xy-plane (Fig. 2.1a), whereafter the

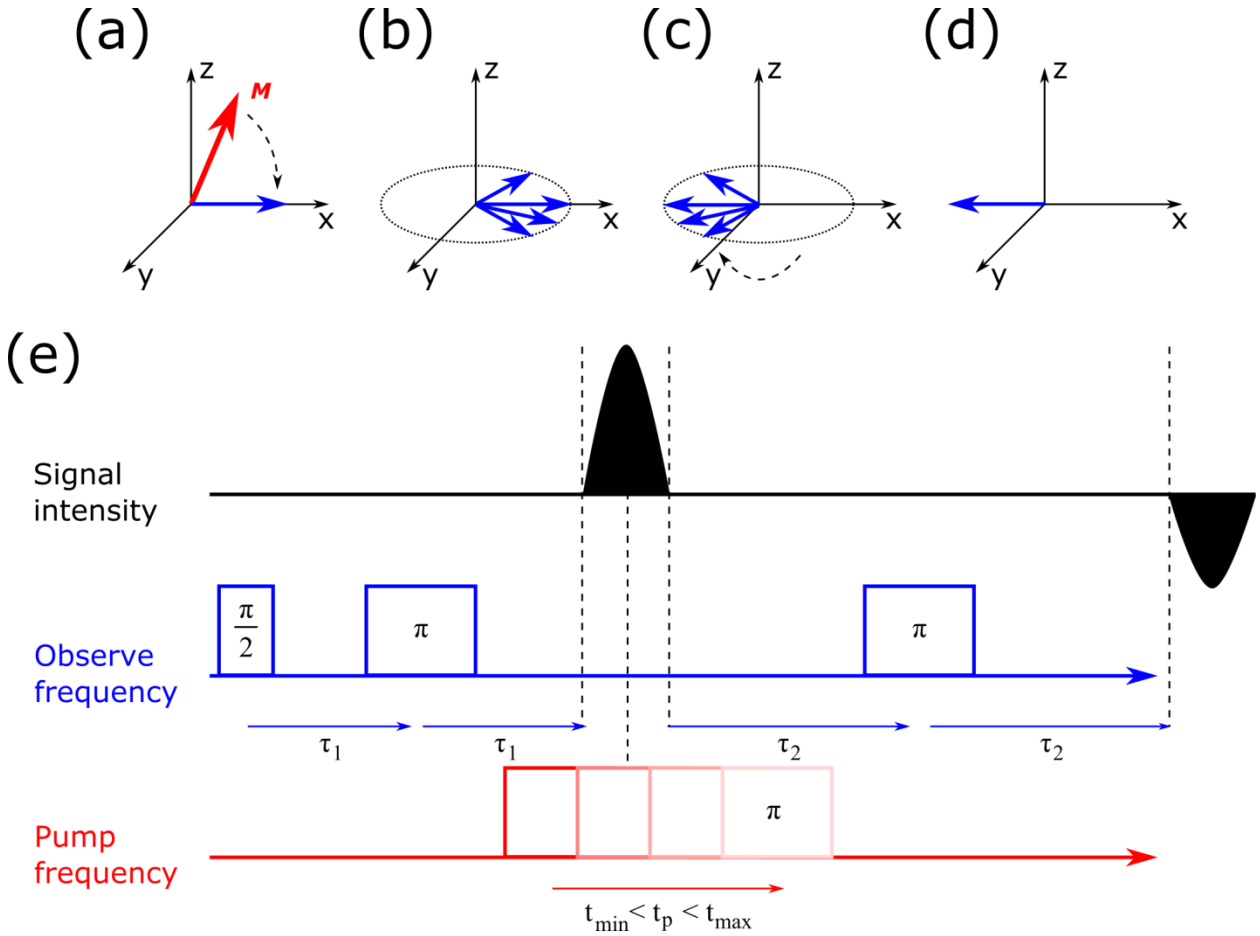

Figure 2.12 In (a-d), the spin echo experiment is shown graphically. Spin magnetization is tipped into the xy-plane in (a) by a 90° pulse, which dephase after a time "t" in (b). A 180° pulse is applied, which mirror the spins across the y-axis, allowing them to rephase after an equal time "t". The echo in (d) is then monitored at time "2t" after the start of the experiment. In (e), the four-pulse DEER sequence is shown, which induces transitions in two populations of spins at two frequencies ("observe" and "pump").

spins dephase due to local field inhomogeneities (Fig. 2.1b). After a specified time interval τ_1, a second 180° pulse mirrors the spin directions (in Fig. 2.1c, this 180° inversion is across the y-axis). The spins come back into phase during a second τ_1 interval. The echo monitors the recovery of spin magnetization in the –x axis (Fig. 2.1d).[37]

Four-pulse DEER builds on the ESE pulse sequence by introducing a pulse at a second frequency, called the "pump" frequency (ω_2), in contrast to the 90° and 180° spin-echo pulses at

the "observe" frequency (ω_1). A 180° pulse at ω_2 at a specified time t_p induces a change in the local field of spins at the "observe" frequency (ω_1) according to the strength of dipolar interaction. The spins at ω_1 experience a phase gain related to the dipolar coupling between the nitroxide electrons $\phi = \omega_{dd} t_p$, where ω_{dd} is the dipolar coupling frequency and t_p is the time position of the "pump" pulse, which varies from some t_{min} to t_{max} relative to the primary ESE. A final 180° pulse at ω_1 at a time interval $\tau_2 > t_p$ after the primary ESE inverts the spin precession direction again, which produces a secondary ESE at a time $2\tau_2$. The DEER signal comprises secondary ESE, which is modulated by the dipolar coupling probed by the time position of the "pump" pulse.

Data fitting through a Pake transform[36] or via a library of DEER dipolar evolution functions (DEFs) determines the encoded distance distributions. Because biomolecules are not static structures and the distances derived from DEER data are statistical distributions, DEER gives insight into the structure and dynamics of proteins on long distance scales. Fig. 2 demonstrates the relationship between short and long oscillations and corresponding distance distributions; strong dipolar couplings produce DEER DEFs that are sharp (Fig. 2.13a, blue) are closely spaced (Fig. 2.13b, blue) compared to broader DEFs (Fig. 2b, magenta) which are due to nitroxides spaced further apart (Fig. 2.13b, magenta). Determination of complex distance distributions is a particular strength of DEER EPR, in which the distribution of conformational arrangements can be directly measured (Fig 2.13a and Fig. 2.13b, cyan). Software (LongDistances, written by Christian Altenbach, or DeerAnalaysis, written by Gunnar Jeschke) is available to facilitate analysis of DEER data.

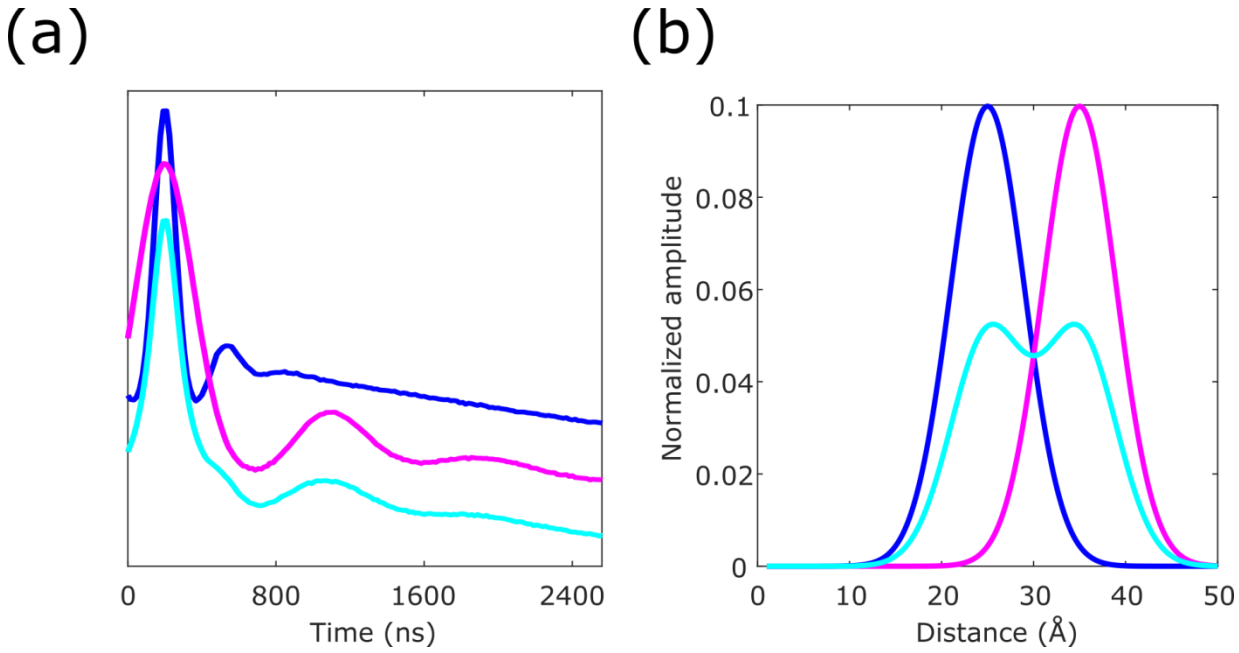

Figure 2.13 In (a), simulated DEER DEFs are shown. Fits to these time-domain distributions provide corresponding distance distributions shown in (b).

2.6 Bibliography

1. Carrington, A. & McLachlan, A. D. *Introduction to magnetic resonance with applications to chemistry and chemical physics*. (Harper & Row, 1967).
2. Marsh, D. Electron spin resonance: Spin labels. in *Molecular Biology, Biochemistry, and Biophysics: Membrane Spectroscopy* (ed. Grell, E.) 51–142 (Springer-Verlag, 1981).
3. Poole, C. P. *Electron spin resonance: A Comprehensive Treatise on Experimental Techniques*. (Dover Publications, 1983).
4. Zavoisky, Y. Paramagnetic relaxation of liquid solutions for perpendicular fields. *J. Phys.* **IX**, 211–216 (1945).
5. Johnson, H. A. Thermal noise and biological information. *Q. Rev. Biol.* **62**, 141–152 (1987).
6. Bloch, F. Nuclear Induction. *Phys. Rev.* **70**, 460–474 (1946).

7. Eaton, G. R. & Eaton, S. S. Resolved Electron-Electron Spin-Spin Splittings in EPR Spectra. in *Spin Labeling: Theory and Applications* (eds. Berliner, L. J. & Reuben, J.) 339–397 (Springer US, 1989). doi:10.1007/978-1-4613-0743-3_7

8. Griffith, O. H., Cornell, D. W. & McConnell, H. M. Nitrogen Hyperfine Tensor and g Tensor of Nitroxide Radicals. *J. Chem. Phys.* **43**, 2909–2910 (1965).

9. Columbus, L. & Hubbell, W. L. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry* **43**, 7273–7287 (2004).

10. Voet, D. & Voet, J. G. *Biochemistry*. (J. Wiley & Sons, 2014).

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

12. Mchaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**, 7692–7704 (1996).

13. López, C. J., Fleissner, M. R., Brooks, E. K. & Hubbell, W. L. Stationary-phase EPR for exploring protein structure, conformation, and dynamics in spin-labeled proteins. *Biochemistry* **53**, 7067–7075 (2014).

14. More, C. *et al.* EPR spectroscopy: a powerful technique for the structural and functional investigation of metalloproteins. *Biospectroscopy* **5**, S3–S18 (1999).

15. Burr, M. & Koshland, D. J. Use of “reporter groups” in structure-function studies in proteins. *Proc Natl Acad Sci U S A* **52**, 1017–1024 (1964).

16. Stone, T. J., Buckman, T., Nordio, P. L. & McConnell, H. M. Spin-labeled biomolecules.

Proc. Natl. Acad. Sci. United States Am. **25**, 1010–1017 (1965).

17. Todd, A. P., Cong, J., Levinthal, F., Levinthal, C. & Hubell, W. L. Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. *Proteins Struct. Funct. Bioinforma.* **6**, 294–305 (1989).

18. Berliner, L. J., Grunwald, J., Hankovszky, H. O. & Hideg, K. A novel reversible thiol-specific spin label: Papain active site labeling and inhibition. *Anal. Biochem.* **119**, 450–455 (1982).

19. Guo, Z., Cascio, D., Hideg, K. & Hubbell, W. L. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci.* **17**, 228–239 (2007).

20. Guo, Z., Cascio, D., Hideg, K., Kálái, T. & Hubbell, W. L. 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 (2007).

21. Fleissner, M. R., Cascio, D. & Hubbell, W. L. Structural origin of weakly ordered nitroxide motion in spin-labeled proteins. *Protein Sci.* **18**, 893–908 (2009).

22. Kroncke, B. M., Horanyi, P. S. & Columbus, L. Structural origins of nitroxide side chain dynamics on membrane protein α-helical sites. *Biochemistry* **49**, 10045–10060 (2010).

23. Cunningham, T. F. *et al.* High-resolution structure of a protein spin-label in a solvent-exposed β-sheet and comparison with DEER spectroscopy. *Biochemistry* **51**, 6350–6359 (2012).

24. Warshaviak, D. T., Serbulea, L., Houk, K. N. & Hubbell, W. L. Conformational analysis of a nitroxide side chain in an α-helix with density functional theory. *J. Phys. Chem. B*

115, 397–405 (2011).

25. Columbus, L., Kálai, T., Jekö, J., Hideg, K. & Hubbell, W. L. Molecular motion of spin labeled side chains in α-helices: Analysis by variation of side chain structure. *Biochemistry* **40**, 3828–3846 (2001).

26. Fraser, R. R., Boussard, G. & Saunders, J. K. Barriers to Rotation about the Sulfur-Sulfur Bond in Acyclic Disulfides. *J. Am. Chem. Soc.* **93**, 3822–3823 (1971).

27. López, C. J., Oga, S. & Hubbell, W. L. Mapping molecular flexibility of proteins with site-directed spin labeling: A case study of myoglobin. *Biochemistry* **51**, 6568–6583 (2012).

28. Budil, D. E., Sanghyuk, L., Saxena, S. & Freed, J. H. Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm. *J. Magn. Reson. - Ser. A* **120**, 155–189 (1996).

29. Schneider, D. J. & Freed, J. Spin Labeling-Theory and Applications. in *Biological Magnetic Resonance* (eds. Berliner, L. J. & Reuben, J.) 1–76 (Springer US, 1989).

30. Liang, Z., Lou, Y., Freed, J. H., Columbus, L. & Hubbell, W. L. A multifrequency electron spin resonance study of T4 lysozyme dynamics using the slowly relaxing local Structure model. *J. Phys. Chem. B* **108**, 17649–17659 (2004).

31. McCoy, J. & Hubbell, W. L. High-pressure EPR reveals conformational equilibria and volumetric properties of spin-labeled proteins. *Proc. Natl. Acad. Sci.* **108**, 1331–1336 (2011).

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

33. Bridges, M., Hideg, K. & Hubbell, W. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* **37**, 363–390 (2010).

34. Fleissner, M. R. *et al.* Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proc. Natl. Acad. Sci.* **108**, 16241–16246 (2011).

35. Percival, P. W. & Hyde, J. S. Saturation-recovery measurements of the spin-lattice relaxation times of some nitroxides in solution. *J. Magn. Reson.* **23**, 249–257 (1976).

36. Jeschke, G. & Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* **9**, 1895–1910 (2007).

37. Hahn, E. L. Spin Echoes. *Phys. Rev.* **80**, (1950).

Chapter 3. Concepts in hydrostatic pressure perturbation

Proteins can exist as a manifold of conformational states in equilibrium under physiological conditions, including the native (ground) and low-lying excited states, with relative populations that correspond to their respective free energies. Because the native state predominates, excited states are considered "invisible" as they have populations below the detection limit of most spectroscopies. These excited states may be conformations essential for protein function[1] and are thus important targets for investigation. Hydrostatic pressure shifts protein conformational equilibrium toward conformations with lower partial molar volumes. Remarkably, low-lying excited states generally have lower partial molar volumes than the corresponding ground states; thus pressure can increase the population of "invisible" excited states to levels suitable for spectroscopic investigations.[2]

3.1 Pressure terms in thermodynamics

The equilibrium of a two-state system, in which a protein exchanges between conformational states A and B, is described by

$$A \rightleftharpoons B \tag{1}$$

for which the general expression of the Gibbs free energy difference between states A and B is

$$\Delta G = \Delta G^0 + \Delta \bar{V}^0 (P - P_0) - \Delta S^0 (T - T_0) \tag{2}$$

where ΔG is the difference in free energy between states A and B, ΔG^0, $\Delta \bar{V}^0$, and ΔS^0 are the standard differences at ambient pressure and ambient temperature in free energy, partial molar volume, and entropy, respectively, P and P_0 are the variable and atmospheric pressures, and T and T_0 are the variable and ambient temperatures. The partial molar volume of a state $\bar{V}^0$ is the sum of its molecular volume and the associated solvent volume or "hydration shell" and can be

decomposed into three component parts[3]

$$\bar{V}^0 = v_W + v_V + \Delta v_h \tag{3}$$

where the molecular volume is the sum of van der Waals volumes of all the protein constitutive atoms v_W and solvent-inaccessible void volume that results from imperfect atomic packing v_V, plus the hydration volume, which is the change in solvent volume associated with protein atomic groups Δv_h. The change in the partial molar volume represents differences in both molecular packing and hydration of the protein and, as will be discussed, the solvent interactions may not always be limited to the solvent accessible surface area of the protein native state.

Eq. 2 can be elaborated to incorporate second-order terms. For experiments performed at constant temperature, this expression becomes:

$$\Delta G = \Delta G^0 + \Delta V^0 (P - P_0) - \frac{\Delta \beta_T}{2} (P - P_0)^2 \tag{4}$$

where $\Delta \beta_T$ is the change in isothermal compressibility. Generally, nonzero compressibility applies to equilibria between folded and unfolded conformations of proteins. Because the extended chain of an unfolded protein is incompressible, the magnitude of $\Delta \beta_T$ is sufficiently large and can be measured experimentally[4]. In comparison, the difference in isothermal compressibility between folded protein states is very small.

The equilibrium constant from Scheme 1 is defined at atmospheric pressure and ambient temperature by

$$\Delta G^0 = -RT \ln K_0 \tag{5}$$

and the pressure-dependent equilibrium constant $K(P)$ relative to K_0 is given by the difference in the $\ln K$ terms of Eq. 5 and Eq. 4 substituted into Eq. 5

$$\ln \left(\frac{K}{K_0} \right) = -P \cdot \frac{\Delta V^0}{RT} + P^2 \cdot \frac{\Delta \beta_T}{2RT} \tag{6}$$

where $P_0 \approx 0$. To illustrate the pressure dependence on protein conformational equilibria,

suppose states A and B make up 10% and 90% of the total protein population at ambient pressure, respectively. The equilibrium constant at ambient pressure and atmospheric temperature is $K_0 = 0.11$. The Gibbs free energy difference is therefore $\Delta G^0 = 1.3$ kcal mol^{-1}. If the difference in partial molar volume between states A and B is -50 ml mol^{-1} (experimental values in the literature are found to be on the order of -25 to -90 ml mol^{-1})[5–7] and the difference in isothermal compressibility is zero, application of 1000 bar repopulates states A and B to 54% and 46% of the total protein population, respectively. This pressure dependence is shown in the top panel of Fig. 3.1a (blue curve). The trend is sigmoidal; in a semilog plot of $\ln K/K_0$ vs. pressure, the trend is linear (Fig. 3.1a, blue line of the bottom panel). In cases when

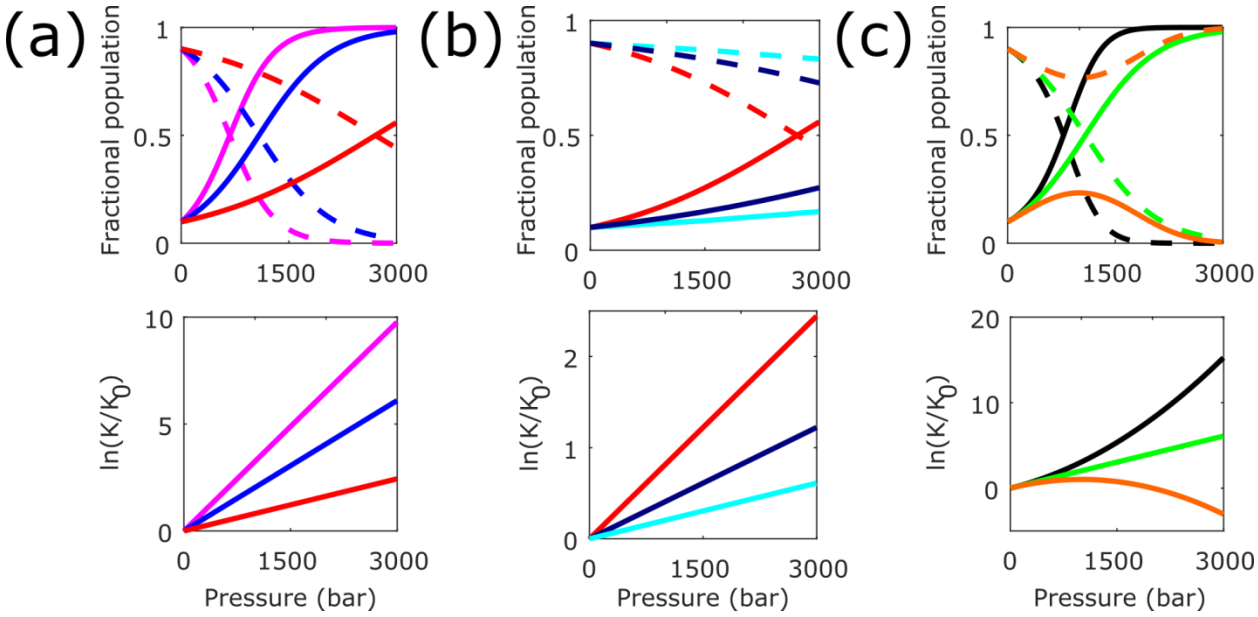

Figure 3.1 For states A (bold lines) and B (dashed lines), pressure-dependent population sigmoids are simulated for different transitions (top) and plotted according to Eq. 6 (bottom). In (a), $\boldsymbol{K_0} = 0.11$, $\boldsymbol{\Delta \bar{V}^0}$ = -80 ml mol^{-1} (magenta), -50 ml mol^{-1} (blue), and -20 ml mol^{-1} (red). The midpoint of each curve shifts due to the change in $\boldsymbol{\Delta G/\Delta \bar{V}^0}$. An extension of the curves in (a) are shown in (b), where pressure dependence appears linear due to small $\boldsymbol{\Delta \bar{V}^0}$ and the midpoint of each sigmoid is not resolved. $\boldsymbol{\Delta \bar{V}^0}$ = -20 ml mol (red), -10 ml mol-1 (navy), -5 ml mol-1 (cyan). In (c), the influence of the difference in isothermal compressibility between states A and B is shown; $\boldsymbol{K_0} = 0.11$, $\boldsymbol{\Delta \bar{V}^0}$ = -50 ml mol^{-1}, and $\boldsymbol{\Delta \beta_T}$ = -0.025 ml mol^{-1} bar^{-1} (orange), 0 ml mol^{-1} bar^{-1} (green), and 0.025 ml mol^{-1} bar^{-1} (black). Generally it is expected that the higher-energy state B has a smaller compressibility than state A (see Section 3.2) such that $\boldsymbol{\Delta \beta_T < 0}$.

$\Delta\bar{V}^0$ is small (0 – 20 ml mol^{-1}), it is often the case that the sigmoid transition is not apparent and the relationship between fractional population and pressure appears linear (Fig. 3.1a, red curves). In these cases, it is not possible to distinguish conformational exchange from compression due to the lack of obvious two-state behavior (Fig. 3.1b). Additional experiments characterizing these equilibrium states are therefore necessary to comment on the mechanism of pressure that produces shallow sigmoids or linear transitions. Lastly, the influence of a difference in compressibility between states A and B is shown in Fig. 3.1c, which distorts the sigmoid shape depending on the magnitude and sign of $\Delta\beta_T$ (refer to Eq. 6). A small, negative $\Delta\beta_T$ was identified for pressure-induced unfolding of T4 lysozyme in mild denaturant conditions[4], consistent with an excited state that is less compressible than the native state.

3.2 The Volume Theorem

So far, the pressure dependence of an equilibrium between two states has been described in terms of thermodynamic parameters, but the structural origin of this effect has not been discussed. The "Volume Theorem," proposed by Kazuyuki Akasaka, is an empirical correlation between the partial molar volume and conformational order of each state.[8] In Chapter 1, the importance of conformational order and flexibility for protein function was emphasized. The Volume Theorem connects the current thermodynamic framework to the pressure-dependent structural properties of conformational states of a protein. The theorem states:

> "*Partial molar volume of a globular protein decreases in parallel with the loss of its conformational order*."

Although the Volume Theorem was proposed in light of pressure denaturation (i.e. population of a non-native conformational state) of apo myoglobin, interpretation of changes in structure for

pressure denaturation for a large variety of proteins, including ß-lactoglobulin[7], dihydrofolate reductase[5], ubiquitin[6], and prion protein[9] by NMR and chymotrypsinogen[10], lysozyme[10], riboflavin binding protein[11], staphylococcal nuclease[12], and Trp repressor[13] by intrinsic fluorescence among many others are all consistent with the Volume Theorem, irrespective of secondary structure or global fold. In Chapter 1, the conformational landscape of a protein was described as funnel-shaped, in which the native state at the bottom is separated from unfolded conformations at the top by an increasing number of intermediate conformations. These conformational states are organized bottom-up with decreasing order and decreasing volume, in parallel with the Volume Theorem and the effect of pressure perturbation on protein structure.

Pressure denaturation of proteins was first demonstrated over 100 years ago by egg white coagulation.[14] Since the mid-1990s, however, interest in the effects of pressure on protein structure has renewed due to the development of instruments capable of implementing variable pressure, such as specialized windows for optical spectroscopy. An on-line high-pressure cell for 2D NMR has been particularly important.[15]

From these advancements, pressure-denaturation of many globular proteins is often found to be stepwise, revealing conformational intermediates at moderate pressures. According to the Volume Theorem, these intermediates have increasingly smaller partial molar volumes and reduced conformational order. The first-order term of Eq. 4 demonstrates this relationship:

$$\Delta G \propto \Delta \bar{V}^0 (P - P_0) \tag{7}$$

3.3 Protein compressibility

In addition to shifting conformational equilibria, pressure exerts an effect on the conformational states themselves. At elevated pressures, states A and B may have marginally

smaller partial molar volumes than they do at ambient pressure due to their partial molar compressibilities. The isothermal compressibility of a protein conformational state is

$$\beta_T = -\left(\frac{\delta V}{\delta P}\right)_T \tag{8}$$

This value, normalized by $\bar{V}^0$, is on the order of $7\times10^{-6} - 10\times10^{-6}$ bar^{-1} for globular proteins.[16] In the Frauenfelder conceptual framework from Chapter 1, structural fluctuations of a conformational state are due to exchange between statistical substates. As such, the change in partial molar volume due to compression corresponds to a shift in the relative population of statistical substates within a conformational state energy well. Fig. 3.2 is a graphical depiction of the difference between pure compression of a protein conformation and shifts in protein conformational states as a result of pressure.

From Eq. 3, it was shown that the partial molar volume of a state is the sum of molecular and solvent volumes. Since the constitutive van der Waals volume v_W is considered incompressible, only cavity and hydration volumes are susceptible to pressure perturbation. A statistical mechanics approach shows the mean squared volume fluctuation within a conformational state is related to the compressibility of said state by

$$\langle(\delta V)^2\rangle = \beta_T k_B T \tag{9}$$

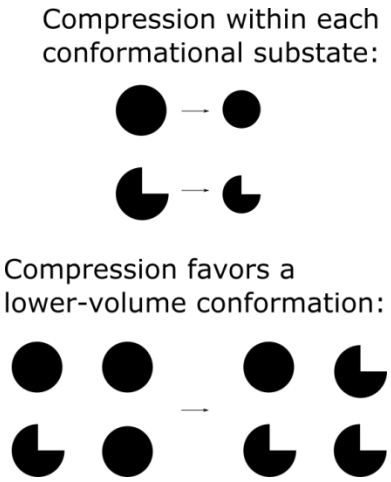

Figure 3.2 Transitions between statistical substates, above, and transitions between conformational states, below. Two conformational states are indicated by different schematic shapes. Compressible states have smaller partial molar volumes at higher pressures. A protein in equilibrium between two states has fractional populations related to their relative free energies.

Shown diagrammatically in Fig. 3.3a, the fluctuation width of a protein conformation about a mean volume is made up of a distribution of statistical substates (solid Gaussian curve). Compression of this state reduces its partial molar volume to a small degree, shifting the average volume of the distribution (dashed Gaussian curve). Fig 3.3b depicts two conformational states A and B and their respective volume fluctuations. At high pressure, the areas under each curve change as the populations shift between states, as do the volume distribution centers as a function of compressibility.

The "adhesive-cohesive" model for protein compressibility relates the partitioning of charged amino acid side chains between the solvent accessible surface area (SAS) and solvent excluded surface area (SES) of a protein[17] (Fig. 3.4). The relative compressibilities of different proteins were found to correlate with localization of charged amino acids, leading to a model in which strong electrostatic interactions between protein SAS and solvent form attractive forces

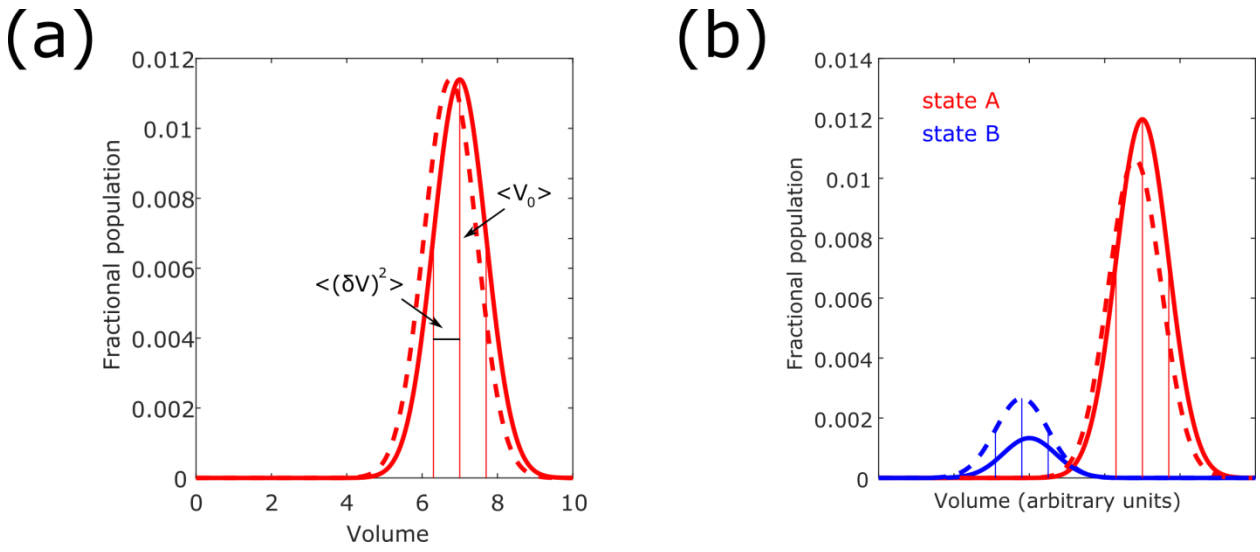

Figure 3.3 A schematic description of compressibility of a single conformational state (a) and pressure-induced shifts between conformational states (b). Individual instantaneous volumes of taxonomic substates form a Gaussian distribution centered around the mean volume $\langle \boldsymbol{V} \rangle$. The compressibility of each state is a function of the distribution width (Eq. 9). In this notation, the change in partial molar volume is $\boldsymbol{\Delta \bar{V}^0 = \langle V_{B,0} \rangle - \langle V_{A,0} \rangle}$ and the change in isothermal compressibility $\boldsymbol{\Delta \beta_T = \langle (V_B)^2 \rangle / k_B T - \langle (V_A)^2 \rangle / k_B T}$.

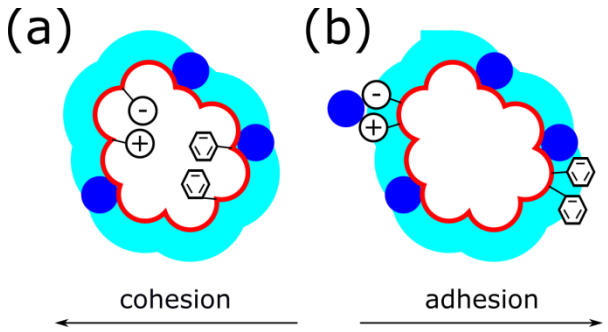

Figure 3.4 Partitioning of charged side chains in the SE volume and SAS of a protein. In (a), internal charged and hydrophobic side chains both contribute to protein stability. In (b), solvent exposure of charged and hydrophobic side chains is correlated with reduced isothermal compressibility $\boldsymbol{\Delta\beta_T}$ and heat capacity $\boldsymbol{\Delta C_P}$.

(adhesive forces) that compete with favorable intramolecular interactions in the hydrophobic protein interior (cohesive forces). Molecular dynamics simulations suggest the per-residue hydrophobic van der Waals energy on the SES of small globular proteins is nearly equal, while the per-residue electrostatic energy from buried salt bridges was found to exceed the per-residue hydrophobic energy by 5-10 times. This difference supports the significant stabilizing effect of few charge-charge interactions contained in the SES and infers that enthalpic stability is gained for proteins with even charge distributions between SES and SAS. Proteins with small compressibility have balanced adhesive and cohesive forces, while proteins with high compressibility have an excess of surface charge, such that the adhesive forces outweigh the cohesive forces. In other words, contraction of protein molecular volume via cohesive interactions is counteracted by adhesive interactions between SAS and solvent which expand the protein volume. The adhesion-cohesion model also accounts for correlations between compressibility and heat capacity, as well: because compressibility, a function of molecular volume fluctuations, and heat capacity, a function of energy absorption, are sensitive to partitioning of charged side chains in a protein with a desolvated core, both reflect structural stability of a conformational state. It is widely recognized that hydration of apolar groups increases ΔC_P, where hydration of charged groups decreases ΔC_P. An excess of surface charge compared to the protein interior is therefore dually associated with large isothermal

compressibility and large, positive heat capacity.

3.4 Protein structural changes with pressure

A number of models for pressure effects on protein structure have been described in the literature. An early model for pressure perturbation of proteins involves forced hydration of the hydrophobic core, leading to protein "denaturation"[18] by forming states described as compact-denatured or fully unfolded. Because this model fails to account for structural changes that occur outside of the hydrophobic core, the model was adapted to include the pressure dependence of polar and nonpolar moieties: "pressure denaturation" was reframed by focusing on transfer of the protein backbone from the interior (SES) to solvent (SAS).[19] Although solvation of hydrophobic side chains is increasingly unfavorable as pressure increases, solvation of polar side chains is neither more favorable nor unfavorable[19] and at elevated pressure, the protein "denatures" but maintains a sizeable hydrophobic core. Two implications were made clear in this adapted model: first, that reduction of internal protein void volumes is linked to the free energy of transfer of side chains to water or to the SES, and second, that tightly-packed proteins are "pressure-resistant" (i.e. incompressible). These concepts capitalize on the variability of packing defects among a variety of globular proteins as a means to explain different pressure responses. In a related model, differences in solvent density due to pressure were proposed to cause a SES-to-SAS repartitioning of nonpolar groups by weakening the hydrophobic effect.[20] However, predictions for the volume changes of proteins by these few models are inconsistent with experimental values for changes in partial molar volume, which are found to be negative but small.[21,22] An emerging view of structural changes that does not require a large unfolding transition (hence the ambiguous definition of "pressure denatured" states) involves elimination

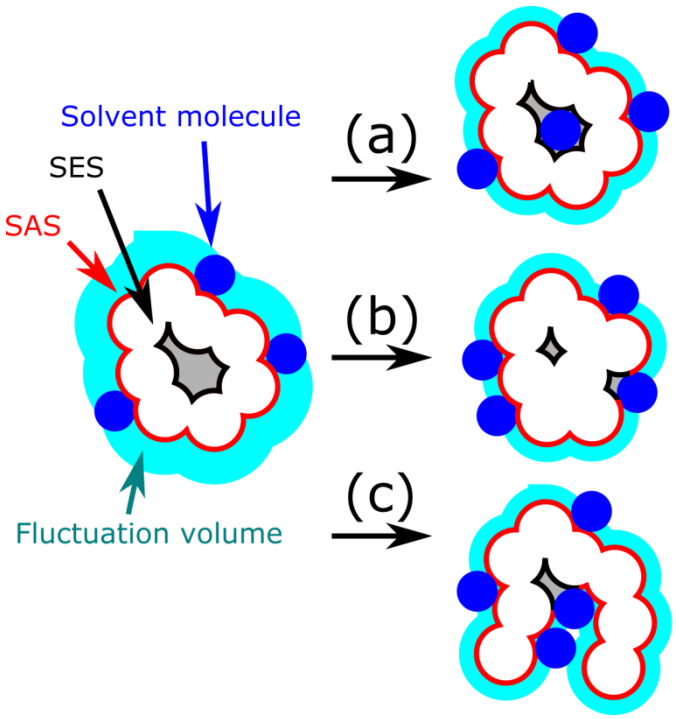

Figure 3.5 On the left, a dynamic protein fluctuates about the mean volume of the native conformational state (cyan boundary). The total volume of is the sum of van der Waals (white) and solvent-excluded (grey) void volumes, which are bounded by SAS and SES, plus the hydration volume of the protein, which includes interacting solvent molecules (sparsely shown as blue spheres). On the right, possible responses of a protein to pressure perturbation include (a) cavity hydration, (b) structure-relaxation, and (c) solvent exposure of SES. These mechanisms are not exclusive and combinations of these transitions can occur heterogeneously throughout a protein.

of native internal protein cavities by hydration[23,24] or by structure-relaxation[24–28]. Fig. 3.5 compares three mechanisms by which new protein conformational states are assumed.

Of all the models discussed above, the elimination of cavities by hydration or structure-relaxation at moderate pressures (1 – 3000 bar) are the currently favored models of pressure perturbations on proteins. The effect of pressure on proteins with enlarged cavities shows increased volume changes on pressure denaturation compared to wild-type structures, indicating that the conformational transition due to pressure is directly related to internal cavity volume.[26,29] Both cavity hydration and structure-relaxation are consistent with population of conformations among a pre-equilibrium of states that have smaller partial molar volumes.

In addition to protein-solvent interactions and intraprotein hydrophobic interactions, pressure may also modulate intraprotein electrostatic interactions. Disruption of an ionic

interaction on the surface of a protein increases the number of hydration waters around the ions. Because the partial molar volume of the hydration waters around an ion is smaller than for bulk water due to electrostriction, pressure acts to dissociate ionic interactions. Often, such electrostatic interactions in proteins are weakly stabilizing, though electrostatic networks have a stronger and often essential stabilizing effect.[30,31]

3.5 Lineshape analysis of variable pressure EPR

The total motion reflected in the EPR line shape of a nitroxide is shown in Chapter 2 to dependen on three isotropic correlation times arising from overall rotary diffusion of the protein (τ_r), backbone motion (τ_{bb}), and internal motions of the nitroxide side chain (τ_{int}). The relationship between these correlation times and the overall observed correlation time is the sum of inverse parts, assuming that the motions are uncoupled and independent:

$$\frac{1}{\tau_R} = \frac{1}{\tau_r} + \frac{1}{\tau_{bb}} + \frac{1}{\tau_{int}} \quad (9)$$

The effect of pressure on the motion of the nitroxide R1 in an α-helix has been investigated by McCoy and Hubbell.[4] For a noninteracting helix surface site at atmospheric pressure, a nitroxide exhibits simple anisotropic motion where $\tau_R = 1 - 2$ ns.[32] Extracting this correlation time from variable-pressure EPR spectra is possible by line shape analysis via spectral simulation (refer to Chapter 2) using a MOMD model with a fixed order parameter. It has been suggested for nitroxides attached to rigid sites on a protein, increasing pressure results in an increase in τ_R reflecting a change predominately in the internal motion of the nitroxide (τ_{int}).[33] More flexible sites of proteins with segmental motion in the α-helices on this time scale can be pressure sensitive, in which case both the internal and backbone motions (τ_{int} and τ_{bb}) will change with pressure.[33] The pressure sensitivity of the internal motion of the nitroxide is interpreted in a

model according to transition state theory as an expansion of the solvent cage around the nitroxide side chain that permits rotation of the spin label.[4,34,35] The overall motion τ_R and the transition state activation volume $\Delta V^{\ddagger}$ are related by

$$\frac{1}{\tau_R(P)} = \frac{1}{\tau_{R,0}} e^{-\frac{P\Delta V^{\ddagger}}{RT}} \tag{10}$$

where $\tau_R(P)$ and $\tau_{R,0}$ are the overall rotational correlation times at applied and atmospheric pressures, respectively. Eq. (11) is simplified for small values of $\Delta V^{\ddagger}$ and P (i.e., < 5 ml/mol and < 2 kbar, respectively) to yield

$$\frac{1}{\tau_R(P)} = \frac{1}{\tau_R,0}\left(1 - \frac{P\Delta V^{\ddagger}}{RT}\right) \tag{11}$$

Site-specific differences in $\Delta V^{\ddagger}$ suggest that variations in τ_R are indicative of variability in the pressure sensitivity of backbone motion across protein structures. Combining Eq. 9 and Eq. 11 yields an expression which separates different types of motion

$$\frac{1}{\tau_R(P)} = \frac{1}{\tau_R,0} - \frac{1}{\tau_{bb},0}\left(\frac{P\Delta V_{bb}^{\ddagger}}{RT}\right) - \frac{1}{\tau_{int},0}\left(\frac{P\Delta V_{int}^{\ddagger}}{RT}\right) \tag{12}$$

where $\Delta V_{bb}^{\ddagger}$ and $\Delta V_{int}^{\ddagger}$ are the activation volumes arising from protein backbone fluctuations in the transition state and solvent reorganization in the transition state. The smallest pressure-dependent change in τ_R observed in the rigid protein holomyoglobin was used as a benchmark to determine an activation volume for the sidechain[33]: $1/\tau_{int,0} = 5\times10^8$ sec^{-1} for a nitroxide placed in a rigid helix, i.e. $1/\tau_{bb,0}$ is very small. For more mobile helices, the quantities $1/\tau_{bb,0}$ and $\Delta V_{bb}^{\ddagger}$ were solved for given values of $\tau_{int,0}$ and $\tau_{int}(P)$ from simulations of holomyoglobin spectra. With this model, the local structure of a protein can be surveyed in order to find regions of high and low pressure sensitivity, which may be coupled to protein function.

Conformational states exchanging slower than 100 ns are observed as discrete spectral components in EPR spectra at X-band. Quantification of the pressure-dependent populations of

these states by simulation provide equilibrium constants as a function of pressure, which can be used to determine thermodynamic parameters that relate the states (refer to Eq. 6). A few considerations when interpreting EPR line shapes must be mentioned. Although time scale is used to discriminate between slow conformational exchange and fast backbone fluctuations and rotameric exchange, the amplitude of motion is not determined by CW EPR. Additionally, only spectral states with motions that are sufficiently distinct can be separated in the EPR spectrum. To this point, it is possible that a multicomponent line shape indicates exchange between statistical substates on the high nanosecond time scale rather than conformational exchange on the μs-ms time scale. Alternatively, two states could be undetectable by virtue of nitroxide motion with similar correlation times. In practice, and despite these potential insensitivities, SDSL EPR is shown to be a sensitive, quantitative technique to study protein structure and conformational exchange in combination with pressure perturbation.

3.6 Bibliography

1. Baldwin, A. J. & Kay, L. E. NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5**, 808–814 (2009).
2. Akasaka, K. Probing Conformational Fluctuation of Proteins by Pressure Perturbation. *Chem. Rev.* **106**, 1814–1835 (2006).
3. Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. in *Advances in Protein Chemistry* (eds. Anfinsen, C. B. J., Anson, M. L., Bailey, K. & Edsall, J. T.) 1–63 (Academic Press, 1959).
4. McCoy, J. & Hubbell, W. L. High-pressure EPR reveals conformational equilibria and volumetric properties of spin-labeled proteins. *Proc. Natl. Acad. Sci.* **108**, 1331–1336

(2011).

5. Kitahara, R. *et al.* High pressure NMR reveals active-site hinge motion of folate-bound Escherichia coli dihydrofolate reductase. *Biochemistry* **39**, 12789–12795 (2000).

6. Kitahara, R., Yamada, H. & Akasaka, K. Two folded conformers of ubiquitin revealed by high-pressure NMR. *Biochemistry* **40**, 13556–13563 (2001).

7. Kuwata, K. *et al.* High pressure NMR reveals a variety of fluctuating conformers in β-lactoglobulin. *J. Mol. Biol.* **305**, 1073–1083 (2001).

8. Kitahara, R., Yamada, H., Akasaka, K. & Wright, P. E. High pressure NMR reveals that apomyoglobin is an equilibrium mixture from the native to the unfolded. *J. Mol. Biol.* **320**, 311–319 (2002).

9. Kuwata, K. *et al.* Locally disordered conformer of the hamster prion protein: A crucial intermediate to PrPSc? *Biochemistry* **41**, 12277–12283 (2002).

10. Thomas, M. L., Hook, J. W., Drickamer, H. G. & Weber, G. Plurality of Pressure-Denatured Forms in Chymotrypsinogen and Lysozyme. *Biochemistry* **15**, 5571–5580 (1976).

11. Thomas, M. L., Weber, G., Hook, J. W. & Drickamer, H. G. Effects of Pressure upon the Fluorescence of the Riboflavin Binding Protein and Its Flavin Mononucleotide Complex. *Biochemistry* **15**, 3205–3211 (1976).

12. Vidugiris, G. J. A., Markley, J. L. & Royer, C. A. Evidence for a molten globule-like transition state in protein folding from determination of activation volumes. *Biochemistry* **34**, 4909–4912 (2005).

13. Desai, G., Panick, G., Zein, M., Winter, R. & Royer, C. A. Pressure-jump studies of the folding/unfolding of trp repressor. *J. Mol. Biol.* **288**, 461–475 (1999).

14. Bridgman, P. W. The coagulation of albumen by pressure. *J. Biol. Chem.* **19**, 511–512 (1914).

15. Yamada, H. *et al.* Pressure-resisting cell for high-pressure, high-resolution nuclear magnetic resonance measurements at very high magnetic fields. *Rev. Sci. Instrum.* **72**, 1463–1471 (2001).

16. Gekko, K. Protein Dynamics and Function as Revealed by Compressibility. in *High Pressure Bioscience and Biotechnology* **1**, 13–19 (2007).

17. Dadarlat, V. M. & Post, C. B. Adhesive-cohesive model for protein compressibility: An alternative perspective on stability. *Proc. Natl. Acad. Sci. USA* **100**, 14778–14783 (2003).

18. Hummer, G. *et al.* The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proc. Natl. Acad. Sci. USA* **95**, 1552–1555 (1998).

19. Chalikian, T. V. & Macgregor, R. B. J. Origins of Pressure-Induced Protein Transitions. *J. Mol. Biol.* **394**, 834–842 (2009).

20. Grigera, J. R. & McCarthy, A. N. The behavior of the hydrophobic effect under pressure and protein denaturation. *Biophys. J.* **98**, 1626–1631 (2010).

21. Chalikian, T. V. & Breslauer, K. J. On volume changes accompanying conformational transitions of biopolymers. *Biopolymers* **39**, 619–626 (2002).

22. Royer, C. A. Revisiting volume changes in pressure-induced protein unfolding. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1595**, 201–209 (2002).

23. Nucci, N. V, Fuglestad, B., Athanasoula, E. A. & Wand, A. J. Role of cavities and hydration in the pressure unfolding of T4 lysozyme. *Proc. Natl. Acad. Sci.* **111**, 13846–13851 (2014).

24. Lerch, M. T. *et al.* Structure-relaxation mechanism for the response of T4 lysozyme cavity mutants to hydrostatic pressure. *Proc. Natl. Acad. Sci.* **112**, E2437–E2446 (2015).

25. Vidugiris, G. J. A., Truckses, D. M., Markley, J. L. & Royer, C. A. High-pressure denaturation of staphylococcal nuclease proline-to-glycine substitution mutants. *Biochemistry* **35**, 3857–3864 (1996).

26. Lopez, C. J., Yang, Z., Altenbach, C. & Hubbell, W. L. Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc. Natl. Acad. Sci.* **110**, E4306–E4315 (2013).

27. Lerch, M. T., Horwitz, J., McCoy, J. & Hubbell, W. L. Circular dichroism and site-directed spin labeling reveal structural and dynamical features of high-pressure states of myoglobin. *Proc. Natl. Acad. Sci.* **110**, E4714–E4722 (2013).

28. Maeno, A. *et al.* Cavity as a source of conformational fluctuation and high-energy state: High-pressure NMR study of a cavity-enlarged mutant of T4lysozyme. *Biophys. J.* **108**, 133–145 (2015).

29. Bouvignies G, Vallurupalli P, Hansen DF, Correia BE, Lange O, Bah A, Vernon RM, Dahlquist FW, Baker D, K. LE. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111–114 (2013).

30. Nakamura, H. Roles of electrostatic interaction in proteins. *Q. Rev. Biophys.* **29**, 1–90 (1996).

31. Kumar, S. & Nussinov, R. Close-Range Electrostatic Interactions in Proteins. *ChemBioChem* **3**, 604–617 (2002).

32. Columbus, L., Kálai, T., Jekö, J., Hideg, K. & Hubbell, W. L. Molecular motion of spin labeled side chains in α-helices: Analysis by variation of side chain structure.

Biochemistry **40**, 3828–3846 (2001).

33. Lerch, M. T., Yang, Z., Altenbach, C. & Hubbell, W. L. *High-Pressure EPR and Site-Directed Spin Labeling for Mapping Molecular Flexibility in Proteins*. *Methods in Enzymology* **564**, (Elsevier Inc., 2015).

34. Dadali, A. A., Buchachenko, A. L. & Irzhak, V. I. Effect of pressure on the rotational mobility of spin probes in polymers. *Eur Polym J* **17**, 525–532 (1981).

35. Dadali, A., Barashkova Lastenko, I. P. & Wasserman, A. M. Effect of pressure on the rotational mobility of spin label in polymer. **27**, 1097–1100 (1991).

Chapter 4. Conformational equilibria of the intestinal fatty acid binding protein

The "invisible" conformational state of intestinal fatty acid binding protein (I-FABP) that permits ligand entry into the protein interior may be the molten globule (MG) populated at low pH. Despite a wealth of X-ray crystal and NMR solution structures (18 deposited in the RCSB PDB database) published between 1989[1] and 2014[2], the conformational rearrangement that opens a channel to the cavity interior has not been unequivocally identified. A variety of solvent channels, identified both from these structures as regions of high mobility and others from molecular dynamics simulations are possible points of entry for a fatty acid ligand. That the MG state of I-FABP is capable of binding fatty acids suggests that the conformational state is heterogeneous and is difficult to characterize by classical biochemical methods. In this chapter, the native and acid-pH stabilized MG states of I-FABP are investigated by site-directed spin labeling (SDSL) paired with electron paramagnetic resonance (EPR) to reveal significant folded structure at all pH conditions and conformational rearrangements in putative portal regions. This work emphasizes the versatility of SDSL EPR to study the MG of a β-sheet protein, which is a conformational state not well characterized due to the complexity and variability of β-sheets and their motions.

4.1 Background to conformational states of I-FABP

Intestinal fatty acid binding protein (I-FABP) is a small cytosolic protein that binds long-chain fatty acids stoichiometrically. As a cytosolic lipid chaperone, I-FABP extracts fatty acids from intestinal cell membranes via a collisional mechanism as a key component of fatty acid metabolism.[3,4] It belongs to a superfamily of tissue-specific intracellular lipid binding proteins (iLBPs) which have characteristic β-barrels that bind hydrophobic ligands internally and short

peripheral helices. The FABP family within the iLBP superfamily has a highly conserved fold, consisting of a flattened 10-stranded β-barrel formed by two 5-stranded antiparallel β-sheets (Fig 4.1). Three secondary structure features are found across all FABP isoforms: first, the N-terminal half of strand A forms backbone hydrogen bonds with strand B until a kink in strand A, which then hydrogen bonds with strand J, closing the barrel structure. Second, strands A and B are connected by a helix-turn-helix domain, unlike the short β-turns that connect the remaining strands. Third, strands D and E are non-hydrogen-bonded and are stabilized instead by π-stacking interactions between aromatic side chains facing the cavity interior. The X-ray crystal structures of apo and holo I-FABP are shown in Fig 4.1. From the structural biology perspective, I-FABP is an ideal model protein for studying beta sheet structure and dynamics. The rat and human isoforms contain no native cysteine or proline residues, which aid its fast and reversible folding and unfolding pathways.[5]

The ligand binding surface of apo I-FABP is a large hydrated cavity with a volume of 540 $Å^3$ occupied by and a very rough internal surface.[6] Unexpectedly for a protein with specificity towards binding hydrophobic ligands, nearly 25% of the cavity surface is hydrophilic. Bound long chain fatty acids are dynamic within this cavity[7] and have van der Waals volumes (220 – 260 $Å^3$ for 14:0 and 16:0 fatty acids) which make up less than half of the available volume. The remaining volume in the fatty acid-bound X-ray crystal structures of I-FABP is filled with 22 ordered crystallographic waters[8], 4 of which are found by NMR to be tightly associated with the protein with long residence times comparable to the timescale of I-FABP diffusion.[6] Although the fatty acid ligand has considerable conformational entropy in the complex, ligand binding is an enthalpic process, suggesting that fluctuating hydrophobic contacts play a more significant role in stabilizing the holoprotein than the gain in entropy due to cavity

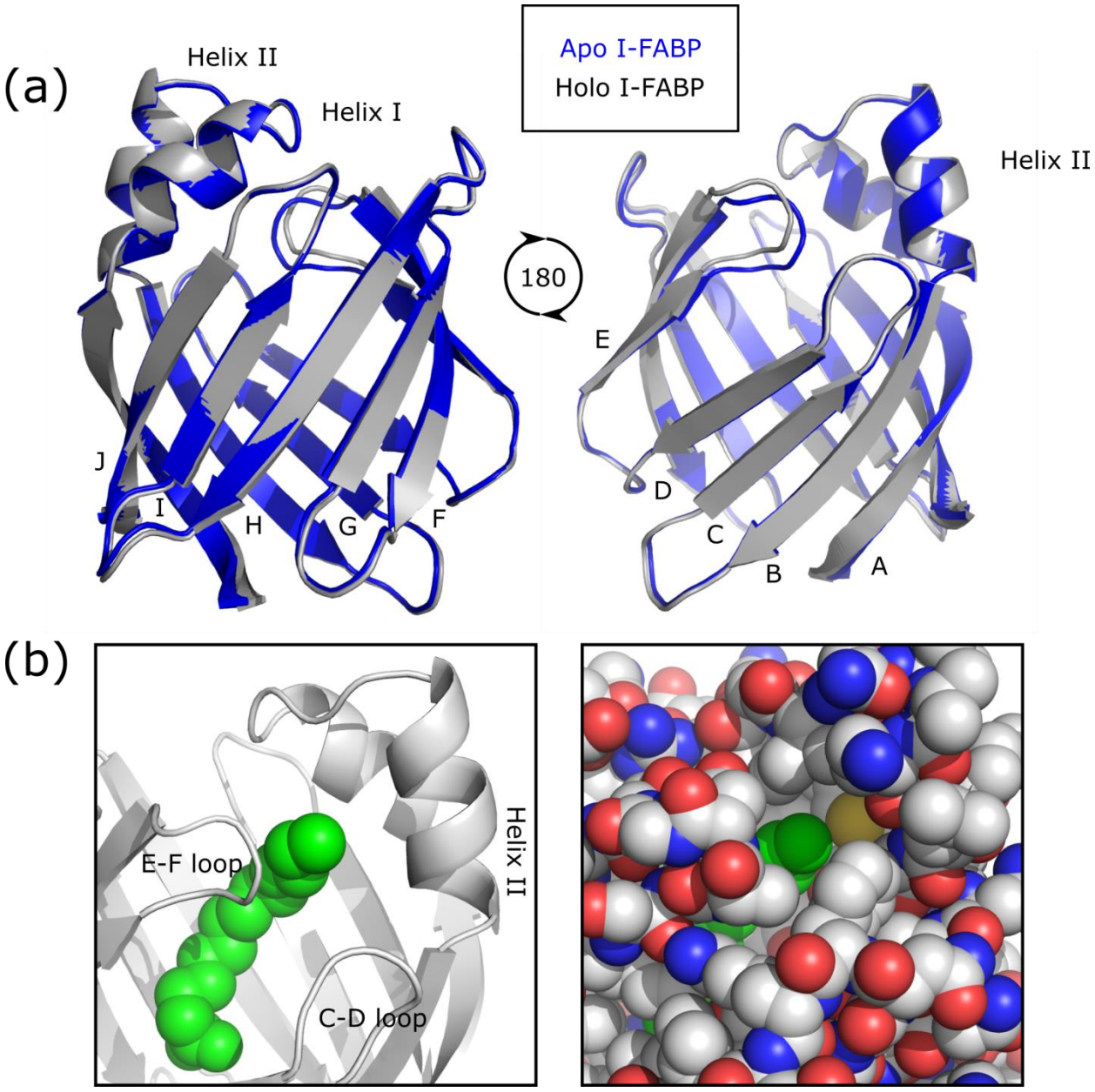

Figure 4.1 X-ray crystal structures of I-FABP in the apo and myristate-bound forms have indistinguishable conformations. In (a), the apo and holo crystals are overlaid in blue and grey, respectively (RCSB PDB accession codes 1IFC and 1ICM). The strands are lettered from A – I in sequential order where each strand is connected by a short β-turn except between strands A and B, which are connected by a helix-turn-helix motif. In (b), the myristate-bound structure is shown to highlight the putative portal. On the left, myristate is modeled in green within the cartoon representation of I-FABP. On the right, the space-filling model of I-FABP shows limited accessibility of solvent to the myristate molecule bound within the cavity.

dehydration, as fatty acid binding displaces all but 8 water molecules.[9] The unfavorable entropy

of fatty acid binding (-10 kcal mol^{-1}) is compensated by this solvent dehydration entropy (10 kcal mol^{-1}), emphasizing the importance of the remaining favorable interaction enthalpy between the fatty acid and internal side chains.[10] In fact, backbone dynamics of the holoprotein are greater than in the apoprotein[11], possibly as a mechanism to offset unfavorable entropy loss. Small differences between the entropic cost and gain of fatty acid binding are thought to be give rise to variations in fatty acid specificity and binding affinity among different FABPs.[12]

Despite the thorough characterization of the apo and holo states of I-FABP, the nature of the "open" conformation remains a mystery. Although the crystal structures of apo and holo I-FABP are nearly superimposable, careful analysis of the space-filling representation of these crystal structures reveals a small channel in both apo and holo I-FABP near the helical domain, which led to identification of a point of entry for ligand binding bounded by the C-terminus of helix II, the C-D loop, and the E-F loop, called the "portal." Evidence for flexibility in this region includes marginally higher thermal crystallographic B-factors, whereas the NMR structure of apo I-FABP also shows somewhat higher backbone flexibility in the portal.[13] It has been hypothesized that small-amplitude shifts in the local structure of the portal could sufficiently open the channel to allow fatty acid access into the cavity. In the absence of direct experimental evidence of an intermediate conformation, simulations have identified fluctuations with short lifetimes in various regions of the protein:

1. The portal is dynamic in most simulations. This region has been identified as the most likely point of entry for ligand binding to I-FABP[14,15] and has been shown to facilitate palmitate exchange into a membrane bilayer by molecular dynamics simulations.[16] However, no passage of fatty acids from solvent into the protein have been identified, although I-FABP is capable of capturing sparingly soluble fatty acids from the

aqueous environment.[17]

2. An alternative portal in myelin FABP between strands D and E, thought to be common to all FABPs, involves unfolding of these strands, forming true edge strands and a discontinuous β-barrel.[18] This movement exposes the fatty acid to solvent laterally and is termed the "cleft" or "gap."
3. Lastly, an "antiportal" between the beta turns opposite the helical domain[1] is a putative solvent channel[19], a region more strongly implicated in ligand binding for adipocyte FABP than I-FABP.[20]

No simulation to date has successfully modeled ligand entry, but most simulations predict widening of various channels that would permit movement of solvent into the cavity. The rate of palmitate dissociation from I-FABP to an acceptor FABP modified with a fluorescent reporter is particularly slow, on the order of 1.8 M^{-1} s^{-1}. Palmitate association, on the other hand, was found to be 23×10^7 M^{-1} s^{-1}.[21] These rate constants were fit to a model which assumes free fatty acids are transferred through the aqueous phase, where the kinetics of the reporter FABP and I-FABP are assumed to be independent. In a separate study, direct measurement of the kinetics of palmitate binding using stopped-flow intrinsic fluorescence shows biphasic behavior, in which a slow conformational transition interpreted as portal opening occurs on the order of 1000 s^{-1} and subsequent ligand binding with a rate constant of 1.2×10^4 M^{-1} s^{-1}.[22] It is clear from these studies that without a full understanding of the mechanism by which FABPs interact with ligands, interpretation of the kinetic behavior of I-FABP can be misconstrued. So far it has been implied that fatty acids use the same routes as water molecules to access the binding cavity.

To study dynamics of partially unfolded intermediates of I-FABP, fluorescence correlation spectroscopy (FCS) with acid pH-denaturation was used to avoid refractive index

differences that complicate similar urea or guanidinium unfolding experiments.[23] The structural change by monitoring FCS of Alexa488Maleimide-modified I-FABP at V60 (an internal site in strand D) was found to require two states to model changes that accompanied lowering the pH from 7 to 2. While the diffusion coefficient of the protein ensemble decreases sharply between pH 3 and 2, suggesting global unfolding of the protein, a dynamics term in the fit function to the FCS data, used to account for quenching, is present even at the apparent "unfolded" pH 2 state.

Since this work, the acid state of I-FABP has attracted attention as a putative "open" conformation. pH-dependent structural changes have been closely investigated. Fatty acid transfer from I-FABP to acceptor phospholipid vesicles is found to accelerate ligand exchange in a concentration-dependent manner.[24] In particular, collisional fatty acid transfer is dependent on both ionic and hydrophobic interactions[25]; charged phospholipid head groups could create a strong electrostatic potential, forming a proton boundary layer with significantly lower local pH.[26] It has been demonstrated that the MG state of proteins are capable of "melting" into membranes via helical domains[27], where the helix-turn-helix motif of I-FABP in the portal could fit this model. Bulk acidic pH is thought to emulate the local proton boundary layer. This model is in contrast with initial studies which investigated the role of direct interactions between cationic side chains on I-FABP and anionic head groups of membrane lipids[24,28] in mediating the collisional mechanism for ligand transfer, which gave rise to complex trends when incorporating varying lipids into vesicles to alter the charge density.[25] Together, these studies all support the portal as the site for ligand entry.

The conformation of the acid state of I-FABP was investigated by selective labeling of phenylalanines in I-FABP by ^{19}F fluorine NMR to provide a coarse map of side chain dynamics and local structure.[29] Line shape analysis of the 1D NMR spectra indicate changes in tertiary

structure in the protein at pH values below pH 4.1, while chemical shift changes are apparent below pH 3.9. At pH 2.8, the eight phenylalanines of apo I-FABP produce ~15 peaks, indicating extensive heterogeneity on the ms time scale, with chemical shifts that are dissimilar to those in the urea-denatured state.[30] At pH 2.3, the conformation of apo I-FABP is more similar to the denatured state. Despite poor chemical shift dispersion, the data indicated that the acid-stabilized conformation of I-FABP is in equilibrium with the native state, seen by small resonance outside the denatured range. It was found that the apo I-FABP ensemble at pH 2.3 is capable of binding oleate, though with lower affinity, either by the MG state or by the small population of the native conformation.

The far-UV CD spectrum of I-FABP at pH 2.8 is superimposable with the pH 7.3 state, indicating native-like secondary structure.[29] At pH 2.3, the secondary structure content is lower, implicating population of an unfolded conformation. However, typical MGs can maintain native-like secondary structure content ranging from 85%[31] to 65%[32], and as such, the loss of secondary structure at pH 2.3 could be within this range. A survey of CD spectra of acid pH-stabilized MGs show that, for the closely related retinol-binding protein (RBP, an iLBP), and other predominately β-sheet proteins such as ribonuclease A and carbonic anhydrase, the CD spectra show less secondary structure than the corresponding native conformations.[33] Although secondary structure quantification of the far-UV CD spectra at any pH for apo I-FABP is not reported, possibly due to significant noise in the 190 – 200 nm range, quantification of the secondary structure content for RBP determined by the computational Levitt and Greer method[34] indicates that in the native state at pH 7.4, the protein contains 11% a-helix and 59% b-sheet, which decrease to 6% a-helix and 54% β-sheet in the acid form at pH 2.0, corresponding to 86% of the native secondary structure in the MG.[33] Qualitatively, the far-UV CD of apo I-FABP in the

native and MG states resemble the CD spectra of these small β-sheet conformations at equivalent pH conditions, as might be expected for similar structures.

This work investigates the MG-like structure of the acid state of I-FABP. First, the native state at pH 7.3 is characterized using SDSL EPR, a highly sensitive spectroscopy that encodes both structure and dynamics of a protein on a site-specific basis. This work extends the sparse sampling by fluorescence and 19F NMR methods to investigate the structure and dynamics of the helical domain, the solvent-facing strand surfaces, the cavity interior, and the cleft strands as a function of pH.

4.2 Structure of the β-domain: CW EPR of the solvent-facing surface of I-FABP

The conformation of the acid state of I-FABP is thought to be heterogeneous and dynamic[29]. To characterize the MG of I-FABP, the faster time scale of EPR reveals motions not detected by ^{19}F NMR or FCS[23,29], which allows for characterization of fast-fluctuating protein structure. The nitroxide line shape is sensitive to side chain packing, which makes significant contributions to motion of the R1 side chain in the solvent-facing β-sheet. Because the R1 side chain structure is aprotic, the line shape should have no intrinsic pH dependence. Changes in the EPR spectra therefore result from changes in protein conformation and local side chain orientation due to protonation. For slow exchange between states (at X-band, > 100 ns), EPR line shapes are resolved as complex or multicomponent spectra.

As a point of comparison to line shapes observed in β-strands, a variety of nitroxide line shapes are observed in an α-helical MG such as the pH 4.1 state of apomyoglobin (apo Mb). All R1 spectra in the apo Mb MG were found to be multicomponent, although many of the corresponding line shapes in the holo- and apoprotein states at pH 6.1 were single component.[35]

Because side chains in the β-sheet are more tightly packed than the splayed arrangement in α-helices, local contacts have significant influence on R1 mobility. This complexity on the EPR time scale is consistent with a dynamic and fluctuating MG state, in which the tertiary structure of the MG is perturbed and the new spectral components were found to be indicative of new conformational arrangements. As such, the rotamer exchange of the nitroxide and the backbone motion of the protein are difficult to distinguish.

In the β-domain, 21 single cysteine substitution mutations were prepared and modified with MTSL to generate proteins containing the R1 side chain (Fig. 4.2). The sites were selected to survey the solvent-facing surface of four beta strands: strands A, B, C, and D on the N-terminal half of the β-domain, and strands F, G, H, and I on the C-terminal half.

To avoid the protein aggregation and precipitation that often occurs in MGs including that of I-FABP, spin labeled mutants were immobilized on a solid support in a scheme

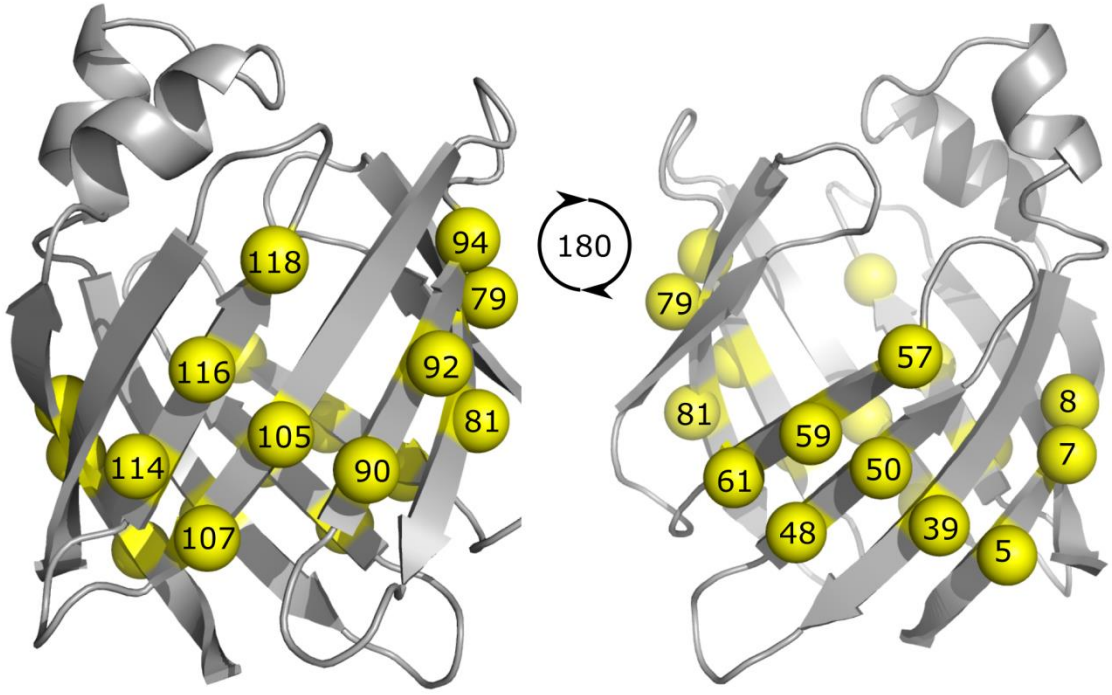

Figure 4.2 Sites in the β-domain of apo I-FABP subject to SDSL EPR.

developed by Lopez *et al*.[36] For this purpose, an unnatural amino acid, para-azidophenylalanine, was incorporated after the C-terminal residue during protein expression using the method of Schultz.[37] The para-azidophenylalanine side chain is modified after protein purification with dibenzocyclooctyne-amine-biotin, a Click Chemistry reagent, to biotinylate the C-terminus with a minimal linker length of 18 carbons, which was preferable to an alternative method for enzymatic C-terminal biotinylation (Avidity AviTag™, which requires a 15 amino acid recognition sequence for the biotin ligase enzyme). The biotinylated construct is immobilized onto streptavidin-functionalized Sepharose beads. Both the covalent attachment of the biotin and the streptavidin-biotin interactions are stable at pH 2.3 and all spectral changes at pH 2.3 were completely reversible, thus solving the problem of protein aggregation.

Lietzow and Hubbell took a systematic approach to investigate nitroxide side chain dynamics in β-sheets[38], which deviates from the "X_4/X_5" model that describes R1 motion in the α-helix (refer to Chapter 2). Mutagenesis of neighboring amino acids revealed nitroxide motions that vary with the identity and position of sidechain contacts. SDSL of solvent-facing "convex" β-sheets can be summarized by three main points:

1. β-branched and bulky residues in the non-hydrogen bonded (NHB) position give rise to ordered states in the EPR line shape. At the hydrogen-bonded (HB) position, this immobilizing effect is weaker, possibly due to the ~1 Å increased distance between the contact pair.
2. Strand twist (and at strand termini) reorganizes the "ideal" β-sheet and brings different sidechains into contact with the nitroxide, resulting in a context-dependent origin of new ordered states in the EPR line shape.
3. In some unusual cases, polar interactions between R1 and a neighboring sidechain

produce small populations of a new ordered state in the EPR line shape.

Non-barrel β-sheets have "edge strands" that lack either a HB or NHB partner. In I-FABP, a unique NHB cleft between strands D and E is stabilized through π-stacking interactions on the cavity-facing strand surface, producing an unusual β-sheet structure.

4.2.1 Results: SDSL EPR of the native state of the I-FABP β-domain

According to the nearest-neighbor amino acid identity in the antiparallel β-sheet outlined by Lietzow and Hubbell, the differences in line shapes of solvent-facing nitroxides in I-FABP can be rationalized. In contrast, sites which are unaffected by nearest-neighbor contacts of the types identified by Lietzow and Hubbell (β-branched or bulky NHBs, significant twist, or polar interactions) are expected to have mobilities reflecting backbone motion.

In the native state of I-FABP at pH 7.3, SDSL EPR line shapes of nitroxides on the solvent-facing surface of β-strands exhibit a wide range of mobilities (Fig. 4.3). Because R1 is a sensitive reporter of local structure, variations in packing density are revealed in the line shape. In strand C, 50R1 has no NHB neighbor, while 48R1 NHB does. The line shape of 48R1 is more immobilized than 50R1. In strand F, both 79R1 and 81R1 have no NHB neighbors and have similarly mobile line shapes. In strand G, 90R1 and 92R1 both are adjacent to NHB neighbors, and have line shapes that are more immobilized than 94R1. Lastly, 105R1 in strand H has a NHB (V90) and its corresponding line shape is immobilized compared to the line shape of 107R1, which has no NHB. The roles of HB and NHB neighbor identity on the R1 line shape in I-FABP is consistent with the rationale proposed for SDSL of the antiparallel β-sheet of the model protein used in Lietzow and Hubbell.[38]

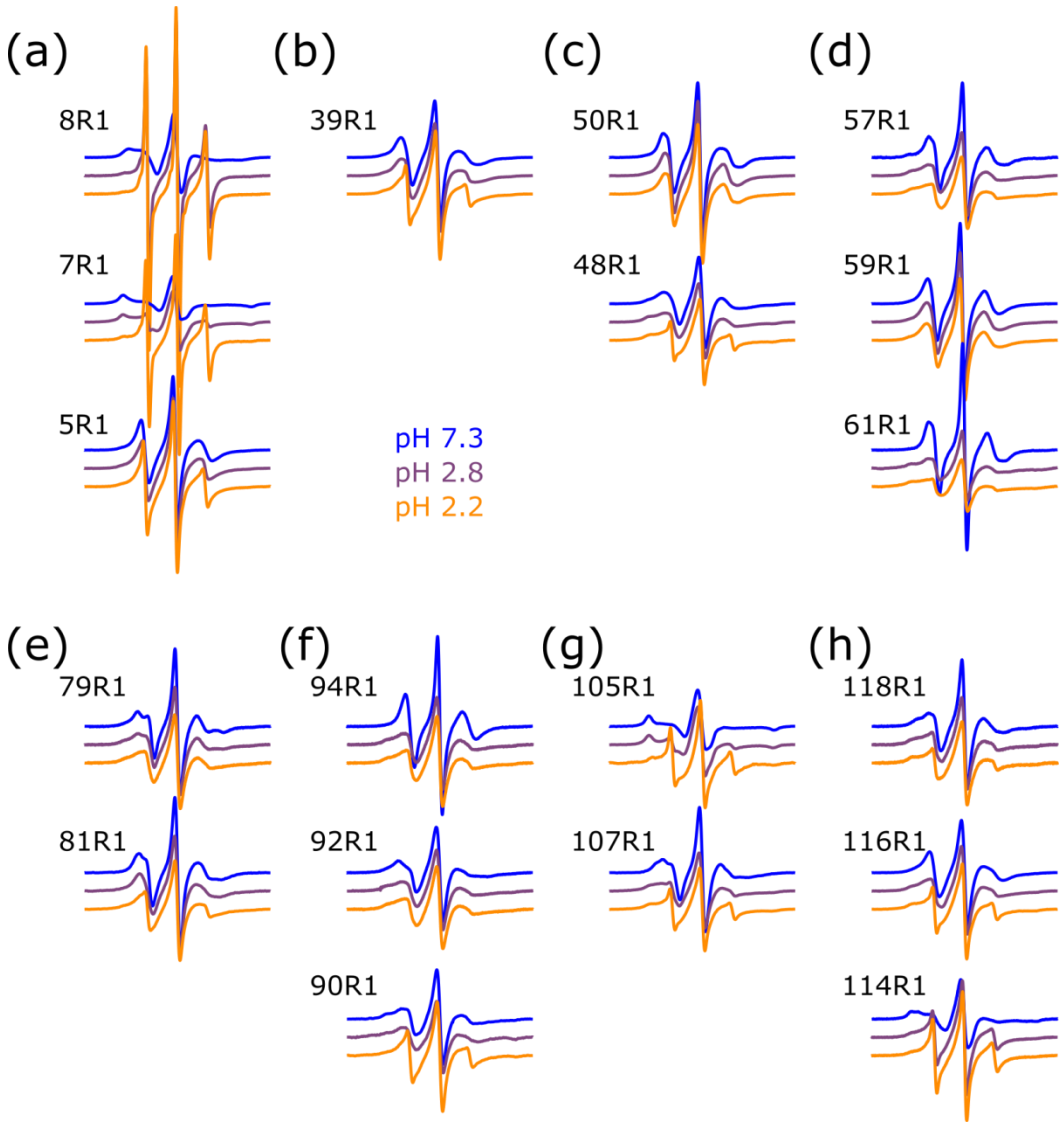

Figure 4.3 CW EPR spectra of apo I-FABP on the solvent-facing surface of β-strands (a) A, (b) B, (c) C, (d) D, (e) F, (f) G, (g) H, and (h) I. Nitroxides close to the portal are at the top of each column.

The mobilities of nitroxides in strands A and D are exceptions to this pattern. First, in strand A, 7R1 is more immobilized than 5R1, even though 5R1 has a β-branched NHB neighbor (T39). These sites are close to the conserved kink in the strand, which disrupts the backbone hydrogen bonding network at V8. This unique local structure may contribute to immobilization of 7R1, positioning the NHB residue K37 in close contact. 8R1, in contrast, is more mobile than

would be expected for a partially buried nitroxide[39] , suggesting that the local structure of strand A at the kink is dynamic. In the second exception, strand D makes up the first half of the NHB cleft, which is edge-strand like. In the FABP-like protein cellular retinol binding protein II (CRBP II, an RBP), similar simple anisotropic line shapes were found in the equivalent strand 4.[38] The motion of nitroxides in the CRBP II cleft were attributed to restriction of R1 rotamers to one that positions the nitroxide in an ordered configuration by aligning the disulfide bond towards the open cleft space.

The most immobilized spectra in this survey are 7R1 and 105R1. Higher local side chain packing density is the most likely origin for these unique line shapes. Similarly to 7R1, 105R1 is near a charged side chain (E107). Although this neighbor is the *i*+2 intrastrand neighbor and not the NHB interstrand neighbor, E107 forms a surface salt bridge with K88, which could be responsible for the higher packing density in the vicinity of 107R1. The R1 motion implied by the sequence-dependent line shapes are therefore compatible with crystal structures.

4.2.2 Results: The low pH state of I-FABP is observed in new spectral components

To assess changes in side chain packing in the β-sheet, and hence local structural changes, CW EPR spectra of nitroxides placed on the solvent-facing surface of apo I-FABP were obtained at pH 2.8 and 2.2. The conformational states populated at each pH include the native state and the MG state, which is partially populated at pH 2.8 and is the predominant state at pH 2.2.[29] Although the line shapes in I-FABP in the native state are complex, as seen by multicomponent spectra, the transition to the MG state is spectrally straightforward: a highly mobile new component appears (Fig. 4.3, purple and orange lines).

In some cases, the appearance of the new mobile component is accompanied by

immobilization of the component observed in the native state. This additional change in the line shape could be due to a change in side chain packing of the native state (i.e. protonation of neighboring solvent-exposed carboxylate side chains) and not population of the MG, which could be evaluated by comparison with spectra at pH ~4 , near the pKa of carboxylate side chains. Alternatively, the MG state could be multicomponent, resulting in the appearance of both mobile and immobile components.

In contrast to the partial unfolding at pH 2.2 indicated by other spectroscopies[23,29], the line shapes at pH 2.2 in strands C, D, G, H, and I are unlike the very mobile spectra anticipated for highly disordered or unfolded conformations.

4.2.3 Discussion: SDSL EPR at variable pH of the β-domain

In the native state, SDSL EPR spectra on the solvent-facing surface of the β-barrel domain of I-FABP are highly variable as expected.[38] In many cases, the relative mobility of nitroxides within strands can be rationalized by identifying bulky or β-branched NHB neighbors, which have been shown to decrease the mobility of R1. Although it is not possible to identify backbone motions on the ps-ns time scale due to the prevailing effects of local side chain contacts, the sensitivity of R1 motion to tertiary structure and packing density is particularly useful in comparing the native state to the acid-stabilized MG in the β-barrel of I-FABP.

Identification of backbone motion is much more straightforward in spectra of R1 in a solvent-exposed α-helix. For example, it has been shown that the line shape of a nitroxide in a rigid helix is generally dominated by a single component that reflects an organized, single dynamic state[40] and that spectra of an acid-stabilized MG in an α-helix are often multicomponent, sometimes with a sharp component that describes a conformational state with

high backbone mobility.[41] To link backbone motion to packing density in the native state of holo Mb, the fraction of buried surface area in the vicinity of the nitroxide, calculated from the X-ray crystal structure, was found to correlate inversely with nitroxide mobility.[35]

In the β-barrel of I-FABP, features of the tertiary structure revealed by R1 R1 is limited to side chain packing on the solvent-facing and cavity-facing surfaces of the protein, rather than contacts between separate secondary structure elements, such as separate α-helices in the globular structure of Mb. Therefore, the MG state of I-FABP is a unique structure unlike other MG conformations studied by SDSL EPR, including the iLBP tear lipocalin, which has tertiary packing between its major α-helix and the solvent-facing β-sheet that is partially released on populating the MG at pH 3.0.[42]

At acid pH, SDSL EPR of I-FABP is able to distinguish the population of a new conformational state in all strands studied herein (strands A, C, D, F, G, H, and I). The appearance of a new mobile component suggests loss of stabilizing interactions between R1 and native side chains, consistent with the definition of the MG (i.e. reduction in tertiary interactions). With decreasing pH, a small gradient in the population of the new highly mobile component appears along strands in the direction away from the helix-turn-helix motif, irrespective of the strand N- to C-terminal direction. This gradient is the most clear in strand I (Fig 4.3), for which 118R1 at the top of the figure is proximal to the helical domain and 114R1 at the bottom of the figure is near the base of the cavity. Although the populations of the new mobile components are very small, the transition is more apparent near the hydrophobic cluster at the base of the cavity that forms the I-FABP core. Because the MG is expected to have a hydrated cavity, this gradient may be indicative of allosteric remodeling in the new conformational state.

This gradient is not apparent in "noncanonical" antiparallel strands A and D, which do not have the regular backbone hydrogen bonding network that defines an antiparallel β-sheet. Strand A has a unique kink that rearranges the hydrogen bonding strand pairs from the N-terminal half of strand A with strand B to the C-terminal half of strand A with strand J. Strand D has a unique structure, as well: its adjacent strand E is not hydrogen bonded, but instead a trio of cavity-facing π-stacking interactions stabilize the two strands. These exceptions demonstrate the high sensitivity of R1 motion to secondary structure regularity in the MG state of I-FABP revealed by strand scanning SDSL.

4.3 Results: SDSL EPR of the native and acid states of I-FABP α-helix I

In contrast with the subtle structural changes in the β-domain, the portal region of apo I-FABP is thought to become more flexible in the "rare" open conformation, involving outward movement of the C-D and E-F loops and local unfolding of the C-terminal half of helix II (Fig. 4.1b). The low pH state is thought to be involved in this structural reorganization as well. To evaluate whether helix II in apo I-FABP has persistent secondary structure at low pH, secondary structure determination by SDSL EPR was carried out by spin labeling consecutive sites between residues 25 – 30 (Fig. 4.4). The mobility of consecutive nitroxides scanning a helix varies with the period of an α-helix (i.e. every 3.6 amino acids). Because nitroxide line shapes are sensitive to local contacts, spin labeling consecutive sites produces a mobility gradient between nitroxide side chains which are oriented towards solvent and towards the protein, where the mobility is higher for noninteracting, solvent-facing sites. The inverse central linewidth, measured as the peak-to-peak distance in the center line (ΔH_{pp}^{-1}), is a semiquantitative measure of nitroxide

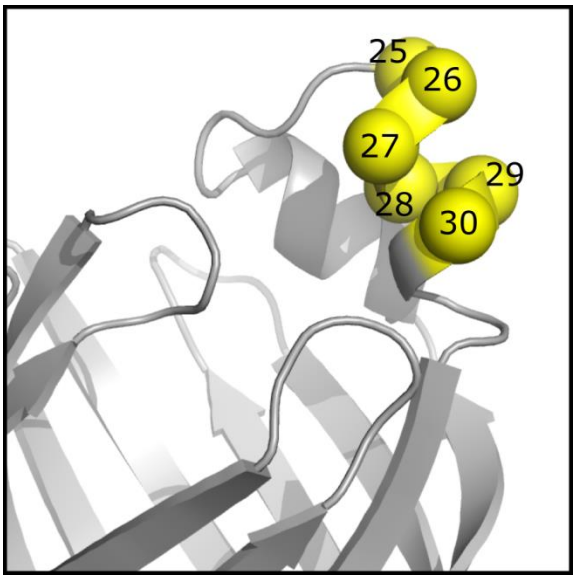

Figure 4.4 Sites in the α-helix II of apo I-FABP subject to SDSL EPR. The orientation of this figure is the same as that in Fig. 4.1b, showing the C-D loop below and the E-F loop to the left.

mobility and is proportional to $1/\tau$. ΔH_{pp}^{-1} is used here to probe sequence-specific changes in the mobility of R1.

The semiquantitative parameter ΔH_{pp}^{-1} for nitroxides in helix II of apo I-FABP shows the expected periodicity at pH 7.3, shown in Fig. 4.5. The phase of the turn matches the expected orientation from the crystal structures. At pH 2.8 and 2.2, the periodicity remains, indicating that the helix does not unfold. The overall upward shift in the values of the inverse central line width result from population of the mobile components in each spectrum and are indicative of a more flexible helix. In the derivative form of the absorbance spectrum, the amplitudes from the mobile components over-represent the populations. Despite this bias towards the sharp mobile component, the maintenance of secondary structure suggests that the mobile components do not reflect unfolding but instead indicate a highly dynamic state in exchange with the immobile state observed at neutral pH. The mobile component that appears at low pH resemble the transitions observed in the β-domain, suggesting that the MG conformation at acid pH is characterized by multicomponent spectra with a population of a highly mobile nitroxide, irrespective of secondary structure type.

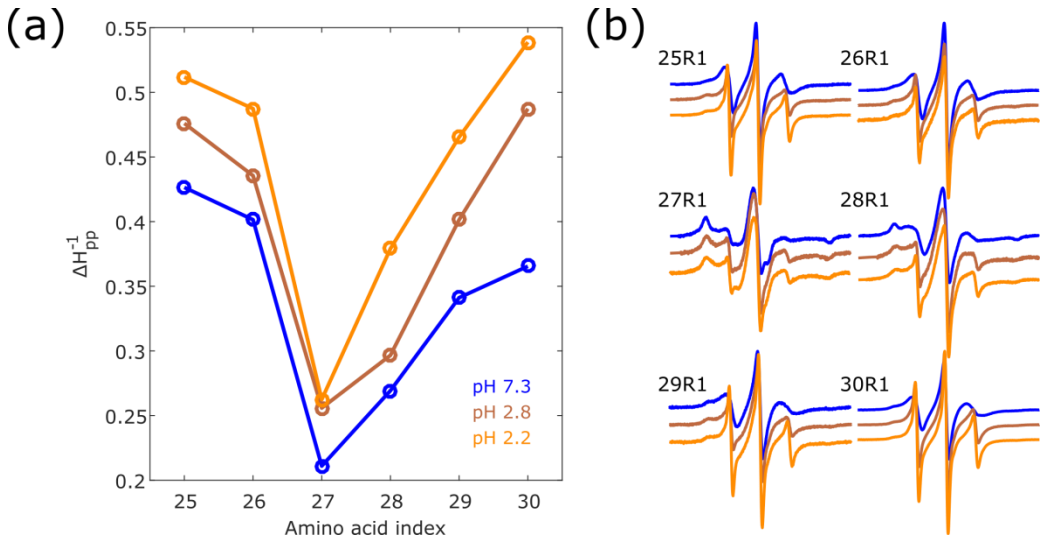

Figure 4.5 In (a), the inverse central line width of nitroxides placed in helix II of apo I-FABP demonstrates that the secondary structure is maintained, even at acid pH. In (b), the CW EPR spectra at each site change with decreasing pH.

In order to evaluate whether the multicomponent spectra arise from R1 rotameric or conformational exchange, saturation recovery (SR) EPR was performed, previously discussed in Chapter 2. SR measures the spin-lattice or T_1 relaxation time of nitroxides, typically in the range of 1-10 μs. In addition, SR can determine the lifetime of exchange processes that occur between two states of R1 that have different T_1s. The T_1 of the nitroxide is sensitive to nitroxide mobility: highly mobile nitroxides have CW spectra which are sharp and have T_1 relaxation ≤ 3 μs, while local contacts produce broad, immobilized CW spectra and slow T_1 relaxation ≥ 7 μs.[43] Thus, SR provides a strategy to estimate exchange lifetimes between states that give rise to mobile and immobile components in a multicomponent CW line shape. Multicomponent spectra from nitroxides that arise from equilibrium between two conformational states have slow exchange rates and generate bi-exponential T_1 decay curves. On the other hand, rotameric exchange is fast on the SR time scale of μs, and multicomponent spectra that arise from rotameric exchange produce single exponential T_1 decay curves due to exchange averaging. In this work, the

uniform penalty (UPEN) method[44] is used to fit the exponential data by taking advantage of a library of exponentials to generate distributions of T_1 relaxation times without the need to determine a model for the required number of exponential functions.[45]

The UPEN fits and resulting distributions for nitroxides in helix II of apo I-FABP show that both in the native state and at acid pH, the multicomponent line shapes reflect exchange between states that are at least in μs exchange, if not slower (Figs. 4.6 and 4.7).

4.3.1 Discussion: Conformational exchange of α-helix II

The populations of these T_1 states, calculated from the area under each distribution peak, do not necessarily reflect the populations of the CW spectral components. In the case of intermediate exchange on the SR time scale the areas under the peaks are functions of both exchange rate and true populations ([46] and Michael Bridges, unpublished work). A framework for estimating exchange rates in the intermediate exchange regime has been developed[46]; the method exploits the differential accessibility of the nitroxides to collision with a paramagnetic exchange reagent to modulate the relative apparent values of T_1. Future investigation into the exchange frequency between these states in helix II of I-FABP can be approached only having an exchange reagent that is stable at low pH. Unfortunately, the commonly used NiEDDA complex dissociates at pH 2.2. As a prelude to such studies, chromium (III) maltolate ($Cr(malt)_3$) was found to be suitable because it is uncharged and remains complexed at pH 2.2 (Fig. 4.8). This sets the stage for future studies, but is beyond the scope of the thesis project.

The slow (> μs) exchange of helix II at pH 7.3 implied by the resolution of two relaxation times is in good agreement with the ~1 ms predicted timescale for protein conformational motion related to ligand binding[22], which was modeled by Cistola *et al.* as a rate-limiting step involving

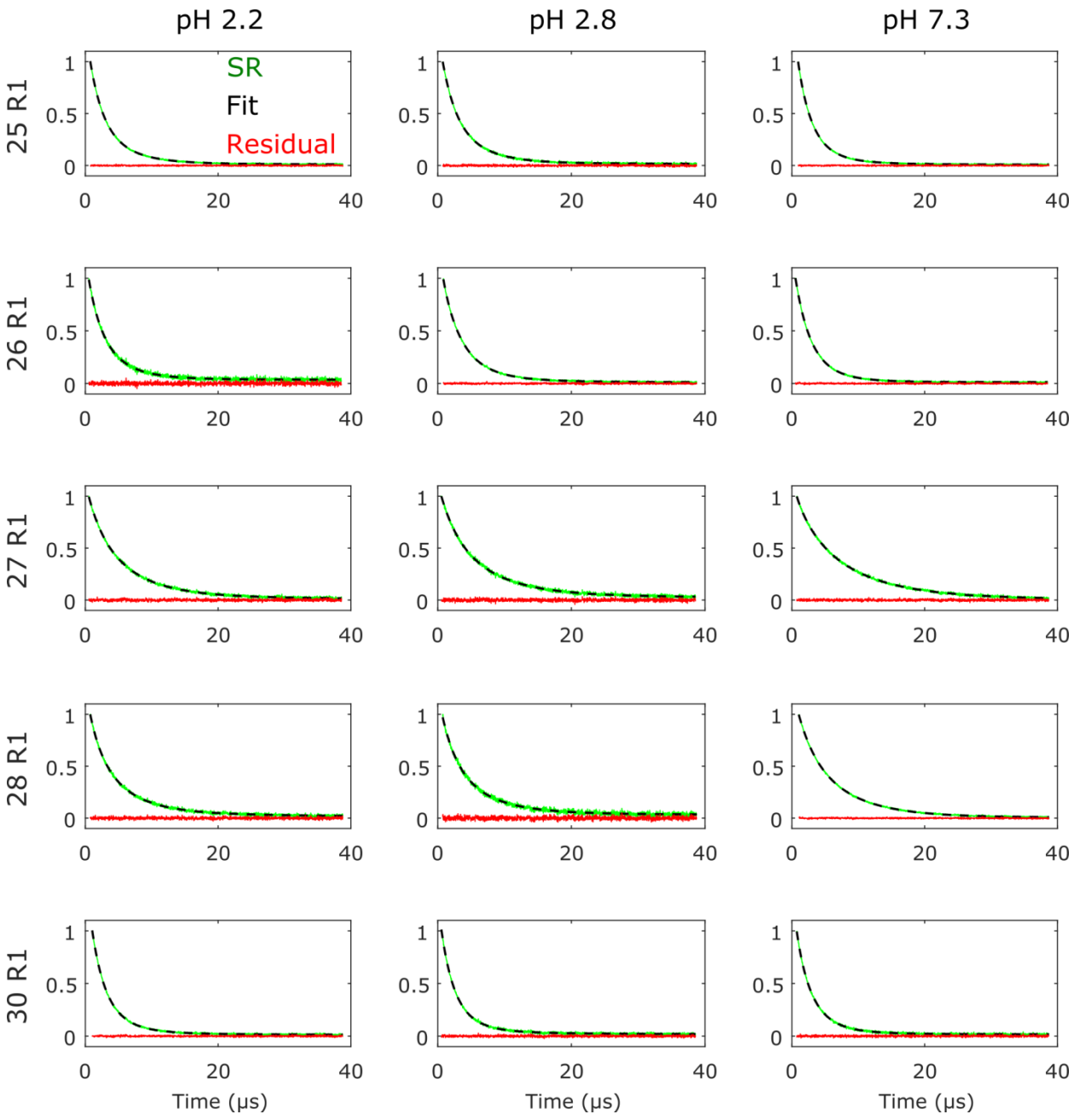

Figure 4.6 Normalized saturation recovery data for sites on helix II of apo I-FABP under nitrogen atmosphere and corresponding UPEN fits.

conformational rearrangement of the portal. This transition to the "rare" open-portal state is expected to occur in between the SR and NMR timescales[13] at neutral pH.

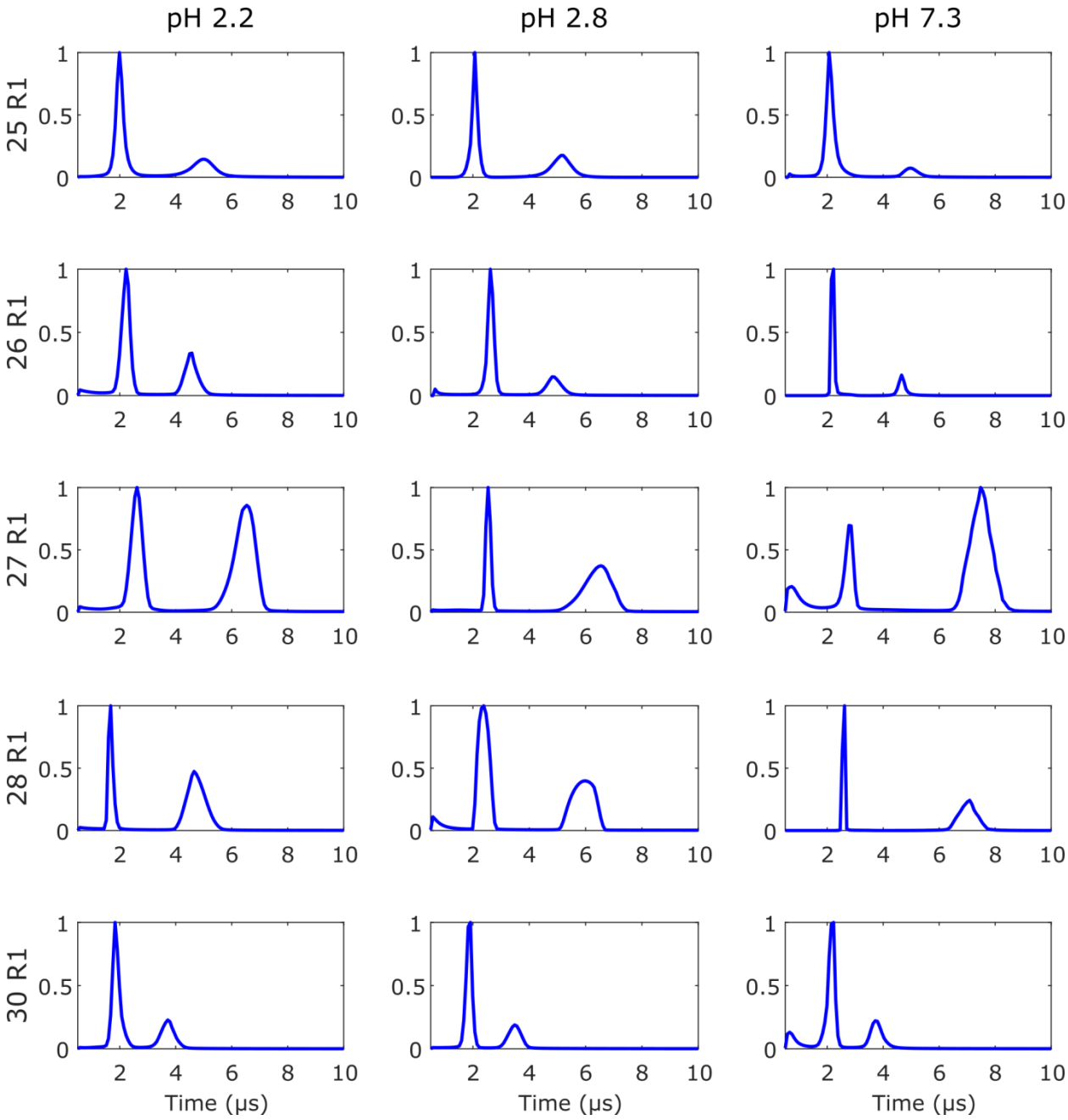

Figure 4.7 T_1 distributions resulting from the UPEN fits to data in Fig. 4.8. Every site in helix II demonstrates two T_1s, which are weighted by the fractional area under each peak. Lack of correlation between the fractional population of each T_1 and the fractional population of CW spectral components suggests intermediate exchange on the SR timescale at all pH values. Note that while the distribution widths of short T_1s are reliable, the distribution widths of long T_1s are less so (Michael Bridges, unpublished work).

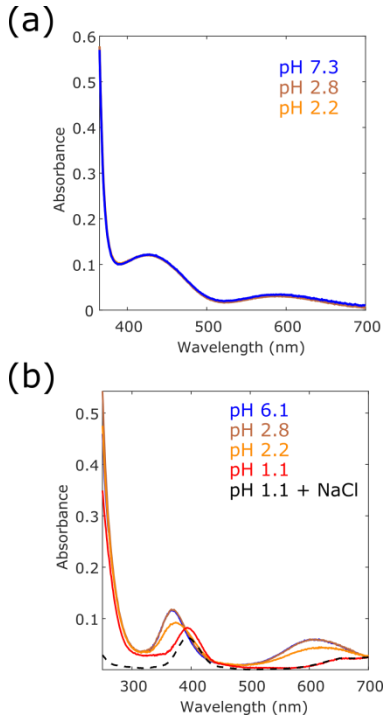

Figure 4.8 UV-visible spectroscopy of (a) 4.8 mM $Cr(malt)_3$ and (b) 17mM NiEDDA in low ionic strength buffers at varying pH conditions indicate the pH dependence on metal ion coordination.

4.4 Results: Electrostatic interactions play a role in stabilizing the α-helical domain

The helical domain of I-FABP is apparently sensitive to pH and is destabilized as pH decreases from pH 4 to pH 2.[29] As a possible mechanism of action, disruption of a salt bridge could free stabilizing electrostatic contacts in the helical domain in a pH-dependent manner. The conformational transition occurs at a pH close to 2, well below the pKa ≈ 4 of a solvent-exposed carboxylate (such as an aspartic or glutamic acid side chain). However, the pKa of a carboxylate in a salt bridge is shifted strongly downward and disruption of a stable salt bridge is a potential candidate for the trigger which leads to the low pH conformational change.

Candidate salt bridges were identified by comparing the crystal structures of I-FABP. A salt bridge network was found to be highly conserved among FABPs that bridges both helix I and helix II (Fig. 4.9). The network connects of K16 – E19 – R28 – E15, from which E15 and K16 are in helix I and E19 and R28 are in helix II. Three salt bridge-breaking mutations were introduced into the helical region of I-FABP in order to investigate specific charge-charge

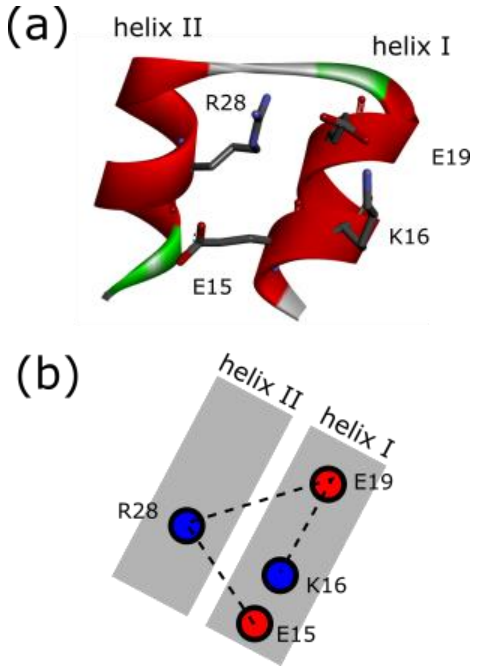

Figure 4.9 Locations of the interhelical salt bridge contacts in I-FABP. In (a), the side chains from the crystal structure 1ICM are shown. In (b), the connectivity of the side chains is shown schematically.

interactions that may be stabilizing. Conservative mutations minimally disrupt the helices were chosen rather than alanine substitutions in order to preserve polarity of these short helices.

The mutation E15Q is intended to break the salt bridge between helices I and II proximal to the C-terminus of helix II, which is the region implicated by NMR as important in opening the portal[13]. The intrahelical salt bridge E19 – R28 remains intact. This mutation is expected to destabilize both helices partially even though the interaction between E19 – R28 may still compensate for the loss of E15 – R28. A E19Q mutation leaves the first salt bridge intact. Lastly, the charge on R28 was mutated, which participates in both salt bridges which span helices I and II. R28I was chosen instead of a charge reversal as isoleucine is found in other FABP isoforms and could participate in an alternative arrangement of the helices.

pH-matched CW spectra of wild-type I-FABP and the three salt bridge mutants were acquired by spin labeling at reference sites in each helix: 20R1 for helix I and 30R1 for helix II. At neutral pH, the line shapes for all mutants are nearly indistinguishable. This confirms that the amino acid substitutions were indeed conservative, even for amino acid mutations that are

adjacent to the spin labeled sites, but also indicates that a new state is not populated at pH 7.3 (Fig. 4.9a); thus breaking these salt bridges does not populate the MG. As the pH is decreased, the line shapes differ more from the wild-type reference spectra. At pH 3.6 (Fig. 4.10b), both the spectra for 20R1/E19Q and 30R1/E19Q are more mobile than wild-type or the other substitutions, clearly visible by the increased center line height. All mutants demonstrate a similar increase in the population of the immobile component of 20R1, however. At pH 2.8, the trend continues as 20R1/E19Q and 30R1/E19Q is more mobile than 20R1/WT and 30R1/WT. However, the other mutations show some dissimilarity: 20R1/E15Q is more mobile than 20R1/WT and 20R1/R28I is less mobile than 20R1/WT. 30R1/E15Q and 30R1/WT are indistinguishable at both pH 2.8 and 2.2 (Fig. 4.10c and 4.10d), but at pH 2.2, 20R1/E15Q is slightly more mobile than 20R1/WT. The overall effect of E19Q and E15Q from the perspective of 20R1 in helix I and 30R1 in helix II is a small destabilization, visible as populating the mobile component more than is found in the wild-type, with E19Q slightly more so than E15Q. R28I

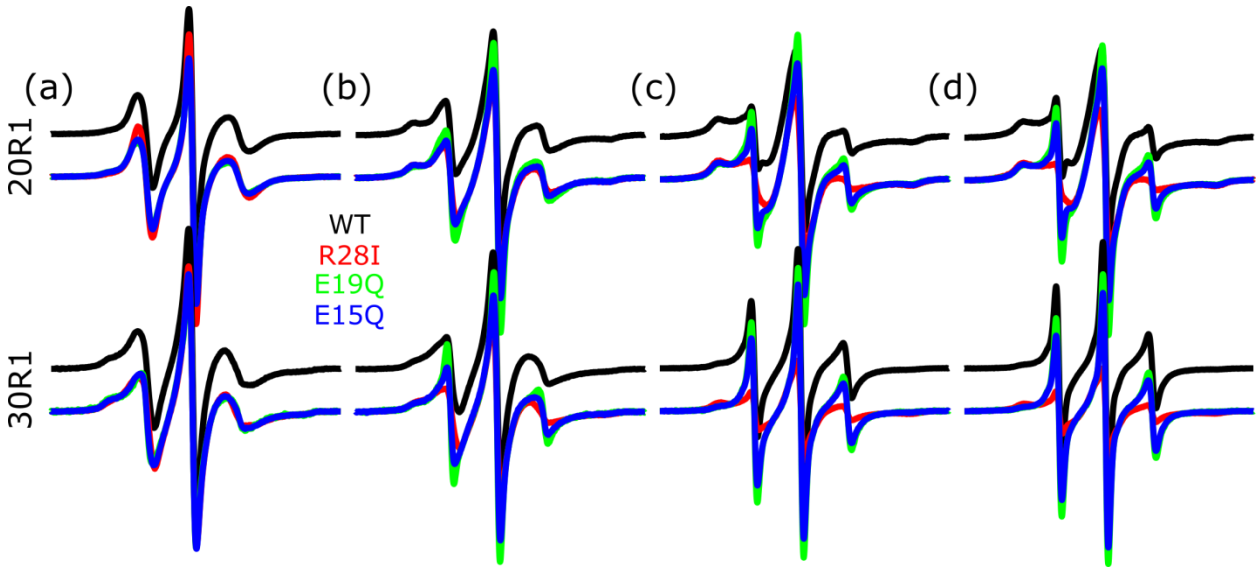

Figure 4.10 CW spectra for charge-reversal mutations in putative salt bridges within helices I and II of apo I-FABP. In (a), at pH 7.3, (b) pH 3.6, (c) pH 2.8, and (d) pH 2.2, no single mutation is shown to pre-populate the line shape anticipated by the wild-type protein at pH 2.2 (black, (d)) at neutral pH (colors, (a)).

produces a stabilizing effect on both helices compared to wild-type. It is not likely that these salt bridges contain the pKa ~2 moiety because they do not show an effect at neutral pH. However, they do participate in maintenance of the wild-type-like conformation at low pH and therefore play a minor role in the MG transition of I-FABP at pH 2.

Because the helices do not have a pH-sensitive intrinsic switch, the location of the salt bridge (or multiple bridges) may be in the β-domain of I-FABP. An initial survey of salt bridge pairs in the interior of the β-barrel show two possible networks: first, between D34 and R126, and second, a more complex network involving E51, R56, and Y70 along with R56 interacting with the backbone of S71. Following from the approach to survey salt bridges in the helical domain, D34 and E51 are acidic amino acids that have pKas that could be lowered if they are form strong salt bridges.

Rather than only assessing the conformational changes in the helical domain, global structure of I-FABP was characterized by dynamic light scattering (DLS) of the protein. At neutral pH, I-FABP is spherically globular. The MG state at low pH has a larger hydrodynamic radius[29] which can be detected efficiently by DLS. DLS indicates the wild-type protein expands slightly at pH 2.8, at the start of the expected structural transition, but significantly increases in hydrodynamic radius at pH 2.3.

The D34A mutation was introduced to abolish the first β-domain salt bridge. Alanine was chosen rather than glutamine because the β-domain interior is sufficiently hydrophobic to accommodate removing the charge. In comparison to wild-type, D34A has the same pH-dependent behavior even at pH 7.3, indicating that protonation of this residue is not the sole driving force behind the structural change at low pH (Fig. 4.11a and 4.11b).

Fitting the DLS size distributions to a single Gaussian provides the average radius of each

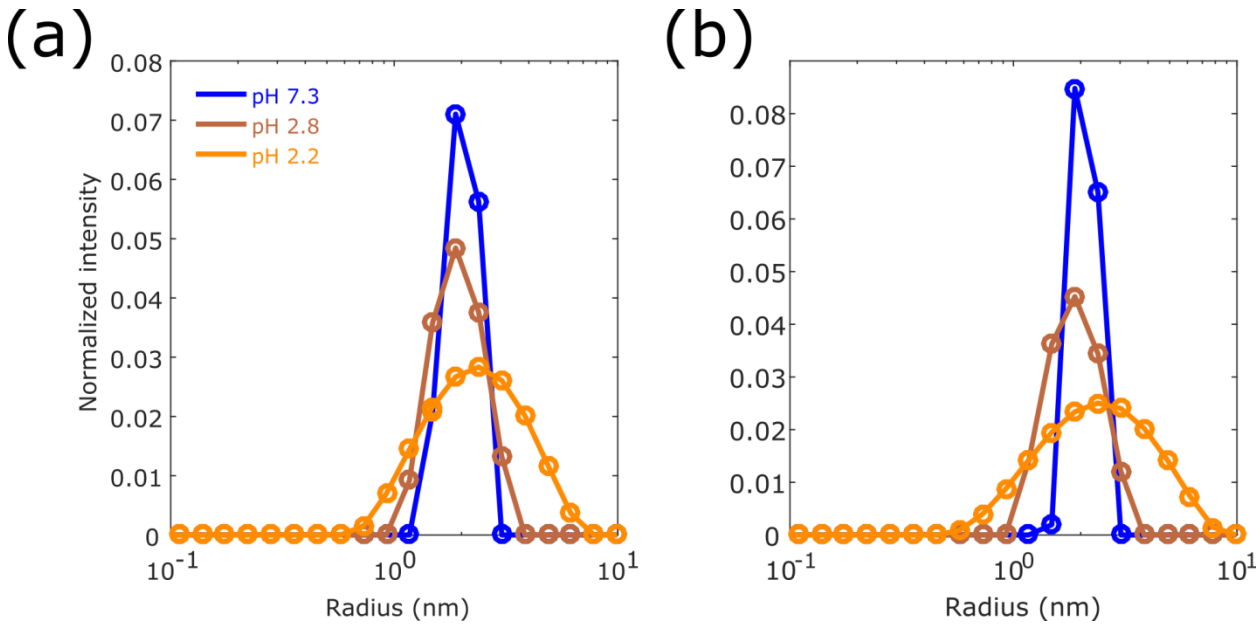

Figure 4.11 Dynamic light scattering profiles show the hydrodynamic radius of I-FABP increases as pH decreases. In (a), wild-type apo I-FABP swells as the putative MG state is populated. In (b), the 34A mutation does not cause a conformational shift at neutral pH; instead, this mutant behaves in a manner identical to the wild-type protein.

particle. The size of I-FABP at acid pH determined here (see Table 4.1) is in striking agreement with the hydrodynamic radius (r_H) determined by pulsed field gradient diffusion NMR.[29]

Table 4.1 The hydrodynamic radius of apo I-FABP at variable pH values

pH	WT r_H	95% CI	D34A r_H	95% CI	WT r_H**
7.3	20.3 Å *	20.1 – 20.5 Å	20.9 Å *	20.9 – 20.9 Å	15.5 Å
2.8	19.1 Å	18.9 – 19.3 Å	18.6 Å	18.4 – 18.8 Å	17.2 Å
2.2	23.0 Å	22.5 – 23.6 Å	23.7 Å	23.1 – 24.3 Å	23.7 Å

* Fits are only to a small number of points and have higher errors than anticipated by the 95% confidence intervals.
** Values for the hydrodynamic radius from Li and Frieden.[29]

A second candidate in the β-domain, E51A, breaks a native salt bridge connecting β-strands C (E51), D (R56), and E (Y70). This region is notable because the C-D loop contains F55, which could be involved in the MG state because it reorients in the presence of ligand and

is thought to serve as a lid to the portal. [11] Hydrogen bonding of the adjacent residue R56 with Y70 and S71 would also contribute to the local electronic field in the portal, providing an additional effect for depressing the observed structural change at around pH 2.5. Delipidation with Lipidex-1000 at 37 °C resulted in complete precipitation of the mutant. This indirectly implies that E51 is important in stabilizing the I-FABP structure. A previous study investigating the E51A mutant also described instability of the protein during expression, but were able to successfully delipidate I-FABP using this method in high ionic strength buffer.[47] This mutant is promising in that its stability is greatly reduced in low ionic strength conditions compared to other mutations investigated so far, including a number of other cysteine substitutions within the β-domain. Future work characterizing E51A or a charge-reversing E51Q mutant could be accomplished through immobilization on a solid support before delipidation, though characterization using optical spectroscopies like DLS would not be possible.

4.4.1 Discussion: The acid MG of apo I-FABP is not populated by mutations in particular salt bridges

The results presented do not implicate breaking any single salt bridge can populate the apo I-FABP MG at pH 7.3. A more exhaustive search could identify a single or multiple salt bridges that are broken at pH 2, perhaps by taking advantage of recent advancements in calculating residue pKas from X-ray crystal structures which recognize the importance of conformational exchange in predicting pKa values[48]. Additionally, search tools using a database of experimental data could facilitate identification of unique salt bridges in I-FABP.[49]

4.5 Results: Structure of the acid MG of apo I-FABP by DEER EPR

The increase in hydrodynamic radius of a protein upon formation of the MG (see DLS distributions in the previous section) is a hallmark of the MG; however, analysis of DLS is limited to particle shapes that are fixed shapes (i.e. light scattering data of apo I-FABP are fit to a spherical particle model). Swelling of the protein is expected to involve movement of the protein backbone and an increase in the size of the associated hydration shell. Hydration of a protein can be described by disengaging tertiary contacts across the protein associated with hydration of charged sidechains, recruiting water to the structure. As changes in the structure of I-FABP appear heterogenous, suggested by differences in the magnitude of line shape changes at different sites, inter-nitroxide distances were measured to determine the amplitude of changes corresponding to population of the MG at acid pH. I-FABP contains two somewhat independent domains: a large β barrel and a small helix-turn-helix domain. For DEER distance mapping of the structural changes in I-FABP with pH, a set of doubly-labeled mutants were designed to monitor both the helical and beta domains. Solvent-facing sites were modified with the R1 spin label in strands A, C, F, and I, as well as helices I and II. These pairwise mutations measure the width of the β-domain, as well as the position of the helices relative to the clamshell and relative to each other, illustrated in Fig. 4.12.

To consider which sites serve as good reference sites, it is important that the position of the nitroxide and the local protein structure changes minimally with pH. The CW line shape is a highly sensitive measure of local dynamics and serves as a guide. Because the line shape of 81R1 is mostly pH-independent, this site in the β-domain serves as the most stationary reference site for this set of DEER distances (Fig. 4.13). The remaining sites all have multicomponent

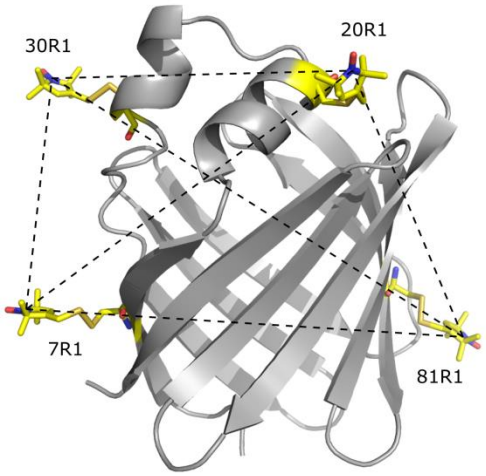

Figure 4.12 Nitroxide sites used to map the structure of apo I-FABP at varying pH conditions.

spectra, which may give rise to additional peaks in the DEER distance distributions from each rotamer. Despite the complex distribution shapes that arise from both internal nitroxide rotamerization and protein conformational heterogeneity, the aim of this section is to compare structural changes at varying pH conditions. Since the nitroxide spin label is aprotic and not intrinsically pH-sensitive, the line shape differences are due to exchange with non-native conformational states, which will be reflected in the DEER distance distributions. Seven pairs of nitroxides were studied to provide a global map of the pH-dependent conformations of apo I-FABP. 7R1/20R1 and 20R1/81R1 monitor helix I, 7R1/30R1 and 30R1/81R1 monitor helix II, 20R1/30R1 monitors the distance between the helices, and 48R1/116R1 and 7R1/81R1 monitor the beta domain.

A measure of the conformation of the β-domain is monitored by 7R1/81R1. At neutral pH, the distance of 7R1/81R1 has a half-maximal width of 3.8 Å, which supports a model for the β-domain as a rigid barrel structure. At pH 2.8 and 2.2, new distances both shorter and longer than the maximal peak reflecting the native state appear in the distribution of 7R1/81R1 (Fig. 4.13, second row). Because the CW line shape of 7R1 is significantly mobile at low pH, these distances more likely reflect disordering of strand A (containing 7R1) rather than strand F

(containing 81R1). The role of strand A in stabilizing the I-FABP fold may be minor, as evidenced by the native-like structure and ligand-binding behavior of the reductive Δ98Δ mutant, which lacks strand A, helix I, and helix II[50]. By estimating the population of the residual native state by fitting the 36 Å-centered peak to a single Gaussian, the low-amplitude distances represent a new heterogenous state that amounts to a sizeable 36% of the total I-FABP population. These broad distances can be distinguished from an unfolded state by comparison with the 7R1/81R1 distance in 8M urea (Fig. 4.14). Despite this overall broadening at low pH, the average distance between 7R1/81R1 remains constant from pH 7.3 – 2.2, suggesting that the fraction of native-like structure is in conformational equilibrium with the MG-like state.

The orientation of the α-helical domain with respect to the β-domain domain is of particular interest for identifying conformational motions related to opening of the putative portal region. Sites for spin labeling were chosen from poorly conserved amino acids in each α-helix, which are expected to minimally perturb the helices: 20R1 and 30R1 were chosen to monitor the positions of helix I and II, respectively.

At neutral pH, the distance distribution of 20R1/81R1 between helix I and strand F is slightly more broad than the distribution width of 7R1/81R1 with a half-maximal width of 6.4 Å (Fig. 4.13, first row). At pH 2.8, however, this distribution shifts to shorter distances, seen as a bimodal distribution centered around peaks at 24.0 Å and ~28 Å. Fitting this distribution to two Gaussians shows that the major peaks comprise 18% and 82% of the total I-FABP population for the short and long peaks, respectively. Decreasing the pH to 2.2 continues to depopulate distances corresponding to the native state at represented by the peak centered near 28-30 Å while increasing the population of the shorter distances centered around 24.0 Å. The population of distances in the short distance regime (fitted to Gaussians) increases to 37%. A

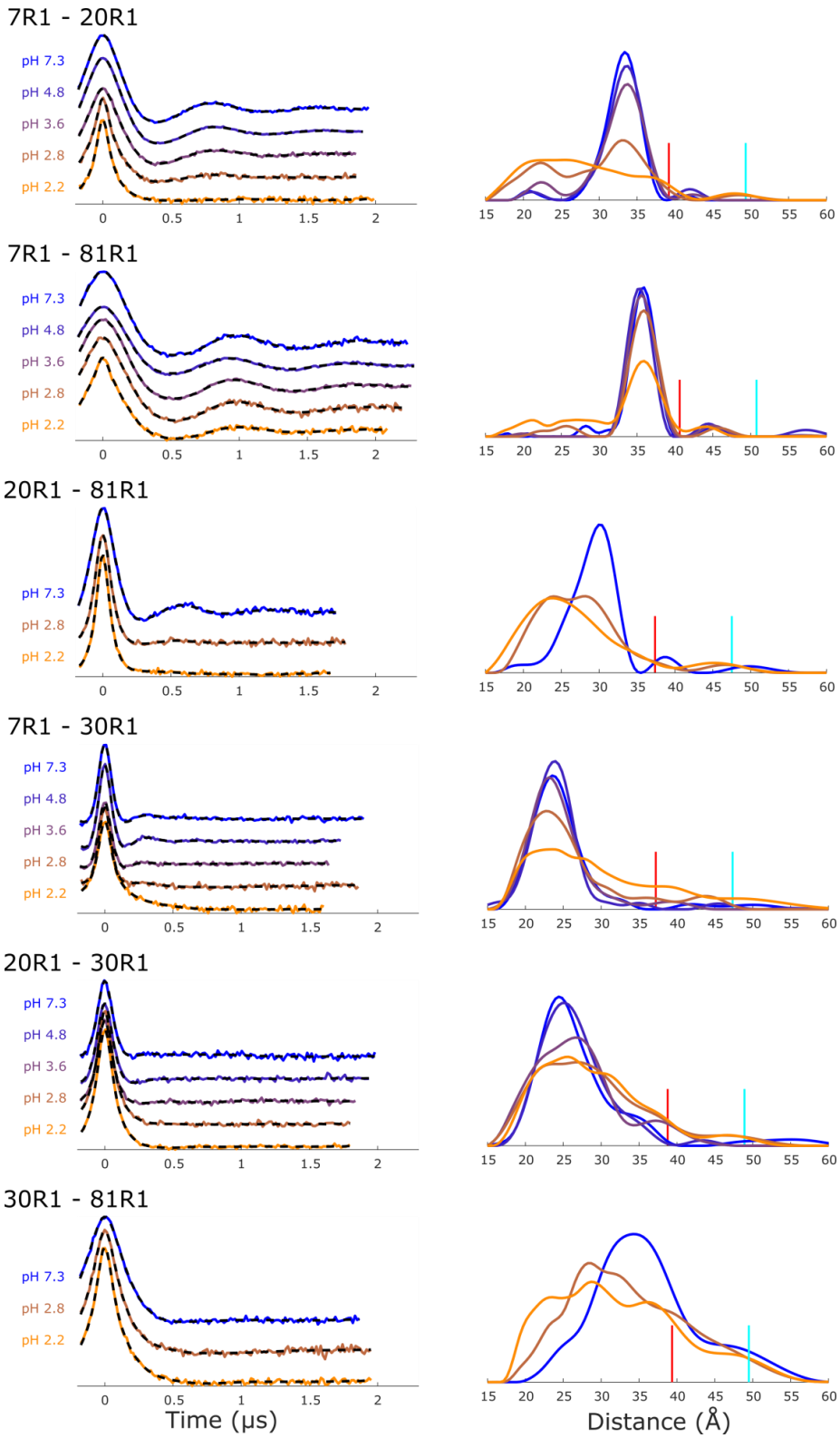

Figure 4.13 DEER dipolar evolutions, left, and fitted distance distributions, right, for apo I-FABP at varying pH conditions. The red bar shows the confidence limit for distribution shape and the cyan bar shows the confidence limit for distribution length.

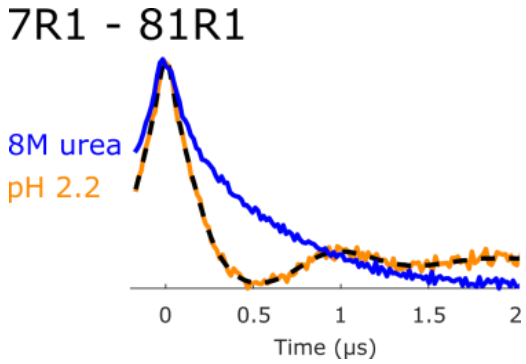

Figure 4.14 DEER dipolar evolutions for the cross-cavity double 7R1/81R1 at acid pH and in 8M urea, demonstrating the lack of structure in chemical denaturant by a broad and smooth DEF compared to the maintained protein structure indicated by oscillations at pH 2.2.

rough estimation of the population of the mobile component in the CW spectrum of 20R1 at pH 2.2 is only ~12%, suggesting that the distances near 24 Å reflect an acid-stabilized conformation of apo I-FABP and are not only a result of a difference in nitroxide rotamers.

In contrast, the DEER pair monitoring helix II, 30R1/81R1, is much broader at neutral pH compared to distances within the β-domain and involving helix I (Fig. 4.13, sixth row). Because the lineshape of 30R1 is multicomponent, the 12 Å distribution width could be due to the rotameric space of R1 at this site, but the literature strongly suggests the C-terminal half of helix II is less ordered even at neutral pH.[13] At pH 2.8 and at 2.2, the distances between 30R1/81R1 shift to shorter distances while depopulating the native state in a manner similar to helix I, but due to the broadness of this distribution, fitting to a sum of Gaussians is less descriptive.

The inter-helical distance between 20R1/30R1 at neutral pH is a sharp distribution centered at 24 Å, with a half-maximal distribution width of 7.3 Å (Fig. 4.13, fifth row). Both the distribution width and the shoulder at longer distances reflect conformational flexibility within the helical domain. With decreasing pH, the native state peak is depleted from the distribution as longer distances are increasingly populated. Because this distribution is sharper than 30R1/81R1 at pH 2.2 (14.9 Å < 20.9 Å), it is possible that helix I and II have concerted motions.

These same helical reference sites (20R1 and 30R1) were paired with 7R1 in strand A to provide a more complete description of the helical conformations. Although the line shape of 7R1 changes significantly with decreasing pH, the magnitude of change in the line shape is not indicative of the amplitude of change in the DEER distance distributions, though the new spectral component does suggest that the orientation of the nitroxide with the protein changes with decreasing pH. Using 7R1 as a reference site therefore provides additional comparisons with which large shifts in the conformation of I-FABP can be identified. The distance distribution between strand A and helix I (7R1/20R1, Fig. 4.13, first row) and strand A to helix II (7R1/30R1, Fig. 4.13, fourth row) are relatively sharp at neutral pH and broaden significantly at pH 2.2. The width at half-maximum of the distribution between 7R1/20R1 is 20.1 Å, and between 7R1/30R1 is 13.2 Å, indicating that both strand A and helix I could be very flexible at low pH.

Together, the CW lineshape and DEER data suggest that the structure of I-FABP becomes conformationally heterogeneous on the EPR times scale with exchange lifetimes longer than about 100 ns.

4.5.1 Discussion: The conformational change in apo I-FABP at low pH populates a highly heterogeneous state

In contrast to descriptions of the acid MG of apo I-FABP as an "acid unfolded state"[29] and an "intermediate state where secondary structure is relatively intact"[51], distance distributions between pairs of nitroxides show that a fraction of native-like structure is maintained even at pH 2.2 and the new conformational state (or states) are highly heterogeneous but not unfolded. DEER distribution widths with nitroxides placed in the helical domain are broader than the

reference pair in the β-domain at neutral pH and at low pH, in support of a set of pH-populated conformational states modeled by a pair of dynamic helices and an ordered β-domain. Table 4.2 summarizes the differences between the seven DEER distributions at pH 7.3 and pH 2.2.

Table 4.2 Quantification of DEER distance distributions

	pH	7R1/81R1	20R1/81R1	30R1/81R1	20R1/30R1	7R1/20R1	7R1/30R1
Half-maximal width	7.3	3.8 Å	6.4 Å	12.0 Å	7.7 Å	5.1 Å	6.8 Å
	2.8	4.3 Å	12.4 Å	17.5 Å	14.9 Å	16.2 Å	8.5 Å
	2.2	5.1 Å	12.4 Å	20.9 Å	14.9 Å	20.1 Å	13.2 Å
Most probable distance	7.3	35.9 Å	29.9 Å	34.2 Å	24.4 Å	33.4 Å	23.5 Å
	2.8	35.9 Å	28.2 Å	28.2 Å	27.0 Å	32.9 Å	22.7 Å
	2.2	35.9 Å	24.0 Å	28.7 Å	25.7 Å	25.3 Å	23.5 Å

In summary, α-helix I shifts closer to strand A and strand F in the MG, while helix II moves away from strand A and closer to strand F. The interhelical distances do not suggest that an "open" portal is a conformation in which the helices splay apart; instead, the helices fluctuate over a wide distance range, forming a number of arrangements but not a static "open" state. The β-domain retains a large fraction of its rigid structure at pH 7.3 even at pH 2.2, supporting an NMR study which interpreted ligand binding at low pH as due to residual population of the native state.[29]

4.5.2 Results: An alternative portal shows high conformational heterogeneity at acid pH

Recent simulations in the myelin isoform of FABP suggest that the non-H-bonded cleft plays an important role in exposing the ligand to solvent. Atomistic molecular dynamics simulations identified dissociation of strands D and E by up to 20 Å at neutral pH and proposed this alternative portal as a mechanism for ligand entry and exit[18]. In this model, F57 in the C-D loop (the equivalent of F55 in I-FABP) was found to adopt different positions in different

conformations. In earlier studies, F57 in adipocyte FABP was implicated as part of the portal region of all FABPs.[52] Not only is the position of this phenylalanine conserved across FABPs, rotation of the phenylalanine side chain is found in the apo to holoprotein transitions of many isoforms, including I-FABP.[13] Laulumaa *et al.* suggests a new role for this conserved phenylalanine in fluctuations of the β-domain (the alternative portal) rather than the α-helical portal.[18]

To investigate the alternative portal and its proposed large-amplitude conformational movement, measurement of distances between strands D and E by CW EPR and DEER was carried out for doubly-labeled apo I-FABP at three points along the strands, shown in Fig[]. Solvent-facing nitroxides were chosen to monitor interstrand distances. The cross-cleft pairs are 57R1/73R1, 59R1/71R1, and 61R1/69R1.

At neutral pH, all crystal of I-FABP structures suggest that the interstrand distance is within 20 Å. In this distance window, dipolar broadening of the CW line shape has been shown to provide detailed structural information.[53] The extent of dipolar broadening of doubly labeled proteins is proportional to the inverse third power of the interspin distance (see Chapter 2). The

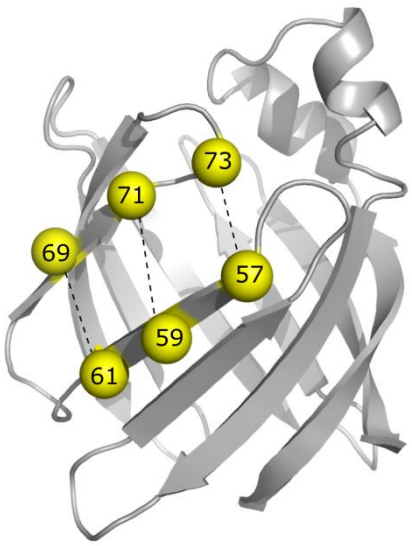

Figure 4.15 Nitroxide sites used to measure the interstrand distance in apo I-FABP as a function of pH.

distribution of interspin distances can be determined from the broadened line shape and the corresponding unbroadened line shape by a spectral deconvolution method that assumes a random orientation of the interspin vector with respect to the magnetic field.[53] This method applies well to nitroxides which are between 8 – 20 Å apart. For distances shorter than 8 Å, electron exchange interactions dominate the dipolar interaction; beyond 20 Å, the dipolar coupling is too weak to affect the line shape. In the slow tumbling limit (i.e. attached to solid support) and at room temperature, this method has been shown to robustly determine intra-protein distances in T4 lysozyme by comparing the line shape of R1-labeled protein (the broadened, or interacting, line shape) to that of the same double-cysteine mutant labeled with a mixture of R1 and a diamagnetic analogue; here, the N-acetylated form of R1. This effectively removes the dipolar interaction and is essentially the algebraic sum of the individual spectra in the absence of magnetic interaction (the "sum of singles" line shape).[53]

Fig 4.16a show "interacting" and "sum of singles" EPR spectra of 57R1/73R1, 59R1/71R1, and 61R1/69R1 at pH 7.3; the derived distance distributions are shown in the insets. The distance distributions indicate that the average distance between strands D and E at each pair are equal (the strands are parallel) and have similar distribution widths (the pair closest to the ß-turn, 61R1/69R1, is not more rigid than the pair furthest away from the turn, 57R1/73R1). At pH 7.3, these strands are well-ordered, despite the lack of backbone H-bonding.

At pH 2.2, the dipolar broadened EPR spectra suggest the distances exceed the limit measureable by CW dipolar broadening (Fig. 4.16b). The distance window for DEER, between 16-75 Å[54–56], is well-suited to measure the distance distributions beyond 16 Å. In Fig. 4.16c, the DEER dipolar evolution functions (DEFs) at pH 2.2 are shown. Long distances due to local flexibility in the cleft at neutral pH, undetectable by dipolar broadening, would be within the

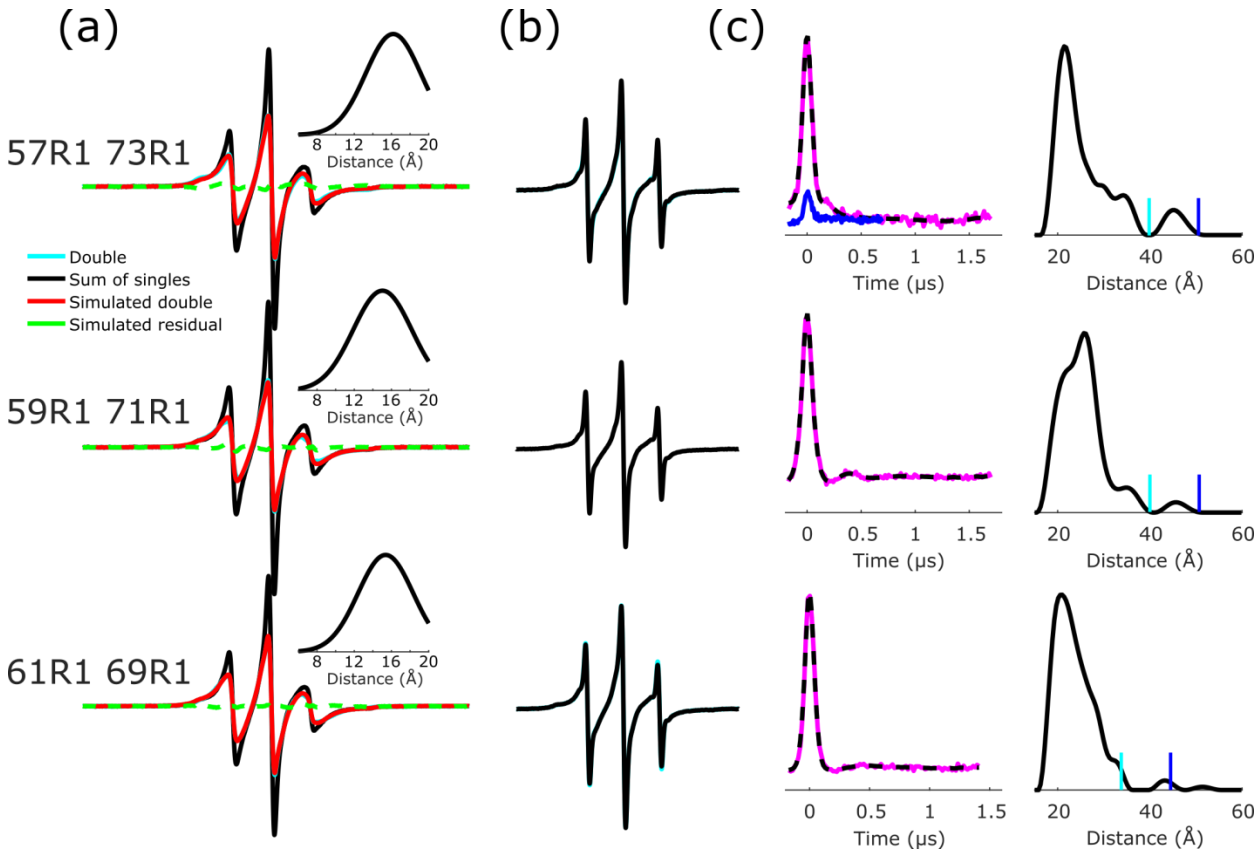

Figure 4.16 Short and long distances determined between nitroxides placed on pseudo edge strands D and E in the non-H-bonded cleft. (a) Determination of distances by dipolar broadening demonstrate parallel, closely-spaced strands at pH 7.3. (b) Line shapes at pH 2.2 show no dipolar broadening. (c) Long distance distributions from DEER DEFs show a significant increase in the distance between strands D and E at pH 2.2. The cyan bar shows the confidence limit for distribution shape, the blue bar shows the confidence limit for the maximum distance. Overlaid in the DEER dipolar evolution of 57R1/73R1 at pH 2.2 in magenta is the scaled DEER signal at pH 7.3 in blue.

distance range of DEER. Overlaid with the DEF for 57R1/73R1 at pH 2.2, magenta, is the DEF for the same pair at pH 7.3. The close proximity of the two labels at neutral pH cause the modulation depth of the signal to decrease significantly[57], suggesting that strands D and E at neutral pH are closely spaced and do not populate the open orientation observed at low pH.

4.5.3 Discussion: The alternative portal region undergoes a conformational change

Dipolar broadening of CW EPR spectra and DEER EPR complement each other to determine distance distributions both in a short distance regime (8 – 20 Å) and in a long distance regime (~25 – 80 Å). At pH 2.2, strands C and D splay apart and their respective fluctuation widths get larger. Although a simulated conformation intended to model ligand access to the β-domain interior describes a nearly 20 Å displacement of the C-D loop from the E-F loop[18], the dipolar broadening and DEER data suggest that the movement of cleft strands D and E do not tilt to open an alternative portal but instead disengage and fluctuate parallel to each other. The present data do not show a gradient in the magnitude of distance increase from what would be the smallest, near the D-E loop (a short, 3-amino acid β-hairpin, proximal to 61R1/69R1), to the largest, near the helical domain (proximal to 57R1/73R1). The "open alternative portal" may still be a component of the acid MG state of I-FABP, but experimental evidence indicates that the conformational state is not identical to that observed by molecular dynamics simulation.

4.6 Results: Acid pH hydrates the cavity of apo I-FABP

Evidence for MG-like structural fluctuations at pH 2.2 so far have been extensive, involving conformational exchange in the α-helices, increased side chain dynamics in the solvent-facing ß-strands, and conformational exchange in strands D and E. With population of a MG at low pH, the hydrophobic core of I-FABP is expected to become hydrated (see Chapter 1). So far, line shape analysis of CW spectra using the spin label MTSL has focused on the mobility of the spin label. In addition to motion, the nitroxide magnetic parameters (A- and g-tensors) are dependent on solvent polarity and proticity.[58] The isotropic hyperfine coupling constant (A_{iso}) and the A-tensor component directed along the π orbital of the nitroxide radical (A_{zz}) are readily

determined and can be used as reference values for describing the local solvent environment of spin-labeled proteins.[59] These parameters are used here to monitor the polarity within the cavity of apo I-FABP.

The four cavity-facing amino acids in strand H are L102, A104, R106, and I108. In the apo I-FABP crystal structure (1IFC), the guanidino group of R106 makes a large number of polar contacts within a partially hydrated cavity: 5 water molecules, which include solvent molecules Sol134, Sol142, Sol147, Sol148, and Sol154, as well as a hydrogen bond with Q115[8]. The alkyl side chains of L102, A104, and I108 are somewhat protected from solvent, though the crystal structure suggests the cavity contains a total of 24 internal ordered water molecules.

SDSL EPR of 102R1, 104R1, 106R1, and 108R1 at pH 7.3 have characteristically immobile line shapes near the rigid limit for X-band, which resemble the line shape of a buried nitroxide in T4 lysozyme.[39] At pH 2.2, a new mobile spectral component appears in the line shapes of 102R1, 104R1, and 108R1 that are easily distinguished from the rigid-limit line shapes at pH 7.3 by their high mobilities (Fig. 4.17a). The origin of the mobile component is presumably from a conformation that either allows entry of bulk water in the cavity or one in which the β-strand has a strong twist to place R1 in water. An estimate for the polarity of the highly mobile components at low pH were determined by measuring the magnitude of the EPR spectral parameter A_{iso}, the isotropic hyperfine splitting measured by the difference between the zero crossings of the low and center field lines (equivalently, between the center and high field lines) of the mobile component. This splitting is defined in Fig. 4.17a. In Chapter 2, the hyperfine coupling $\boldsymbol{A}$-tensor was shown to consist of two terms:

$$\boldsymbol{A} = A_{iso}\mathbf{1} + \boldsymbol{A}_{dipolar} \quad (1)$$

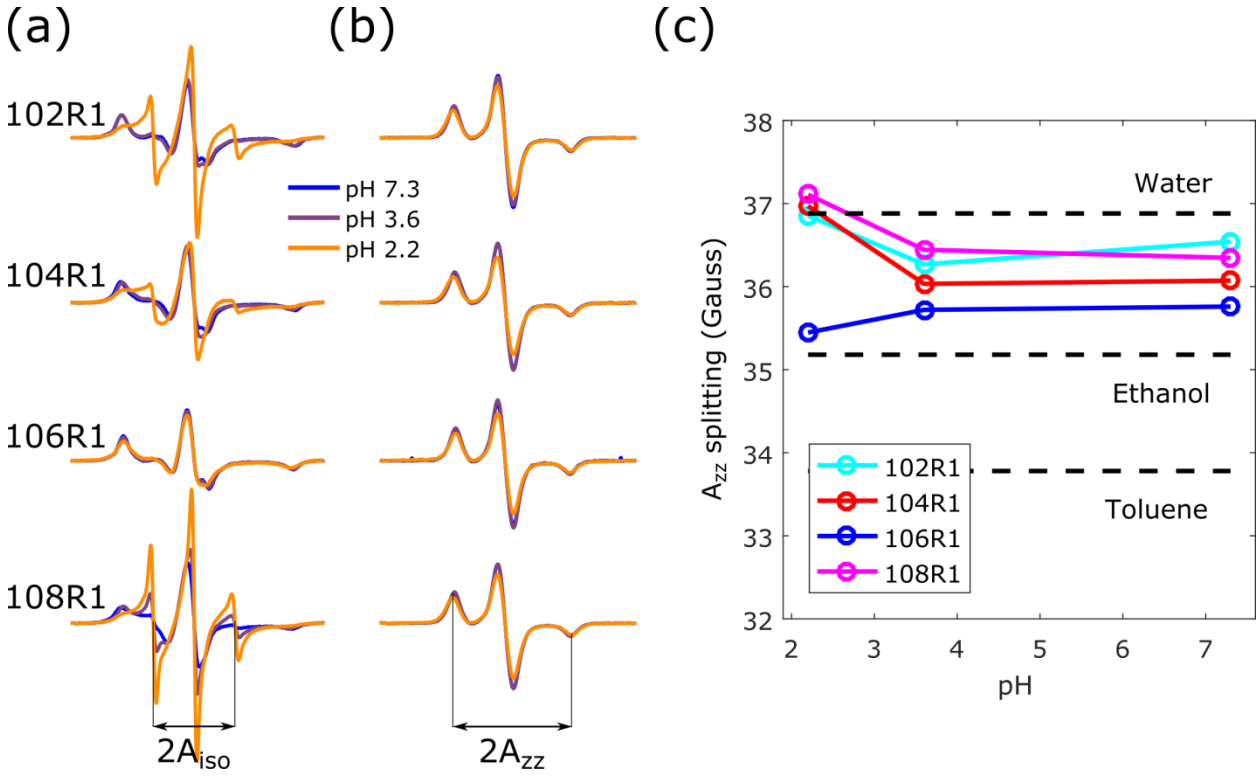

Figure 4.17 Hyperfine splittings of nitroxides are indicative of local polarity. (a) At room temperature, the isotropic splitting of the very mobile component at low pH (b) Frozen EPR spectra of nitroxides placed on the cavity-facing surface of strand H. (b) The splittings change as a function of pH and correspond to environments containing between 1 and 2 hydrogen bonds.

where **1** is the unit vector, A_{iso} is the isotropic term resulting from the Fermi contact interaction in the nitrogen nucleus, and $\boldsymbol{A}_{dipolar}$ is the classical dipolar coupling between the electron and the nitrogen nucleus calculated from a distance-dependent interaction Hamiltonian. A_{iso} is resolved in spectra with motions in the fast time regime (0.1 – 1 ns motions) (Fig 4.17a); in this fast tumbling limit, $\boldsymbol{A}_{dipolar}$ is fully averaged. In the presence of electrostatic interactions and hydrogen bonding, the magnitude of A_{iso} increases by increasing the spin density on the nitrogen atom. In bulk water, which has a high dielectric constant of 80, $A_{iso} \approx 16.1$ G.[58] In comparison, reports for dielectric constant values for the hydrophobic interior of a protein have varied from 4 – 20[60,61], for which $A_{iso} \approx 14.9$ G.[58] Using these values as a guide, and considering that the ligand-binding cavity of I-FABP has a large number of ordered crystallographic water

molecules, isotropic hyperfine splittings for nitroxides in the interior of I-FABP are expected to be large due to considerable polarity in the cavity even at neutral pH. The values of A_{iso} at pH 2.2 are summarized in Table 4.3.

Table 4.3 Splittings for the mobile component of low pH spectra of I-FABP within the ligand-binding cavity

	A_{iso}, mobile component				
	pH	Apo 102R1	Apo 104R1	Apo 106R1	Apo 108R1
Room temp, solid support	2.2	16.1	15.8	N/a	16.1

While the polarity of the mobile line shapes for nitroxides placed near the ends of strand H are similar to bulk solvent, referring to 102R1 near the helical domain and 108R1 opposite the helical domain, the splitting for 104R1 is smaller, suggesting that the mobile component of 104R1 is partially protected from solvent. The line shape of 106R1 at pH 2.2 does not have a mobile component.

The axial A_{zz} value in the rigid limit is another quantitative measure that depends on solvent polarity. The A_{zz} splitting of a nitroxide can be measured directly from EPR spectra at X-band in the rigid motional limit, i.e. at low temperature in liquid nitrogen (at 77 K), by the peak-to-peak splitting between the low field and high field lines, which is symmetric about the minimum of the center field line. Fig. 4.17b shows the relationship between pH and polarity for the same cavity-facing nitroxides in strand H. At pH 7.3 and 77 K, the A_{zz} splittings of 102R1, 104R1, 106R1, and 108R1 reflect a moderately hydrophilic environment, positioned between reference values for the A_{zz} splittings of a nitroxide in water and ethanol. The sites near the strand ends (102R1 and 108R1) are more hydrophilic than the sites in the center of strand H

(104R1 and 106R1). 104R1 and 106R1 are therefore protected from solvent. The hydrophobicity of 106R1 is higher than 104R1 due to replacement of the R106 counterion, which is preserved in the 104R1 mutant. Replacement of R106 by SDSL disrupts the hydrogen-bonding network between ordered waters and polar side chains located in the cavity, suggesting the cavity is partially dehydrated in the 106R1 mutant.

At pH 3.6, the hydration state of the cavity is virtually unchanged for 104R1, 106R1, and 108R1, while 102R1 experiences a small shift to a more hydrophobic environment. At pH 2.2, a measurable shift in the A_{zz} splittings of 102R1, 104R1, and 108R1 suggest increased water penetration into the cavity. This change is coincident with pH-driven population of the MG state of I-FABP. A comparison of the A_{zz} splittings, determined in the rigid limit at 77 K for these sites at pH 7.3, 3.6, and 2.2 are summarized in Table 4.4.

Table 4.4 Splittings for rigid limit spectra at variable pH values of I-FABP within the ligand-binding cavity

		A_{zz}			
	pH	102R1	104R1	106R1	108R1
Room temp, solution	7.3	30.1 G	29.7 G	28.6 G	28.7 G
Room temp, solid support	7.3	34.3 G	34.5 G	33.7 G	33.7 G
	3.6	33.8 G	33.2 G	33.4 G	32.5 G
	2.2	??	??	32.7 G	??
Liquid N2, solid support	7.3	36.5 G	35.9 G	35.7 G	36.4 G
	3.6	36.3 G	36.0 G	35.7 G	34.6 G
	2.2	36.8 G	36.9 G	35.4 G	37.1 G

For 298K sample conditions, table entries marked with ?? have very small high field lines such that the splitting is difficult to measure.

4.6.1 Discussion: Cavity polarity increases as pH decreases, suggesting cavity hydration in the MG state

Hyperfine splitting constants measured from spectra of nitroxides in the cavity of apo I-FABP show the interior of the protein is highly polar. While polarity measurements alone cannot distinguish between increased solvation or local charge, an argument in favor of cavity hydration is proposed to account for the change in internal polarity in the acid state of I-FABP. The lower polarity of 106R1 compared to 102R1, 104R1, and 108R1 suggests that replacement of R106 with a nitroxide side chain suppresses hydration of the cavity in the native state. Furthermore, the decrease in polarity between pH 3.6 and 2.2 at 106R1 site suggests that in the acid state of I-FABP, R106 is responsible for recruiting water molecules into the cavity. Because 102R1, 104R1, and 108R1 are in the vicinity of the static charge on R106 at all pHs investigated, the increase in polarity at pH 2.2 for these sites cannot be due to a change in the protonation state of R106. In the native state, Q115 interacts with R106. In the R106N mutant, Q115 shows a weak propensity to hydrogen bond with internal water molecules[62], suggesting that Q115 in the 106R1 mutant may also weakly hydrogen bond to water. Therefore, the change in internal polarity at low pH can be assigned to a difference in cavity hydration rather than side chain protonation.

At pH 7.3, there is no discernable gradient along strand H; that is, 102R1 near the helical domain and 108R1 at the base of the β-domain do not have significantly different hydrophobicities, though the center strand sites 104R1 and 106R1 are less polar than 102R1 and 108R1. Cavity hydration is therefore not exclusively driven by R106, but this side chain plays an important role in coordinating at least 4 crystallographic waters at neutral pH.

Some internal waters play structural roles: in addition to the stabilizing effect of internal water molecules by hydrogen bonding to buried carboxylate and amide groups[63], binding of

buried structural water is proposed to increase the flexibility of proteins[64], consistent with population of a highly flexible, hydrated MG at acid pH. Such structural waters are tightly associated with the protein. At pH 7.4 and 5.7, only three long-lived solvent molecules associated with apo I-FABP were identified by water ^{17}O and ^{2}H magnetic relaxation dispersion (MRD) NMR.[6] The remaining 21 solvent molecules observed in the crystal structure are suggested to have correlation times faster than the tumbling rate of the protein and therefore play minor structural roles, if at all. Careful consideration of crystallographic solvent molecules led to identification of Sol135 in the D-E loop of the non-H-bonded cleft and Sol137 in the portal as long-lived internal waters, where the third proposed solvent molecule Sol217 interacts with the solvent-facing surface cleft between strands H and I. Nitroxide monitors of internal polarity on strand H are not close to either Sol135 or Sol137; therefore, the high polarity measurements derived from A_{iso} and A_{zz} splittings of these nitroxide line shapes at pH 7.3 correspond to other loosely associated internal solvent molecules, since structural water that is not tumbling has a low dielectric constant.

Small deviations in the A_{zz} splittings could indicate structural changes compared to the native state. The A_{zz} splitting of 108R1 at pH 2.2 is found to exceed the reference value in bulk water, possibly due to interaction of the nitroxide side chain with a local charge. In addition to changes in cavity hydration, a model of increased dynamics of the H-I loop in the acid MG-like state could bring S109, N111, or E112 into contact with 108R1. Additionally, the small hydrophobic shift at 102R1 at pH 3.6 could be due to a local repacking of the nearby aromatic sidechains: neighboring strand G contains 93F and strand I contains 117Y, as well as 17F in helix I and 119Y in the I-F loop. In apo I-FABP, NMR of ^{19}F labeled 17F and 93F at pH 3.6 show line broadening, in support of population of a new state. A simple mechanism for

population of a native-like state unlike the MG at pH ~4 could be due to protonation of all solvent-exposed carboxylates excluding the protein core.

A last comment on the accuracy of the A_{zz} measurements involves separation of spectral components, which was possible at room temperature but not in liquid nitrogen at 77 K. For the spectra of 102R1, 104R1, and 108R1 obtained in the rigid limit at pHs below 7.3, A_{zz} is the weighted average of the splittings from the mobile and immobile components. At pH 3.6, the populations of the mobile components for all sites do not exceed 10%. Additionally, at pH 2.2, the mobile component population of 104R1 is less than 5%. For these spectra, the measured A_{zz} is dominated by the immobile component. However, a small correction to the measured A_{zz} at pH 2.2 can be calculated for 102R1 and 108R1, which have mobile component populations estimated by simulation of approximately ~25% and ~33%, respectively. This correction results in an increase of 0.1 Gs for 108R1 and a decrease of 0.1 Gs for 102R1, which is smaller than the change in the A_{zz} splitting between pH 3.6 and 2.2. Therefore, the A_{zz} measured in the rigid limit is a good measure of the change in the polarity of the immobile component.

Solvation of the I-FABP cavity is also related to the mechanism of ligand binding. I-FABP has low nanomolar affinity for hydrophobic ligands, especially for fatty acids.[65] The binding of these ligands requires reorganization of the internal waters in the protein: the cavity is partially dehydrated on introduction of a fatty acid[9], releasing 16-18 internal water molecules. The binding mechanism is, however, enthalpically driven, and the enthalpy of ligand binding is more significant than the entropy gain upon displacement of internal waters.[10,47] In general, crystal structures of apo FABPs all have a high degree of cavity hydration. In contrast, β-lactoglobulin, a protein with similar structure and function, has a large nonpolar binding cavity devoid of solvent waters[66], but a 100-fold lower affinity for palmitate compared to I-FABP.[67,68]

Alternative portals in I-FABP are thought to provide channels for solvent egress, whereas β-lactoglobulin has only one entry point, necessitating an evacuated cavity primed for ligand binding. Therefore, internal waters in I-FABP are expected to play both structural and functional roles, and structural flexibility in regions of the protein other than the portal are necessary for fatty acid binding and transport.

4.7 Results: Intrinsic fluorescence of the internal I-FABP cavity

Water penetration into the hydrophobic core is a key component of the molten globule. As a second perspective on hydration of the core, intrinsic tryptophan fluorescence was measured to assess the local polarity of the two native tryptophans: W6 and W82. The pH dependence on the maximum emission wavelength of W6 and W82 were determined separately by using single-Trp mutants of I-FABP, which have been characterized previously.[69] The fluorescence spectra of W6 and W82 are nearly additive to that of the wild-type protein , suggesting that they are noninteracting.[69] The intrinsic fluorescence of W6 and W82 are therefore useful reporters for two distinct environments within the ligand binding cavity. The stability loss by mutation to tyrosines measured by urea denaturation is small, with midpoints shifted from 4.2 M to 3.5 M urea and $\Delta\Delta G$ of 0.6 – 0.8 kcal mol^{-1}.[69] Thus, the mutations W6Y and W82Y do not unfold I-FABP and reflect native-like structure of the apoprotein.

W82 is a member of the core of I-FABP, formed by a cluster of hydrophobic side chains in strands C, D, E, F, and G opposite the helical domain (Fig. 4.18b). W82 interacts with two sidechains and is hydrogen bonded to an interior water molecule. W82 forms a T-shaped aromatic interaction with F93, connecting strands F and G. F93 is shown by ^{19}F-Phe labeled NMR to be conformationally sensitive, exhibiting microsecond motion on the intermediate NMR

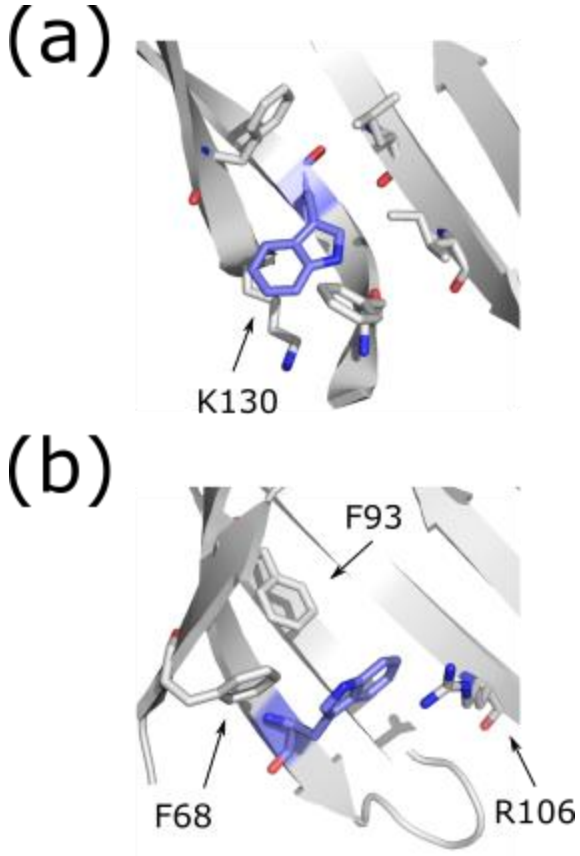

Figure 4.18 Location of tryptophan residues (a) W6 and (b) W82 and interactions with nearby side chains that influence the emission wavelength and amplitude.

timescale indicated by its broad line width, which can be deconvoluted into two peaks at higher temperatures.[70] However, the exchange observed in F93 may be a result of conformational motion adjacent to the sidechain. For example, F68, which forms NOE cross peaks with F93, is also in microsecond exchange. F68 is on strand E in the non-hydrogen-bonded cleft, which could potentially allow for multiple side chain orientations due to lower side chain packing density. As indicated by the discussion earlier, wide distribution widths between dipolar-broadened nitroxides across strands D and E demonstrate a range of side chain flexibility at pH 7.3. The fluorescence emission spectrum of W82 at 22 °C does not indicate more than one orientation of the tryptophan, however, as the profile is fit well to a single log-normal distribution[71] (Fig. 4.19a and 4.19b, insets). W82 makes a second interaction via a parallel cation-π interaction with R106. The guanidinium carbon of R106 is 6 Å away from the aromatic ring of W82, which is within

the range expected to produce a red shift in the emission spectrum of W82.[72]

The second tryptophan in I-FABP is W6, which is 13 Å away from W82 on strand A (Fig. 4.18a). W6 is protected from waters in the cavity by a second cluster of hydrophobic sidechains including F2, L38, I40, L113, and F128. These residues bridge strands A, B, and I, forming a network separate from the core. The fluorescence intensity of W6 is significantly quenched, compared to W82. The backbone carbonyl oxygen of F2 is 2.8 Å away from the W6 indole nitrogen and could give rise to efficient quenching through interactions by excited-state electron transfer with the peptide bond. This mechanism has been proposed using peptides to account for most of the variation in fluorescence intensity across different proteins.[73] The amine group of K130 also forms an interaction with the π electron cloud, quenching by excited-state proton transfer.[74]

The change in fluorescence emission spectra of W6 and W82 as a function of pH is shown in Fig. 4.19. At neutral pH, the maximum emission wavelengths reflect the effects from the proposed interactions with W6 and W82. The fluorescence maximum of W82 is virtually unchanged at all pHs between 7.3 and 2.2 and remains fixed at a value that is red shifted compared to W6 in the native state, though the fluorescence intensity at pH 6.1 increases due to an unknown mechanism. Similarly, the fluorescence maximum of W6 is stable between pH 7.3 and 2.8, but at pH 2.2 a red shift occurs. This shift is coincident with population of the pH-stabilized MG state identified by other methods. Reference values for the fluorescence maximum in the urea-denatured state are provided for context to show that neither the red shifted fluorescence maximum of W82 nor the red shift at acid pH of W6 reflect unfolded states of I-FABP.

Pressure populates a non-native state that at neutral pH exhibits similar fluorescence

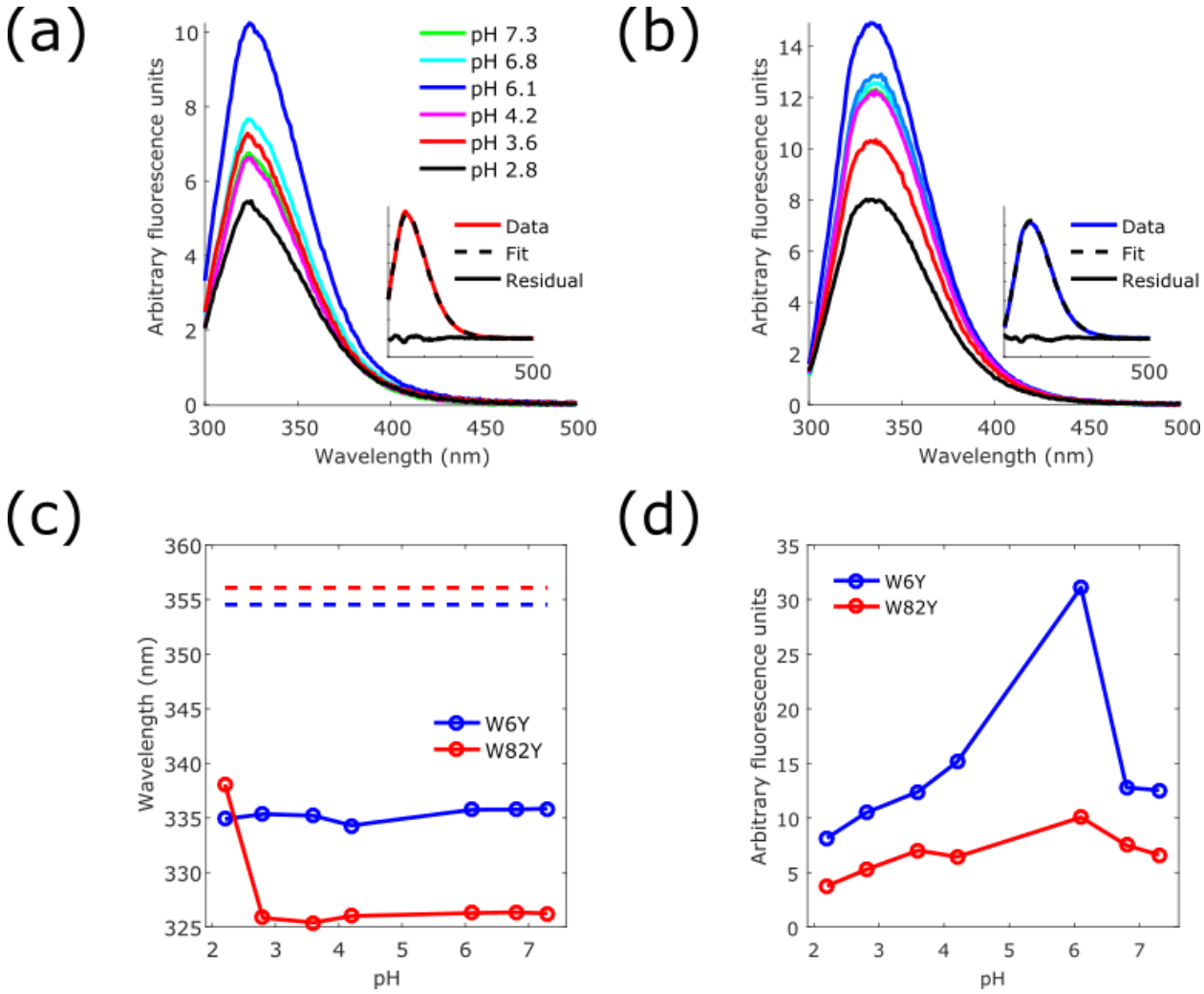

Figure 4.19 pH dependence of the intrinsic tryptophan fluorescence of I-FABP. Fluorescence spectra of (a) W82Y and (b) W6Y are shown. Insets demonstrate the log-normal fits to the respective I-FABP mutant at pH 7.3. The fitted maximal emission spectra wavelengths are plotted against pH in (c), where the dashed lines correspond to the maximal wavelength in 7.8M urea at pH 7.3. The change in fluorescence amplitude is shown in (d) for I-FABP W82Y and W6Y at a concentration of 1.9 µM at 22 °C.

signatures to the acid MG (Fig. 4.20). The fluorescence maximum of W6Y is relatively unchanged between 0 – 3 kbar. The small blue shift in W6Y at pressures above 1500 bar could be due to electrostriction of waters around R106, by which water more tightly associates with the R106 sidechain, shielding the positive charge from W82. This small reorganization of solvent within the cavity of I-FABP would result in a small decrease in polarity. The fluorescence

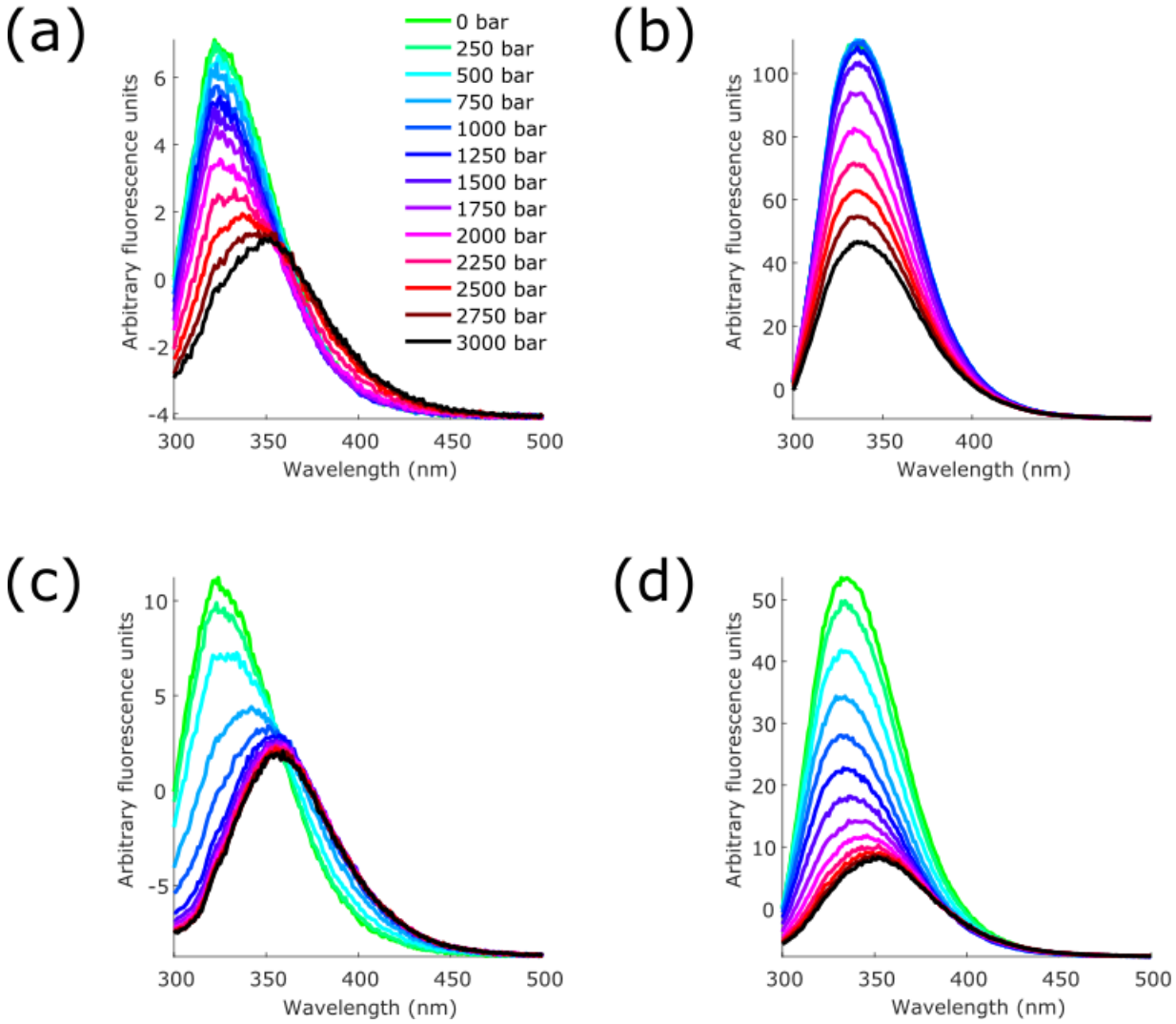

Figure 4.20 Pressure perturbation on the intrinsic tryptophan fluorescence of I-FABP. At pH 7.3, (a) W82Y and (b) W6Y. At pH 3.0, (c) W82Y and (d) W6Y.

maximum of W82Y is unchanged between 0 – 1500 bar, but higher pressures introduce a large red shift. At 3 kbar, the fluorescence maximum of W6 exceeds that of W82. Comparisons of the urea-denatured and the pressure-denatured fluorescence emission spectra also highlight similarities in their line shapes, suggesting that pressure populates a MG in a manner similar to acid pH.

As a first approximation, destabilizing the structure of I-FABP at pH 3.0, which does not

significantly populate the MG, can shift the pressure denaturation curve towards lower pressures. Under this condition, the fluorescence maximum of W82Y reaches the urea-denatured value of 356 nm at 2.0 kbar. Pressure denaturation of W6Y is also apparent at pH 3.0 following the proposed electrostriction effect at 500 bar. It is not possible by fluorescence alone to determine whether this state (pH 3.0 and 2.0 kbar) retains structure. As a future experiment, variable pressure CW EPR of nitroxides in the beta domain could easily distinguish between partially folded and unfolded secondary structure by comparison of lineshapes on the cavity interior and exterior of a single strand at low pH and high pressure. Alternatively, determining the solvent accessibility using a high-contrast exchanger such as $Cr(malt)_3$ in combination with SR.

4.7.1 Discussion: Similarities between low pH and high pressure MG states of I-FABP

W6 in the native state at pH 7.3 has a fluorescence spectrum typical of a buried, well-packed tryptophan. The local polarity at W6 is unchanged even at pH 2.8, indicating that hydrophobic contacts in an acidified state of I-FABP can remain native-like. Because the pI of I-FABP is 7.3, the net charge on I-FABP is positive at any pH below 7. Accumulation of positive charge around pH 4 due to protonation of carboxylates does not produce a change in the structure of I-FABP. Population of the MG state at pH 2.2, however, is associated with hydration of the hydrophobic core, to which W6 is sensitive. In contrast, the stability of the fluorescence maximum of W82 is probably a consequence of the unchanged protonation state of R106. Because the two tryptophans are quenched to different degrees, the previous work fails to show that W6 and W82 report on core hydration and the protonation state of R106, respectively. A decrease in fluorescence intensity at 328 nm of the wild-type protein, which contains both W6 and W82, has been interpreted as a result of formation of an intermediate state.[51] However, it is

clear that the two tryptophans have different pH sensitivities in their emission spectra. The fluorescence intensity changes with pH for W6 and W82 are different as well, complicating the analysis for concurrent fluorescence of W6 and W82 in the wild-type protein.

MG states of some proteins are thought to be related to transient intermediates in the folding pathway.[75] The kinetic refolding process of I-FABP is complex, starting from an extended unfolded form (U_E) that rapidly forms a partially ordered state during a burst phase (U_B) (100 μs)[76,77], which relaxes into a pre-molten globule (I_{TC}) and a molten globule (I_{MG}) (~700 μs), and finally the native state (N) (~30 ms).[78] Double-jump experiments revealed these intermediates indirectly. These studies provided the first indication of a molten globule intermediate for I-FABP, populated under mild denaturant concentrations. Refolding of W6Y proceeds through a simpler pathway which omits the molten globule intermediate, suggesting that W82 is native-like in the MG state and cannot be distinguished from the native state by fluorescence alone.[69] This is recapitulated in the pH titration data, in which W82 has no apparent change in fluorescence maximum between the pH 2.2 MG state and the pH 7.3 native state. Although the urea-populated and pH-populated molten globule states may not be the same, identification of a multitude of conformational states is direct evidence for the ruggedness of the conformational landscape observed for proteins in solution.

4.8 Results: Hydrostatic pressure populates EPR spectral components in R1 at neutral pH similar to those in the acid-stabilized MG state

The high pressure line shapes of 92R1, 105R1, and 116R1 in apo I-FABP resemble the acid pH state, notably the mobile component of 105R1 and the appearance of immobile components in 92R1 and 116R1 (Fig. 4.21). At pressures exceeding 1500 bar, however, the

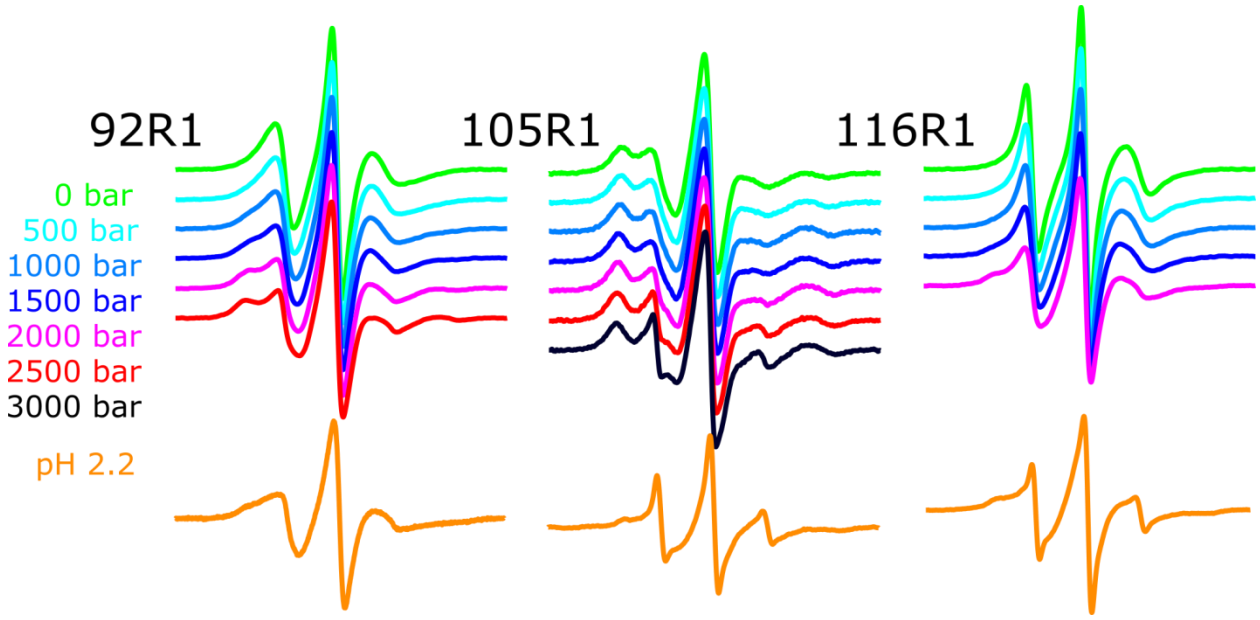

Figure 4.21 SDSL EPR of sites in apo I-FABP with increasing pressure show reversible changes in the lineshape, indicating population of a new state. Orange line shapes are reproduced from Fig. 4.3.

reversibility of the line shape changes were poor, indicating population of conformational states not in thermodynamic equilibrium. In fact, pressurization of 39R1 in strand B, 50R1 in strand C, and 69R1 in strand E all showed population of immobile components with increasing pressure that were irreversible on returning to atmospheric pressure due to aggregation. Aggregation itself indicates that pressure is populating a new state, presumably with exposure of hydrophobic contacts. The instability of site-specific mutations of apo I-FABP suggests the relatively hydrophobic R1 side chain participates in or nucleates aggregation at low pH. In comparison, the reversibility of the intrinsic fluorescence experiments with tryptophan mutagenesis is enhanced by the dilute sample requirements. The similarity between hydrostatic pressure-populated and low pH-populated states indicated by these preliminary results support interpretation of similar MG-like conformational states populated by both experimental conditions. Additionally, the separate mechanistic underpinnings of pressure and pH could imply that I-FABP has a finite

number of conformational states.

4.8.1 Discussion: The apo I-FABP MG

A common mechanism linking protein structural changes at acid pH and high hydrostatic pressure, as discussed in Chapter 3, may be electrostriction of a salt bridge. Candidate salt bridges were investigated in previous sections of this chapter, but no single mutation was shown to pre-populate the MG conformation at neutral pH. However, similarities between the acid pH state and the high pressure state by CW EPR and intrinsic fluorescence suggest a common MG conformation. This initial result is consistent with the pH and pressure-populated MG states of apo Mb, which were shown to have only few structural dissimilarities evinced by small differences in the EPR spectra and in secondary structure content determined by CD.[41] The proposed functions of the MG (i.e. in folding, in ligand binding, and in cellular localization) all take advantage of the loss of tertiary structure and unique heterogeneity of the MG conformational state. This dissertation aims to characterize the extent of structural heterogeneity of the MG from the SDSL EPR perspective, which encompasses fluctuation amplitudes, motional timescale, and the thermodynamic equilibrium state of the MG. That the acid-stabilized MG state of I-FABP is unlike an unfolded conformation, the β-domain is maintained at acid pH, and the α-helices undergo large-amplitude motions on a μs-ms timescale are all properties characterized by SDSL EPR which together confirm the existence of a unique MG in a β-sheet protein.

4.9 Bibliography

1. Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. Crystal structure of rat intestinal fatty-

acid-binding protein. Refinement and analysis of the Escherichia coli-derived protein with bound palmitate. *J. Mol. Biol.* **208**, 327–339 (1989).

2. Patil, R. *et al.* Characterization of two distinct modes of drug binding to human intestinal fatty acid binding protein. *ACS Chem. Biol.* **9**, 2526–2534 (2014).
3. Storch, J. & Corsico, B. The Emerging Functions and Mechanisms of Mammalian Fatty Acid – Binding Proteins. (2008). doi:10.1146/annurev.nutr.27.061406.093710
4. Franchini, G. R., Storch, J. & Corsico, B. The integrity of the α-helical domain of intestinal fatty acid binding proteins is essential for the collision-mediated transfer of fatty acids to phospholipid membranes. *Biochim. Biophys. Acta* **1781**, 192–199 (2008).
5. Ropson, I. J. & Frieden, C. Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by 19F NMR. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7222–7226 (1992).
6. Wiesner, S., Kurian, E., Prendergast, F. G. & Halle, B. Water molecules in the binding cavity of intestinal fatty acid binding protein: Dynamic characterization by water 17O and 2H magnetic relaxation dispersion. *J. Mol. Biol.* **286**, 233–246 (1999).
7. Zhu, L., Kurian, E., Prendergast, F. G. & Kemple, M. D. Dynamics of Palmitic Acid Complexed with Rat Intestinal Fatty Acid Binding. 1554–1561 (1999). doi:10.1021/bi982087v
8. Scapin, G., Gordon, J. I. & Sacchettini, J. C. Refinement of the structure of recombinant rat intestinal fatty acid- binding apoprotein at 1.2-Å resolution. *J. Biol. Chem.* **267**, 4253–4269 (1992).
9. Sacchettini, J. C. & Gordon, J. I. Rat intestinal fatty acid binding protein. A model system for analyzing the forces that can bind fatty acids to proteins. *J. Biol. Chem.* **268**, 18399–

18402 (1993).

10. Richieri, G. V, Ogata, R. T. & Kleinfeld, A. M. Thermodynamics of Fatty Acid Binding to Fatty Acid-binding Proteins and Fatty Acid Partition between Water and Membranes Measured Using the Fluorescent Probe ADIFABP. 15076–15084 (1995).
11. Hodsdon, M. E. & Cistola, D. P. Ligand Binding Alters the Backbone Mobility of Intestinal Fatty Acid-Binding Protein as Monitored by 15N NMR Relaxation and 1H Exchange. *Biochemistry* **36**, 2278–2290 (1997).
12. Richieri, G. V., Low, P. J., Ogata, R. T. & Kleinfeld, A. M. Thermodynamics of fatty acid binding to engineered mutants of the adipocyte and intestinal fatty acid-binding proteins. *J. Biol. Chem.* **273**, 7397–7405 (1998).
13. Hodsdon, M. E. & Cistola, D. P. Discrete Backbone Disorder in the Nuclear Magnetic Resonance Structure of Apo Intestinal Fatty Acid-Binding Protein : Implications for the Mechanism of Ligand. *Biochemistry* **36**, 1450–1460 (1997).
14. Zanotti, G., Feltre, L. & Spadon, P. A possible route for the release of fatty acid from fatty acid-binding protein. *Biochem. J.* **301**, 459–463 (1994).
15. Long, D., Mu, Y. & Yang, D. Molecular Dynamics Simulation of Ligand Dissociation from Liver Fatty Acid Binding Protein. **4**, 1–10 (2009).
16. Mihajlovic, M. & Lazaridis, T. Modeling fatty acid delivery from intestinal fatty acid binding protein to a membrane. *Protein Sci.* **16**, 2042–2055 (2007).
17. Lowe, J. B., Sacchettini, J. C., Laposata, M., McQuillan, J. J. & Gordon, J. I. Expression of rat intestinal fatty acid-binding protein in Escherichia coli. Purification and comparison of ligand binding characteristics with that of Escherichia coli-derived rat liver fatty acid-binding protein. *J. Biol. Chem.* **262**, 5931–5937 (1987).

18. Laulumaa, S. *et al.* Structure and dynamics of a human myelin protein P2 portal region mutant indicate opening of the β barrel in fatty acid binding proteins. *BMC Struct. Biol.* **18**, 1–13 (2018).

19. Friedman, R., Nachliel, E. & Gutman, M. Fatty acid binding proteins: Same structure but different binding mechanisms? Molecular dynamics simulations of intestinal fatty acid binding protein. *Biophys. J.* **90**, 1535–1545 (2006).

20. Friedman, R., Nachliel, E. & Gutman, M. Molecular dynamics simulations of the adipocyte lipid binding protein reveal a novel entry site for the ligand. *Biochemistry* **44**, 4275–4283 (2005).

21. Richieri, G. V., Ogata, R. T. & Kleinfeld, A. M. Kinetics of fatty acid interactions with fatty acid binding proteins from adipocyte, heart, and intestine. *J. Biol. Chem.* **271**, 11291–11300 (1996).

22. Cistola, D. P., Kim, K., Rogl, H. & Frieden, C. Fatty Acid Interactions with a Helix-less Variant of Intestinal Fatty Acid-Binding. **2960**, 7559–7565 (1996).

23. Chattopadhyay, K., Saffarian, S., Elson, E. L. & Frieden, C. Measurement of microsecond dynamic motion in the intestinal fatty acid binding protein by using fluorescence correlation spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14171–6 (2002).

24. Wootan, M. G. & Storch, J. Regulation of fluorescent fatty acid transfer from adipocyte and heart fatty acid binding proteins by acceptor membrane lipid composition and structure. *J. Biol. Chem.* **269**, 10517–10523 (1994).

25. Córsico, B., Franchini, G. R., Hsu, K.-T. & Storch, J. Fatty acid transfer from intestinal fatty acid binding protein to membranes: electrostatic and hydrophobic interactions. *J. Lipid Res.* **46**, 1765–1772 (2005).

26. Prats, M., Tocanne, J.-F. & Teissie, J. Lateral proton conduction at a lipid/water interface. *Eur. J. Biochem.* **162**, 379–385 (1987).

27. Van Der Goot, F. G., González-Mañas, J. M., Lakey, J. H. & Pattus, F. A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature* **354**, 408–410 (1991).

28. Herr, F. M., Matarese, V., Bemlohr, D. A. & Storch, J. Surface Lysine Residues Modulate the Collisional Transfer of Fatty Acid from Adipocyte Fatty Acid Binding Protein to Membranes. *Biochemistry* **34**, 11840–11845 (1995).

29. Li, H. & Frieden, C. Fluorine-19 NMR studies on the acid state of the intestinal fatty acid binding protein. *Biochemistry* **45**, 6272–6278 (2006).

30. Li, H. & Frieden, C. Phenylalanine side chain behavior of the intestinal fatty acid-binding protein: The effect of urea on backbone and side chain stability. *J. Biol. Chem.* **280**, 38556–38561 (2005).

31. Carrey, E. A. & Pain, R. H. Conformation of a stable intermediate on the folding pathway of Staphylococcus aureus penicillinase. *Biochim. Biophys. Acta* **533**, 12–22 (1977).

32. Hughson, F. M., Wright, P. E. & Baldwin, R. L. Structural characterization of a partly folded apomyoglobin intermediate. *Science (80-.).* **249**, 1544–1548 (1990).

33. Vassilenko, K. S. & Uversky, V. N. Native-like secondary structure of molten globules. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1594**, 168–177 (2002).

34. Levitt, M. & Greer, J. Automatic identification of secondary structure in globular proteins. *J. Mol. Biol.* **114**, 181–239 (1977).

35. López, C. J., Oga, S. & Hubbell, W. L. Mapping molecular flexibility of proteins with site-directed spin labeling: A case study of myoglobin. *Biochemistry* **51**, 6568–6583

(2012).

36. López, C. J., Fleissner, M. R., Brooks, E. K. & Hubbell, W. L. Stationary-phase EPR for exploring protein structure, conformation, and dynamics in spin-labeled proteins. *Biochemistry* **53**, 7067–7075 (2014).

37. Xiao, H. *et al.* Exploring the potential impact of an expanded genetic code on protein function. *Proc. Natl. Acad. Sci.* **112**, 6961–6966 (2015).

38. Lietzow, M. A. & Hubbell, W. L. 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 (2004).

39. Guo, Z., Cascio, D., Hideg, K., Kálái, T. & Hubbell, W. L. 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 (2007).

40. Mchaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**, 7692–7704 (1996).

41. Lerch, M. T., Horwitz, J., McCoy, J. & Hubbell, W. L. Circular dichroism and site-directed spin labeling reveal structural and dynamical features of high-pressure states of myoglobin. *Proc. Natl. Acad. Sci.* **110**, E4714–E4722 (2013).

42. Gasymov, O. K., Abduragimov, A. R. & Glasgow, B. J. Evidence for Internal and External Binding Sites on Human Tear Lipocalin. *Arch Biochem Biophys* **468**, 15–21 (2007).

43. MD, B. *et al.* Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* **37**, 363–390 (2010).

44. Borgia, G. C., Brown, R. J. S. & Fantazzini, P. Uniform-Penalty Inversion of Multiexponential Decay Data. *J. Magn. Reson.* **132**, 65–77 (1998).

45. Eaton, S. S. & Eaton, G. R. Multifrequency Pulsed EPR and the Characterization of Molecular Dynamics. *Methods Enzym.* **563**, 37–58 (2015).

46. Bridges, M., Hideg, K. & Hubbell, W. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* **37**, 363–390 (2010).

47. Richieri, G. V, Low, P. J., Ogata, R. T. & Kleinfeld, A. M. Mutants of Rat Intestinal Fatty Acid-binding Protein Illustrate the Critical Role Played by Enthalpy-Entropy Compensation in Ligand Binding. 16737–16741 (1997).

48. Tielker, N., Eberlein, L., Chodun, C., Güssregen, S. & Kast, S. M. pKa calculations for tautomerizable and conformationally flexible molecules: partition function vs . state transition approach. *J. Mol. Model.* **25**, 1–6 (2019).

49. Pahari, S., Sun, L. & Alexov, E. PKAD: a database of experimentally measured pKa values of ionizable groups in proteins. *Database* 1–7 (2019). doi:10.1093/database/baz024

50. Franchini, G. R., Curto, L. M. & Caramelo, J. J. Dissection of a β-barrel motif leads to a functional dimer: The case of the intestinal fatty acid binding protein. *Protein Sci.* **18**, 2592–2602 (2009).

51. Chattopadhyay, K. & Frieden, C. Steady-state and time-resolved fluorescence studies of the intestinal fatty acid binding protein. *Proteins* **63**, 327–335 (2006).

52. Simpson, M. A. & Bernlohr, D. A. Analysis of a Series of Phenylalanine 57 Mutants of the Adipocyte Lipid-Binding. **2960**, 10980–10986 (1998).

53. Altenbach, C., Oh, K. J., Trabanino, R. J., Hideg, K. & Hubbell, W. L. Estimation of inter-

residue distances in spin labeled proteins at physiological temperatures: Experimental strategies and practical limitations. *Biochemistry* **40**, 15471–15482 (2001).

54. Jeschke, G., Bender, A., Paulsen, H., Zimmermann, H. & Godt, A. Sensitivity enhancement in pulse EPR distance measurements. **169**, 1–12 (2004).

55. Schiemann, O. *et al.* A PELDOR-Based Nanometer Distance Ruler for Oligonucleotides. 5722–5729 (2004). doi:10.1021/ja0393877

56. Weber, A., Schiemann, O., Bode, B. & Prisner, T. F. PELDOR at S- and X-Band Frequencies and the Separation of Exchange Coupling from Dipolar Coupling. **285**, 277–285 (2002).

57. Banham, J. E. *et al.* Distance measurements in the borderline region of applicability of CW EPR and DEER: A model study on a homologous series of spin-labelled peptides. **191**, 202–218 (2008).

58. Owenius, R., Engström, M., Lindgren, M. & Huber, M. Influence of solvent polarity and hydrogen bonding on the EPR parameters of a nitroxide spin label studied by 9-GHz and 95-GHz EPR spectroscopy and DFT calculations. *J. Phys. Chem. A* **105**, 10967–10977 (2001).

59. Steinhoff, H.-J. *et al.* High-field EPR studies of the structure and conformational changes of site-directed spin labeled bacteriorhodopsin. *Biochim. Biophys. Acta* **1457**, 253–262 (2000).

60. Gilson, M. K. & Honig, B. H. The Dielectric Constant of a Folded Protein. **25**, 2097–2119 (1986).

61. Antosiewicz, J., McCammon, J. A. & Gilson, M. K. Prediction of pH-dependent Properties of Proteins. *J Mol Biol* **238**, 415–436 (1994).

62. Eads, J., Sacchettini, J. C., Kromminga, A. & Gordon, J. I. Escherichia coli-derived Rat Intestinal Fatty Acid Binding Protein with Bound Myristate at 1.5 Å Resolution and I-FABPArg106→Gln with Bound Oleate at 1.74 Å Resolution. *J. Biol. Chem.* **268**, 26375–26385 (1993).

63. Baker, E. N. & Hubbard, R. E. Hydrogen Bonding in Globular Proteins. *Prog. Biophys. molec. Biol.* **44**, 97–179 (1984).

64. Fischer, S. & Verma, C. S. Binding of buried structural water increases the flexibility of proteins. *Proc Natl Acad Sci U S A* **96**, 9613–9615 (1999).

65. Richieri, G. V, Ogata, R. T., Zimmerman, A. W., Veerkamp, J. H. & Kleinfeld, A. M. Fatty Acid Binding Proteins from Different Tissues Show Distinct Patterns of Fatty. 7197–7204 (2000). doi:10.1021/bi000314z

66. Qvist, J., Davidovic, M., Hamelberg, D. & Halle, B. A dry ligand-binding cavity in a solvated protein. **2008**, (2008).

67. Richieri, G. V., Ogata, R. T. & Kleinfeld, A. M. Equilibrium Constants for the Binding of Fatty Acids with Fatty. *J. Biol. Chem.* **269**, 23918–23930 (1994).

68. Kontopidis, G., Holt, C. & Sawyer, L. The Ligand-binding Site of Bovine β-Lactoglobulin: Evidence for a Function? *J Mol Biol* **318**, 1043–1055 (2002).

69. Dalessio, P. M., Fromholt, S. E. & Ropson, I. J. The Role of Trp-82 in the Folding of Intestinal Fatty Acid Binding Protein. **183**, 176–183 (2005).

70. Li, H. & Frieden, C. NMR studies of 4-19F-phenylalanine-labeled intestinal fatty acid binding protein: Evidence for conformational heterogeneity in the native state. *Biochemistry* **44**, 2369–2377 (2005).

71. Burstein, E. A. & Emelyanenko, V. I. Log-normal description of fluorescence spectra of

organic fluorophores. *Photochem. Photobiol.* **64**, 316–320 (1996).

72. Gallivan, J. P. & Dougherty, D. A. Cation-π interactions in structural biology. *Proc Natl Acad Sci U S A* **96**, 9459–9464 (1999).

73. Pan, C. & Barkley, M. D. Conformational Effects on Tryptophan Fluorescence in Cyclic Hexapeptides. *Biophys. J.* **86**, 3828–3835 (2004).

74. Chen, Y. & Barkley, M. D. Toward Understanding Tryptophan Fluorescence in Proteins. *Biochemistry* **37**, 9976–9982 (1998).

75. Christensen, H. & Pain, R. H. Invited review Molten globule intermediates and protein folding. *Eur Biophys J* **19**, 221–229 (1991).

76. Kim, K., Ramanathan, R. & Frieden, C. Intestinal fatty acid binding protein: a specific residue in one turn appears to stabilize the native structure and be responsible for slow refolding. *Protein Sci* **6**, 364–372 (1997).

77. Chattopadhyay, K., Zhong, S., Yeh, S. R., Rousseau, D. L. & Frieden, C. The intestinal fatty acid binding protein: The role of turns in fast and slow folding processes. *Biochemistry* **41**, 4040–4047 (2002).

78. Yeh, S., Ropson, I. J. & Rousseau, D. L. Hierarchical Folding of Intestinal Fatty Acid Binding Protein. *Biochemistry* **40**, 4205–4210 (2001).

Chapter 5. Effects of an engineered cavity on the structure of T4 lysozyme revealed by pressure perturbation

Lysozyme from the T4 bacteriophage (T4L) is a model protein used for biophysical structural studies. To study the role of internal cavities in protein structure, the L99A cavity-enlarging mutation of T4L was developed, which increases the volume of a native cavity from 40 $Å^3$ to 160 $Å^3$.[1,2] While L99A decreases the packing density of the C-domain, it does not disrupt the native fold, forming a unique example in which loss of van der Waals contacts is not compensated by structural "collapse" or cavity hydration.[3] Because the volume properties of a protein guide its response to pressure, structural changes in T4L wild-type (WT*, which contains C54T/C97A mutations to produce a cysteineless background) and L99A as a function of pressure are investigated in this chapter by site-directed spin labeling electron paramagnetic resonance (SDSL EPR). In particular, this chapter evaluates whether high pressure results in simple compression or populates higher-energy conformational states and how SDSL EPR can distinguish compressibility from conformational exchange. To date, pressure perturbation of T4L L99A has led to different interpretations of the T4L structure at high pressure, including unfolded, highly flexible, or MG-like conformations.

5.1 Background to T4L WT* and L99A

The hydrophobic core of globular proteins are generally well-packed.[4,5] Large-to-small mutations of hydrophobic amino acids in the interior of a protein can create new cavities, which have different effects on protein structure based on the local geometry of the protein and the possibility for a protein to relax to fill the new void. The first foray into the role of cavity-creating mutations in the hydrophobic core of the α-helical protein T4 lysozyme identified a

mutation, L99A, that formed a new cavity *unaccompanied* by any structural change in the X-ray crystal.[6] This large cavity has a volume of 160 $Å^3$ and increases the size of a preexisting cavity in the WT* protein.[2,7] Other engineered cavities with similar volumes were found to induce structural relaxation. Alanine mutagenesis of the remaining nonpolar side chains buried in the core of T4L were consistent with a model for the destabilizing effect of cavity creation: first, that a minimum loss of energy upon amino acid substitution corresponds to the differing free energies of transfers for the native and alanine residues from aqueous to oil phase, and second, that ~22 cal mol^{-1} $Å^{-3}$ is lost due to reduction of van der Waals contacts, which scales with the size of the resulting cavity.[8]

Of the X-ray crystal structures resulting from this survey of interior cavities, only the WT*, I100A, and L99A T4L structures are virtually identical in the crystalline state.[8] However, T4L L99A has received particular interest as a gain-of-function mutation due to its ability to bind benzene within its enlarged cavity[9], independent of the native hydrolytic activity of T4L.[10] Benzene binding is accompanied by only minor structural adjustments, but no indication of the conformational rearrangement that must occur to permit benzene access to this cavity is apparent in the X-ray crystal structure.[9]

Identification of protein dynamics in T4L L99A by solution NMR of side chain methyl ^2H spin relaxation rates and side chain methyl ^{13}C and ^1H spin relaxation revealed motions on the μs-ms timescale in the vicinity of the cavity.[11] These motions increase upon removal of ligand[11] (in this case, xenon), in parallel with findings for other hydrophobic ligand-binding proteins, such as I-FABP.[12] Despite evidence of slow motions on a timescale corresponding to conformational exchange[13], the increased flexibility in L99A was ascribed only to subtle changes in structure of the closed conformational state because the open conformation is rare under the

experimental conditions.[3] In support of this model, backbone dynamics on the ps-ns timescale derived from ^{15}N relaxation were unchanged in L99A compared to WT*[11], suggesting that the cavity in L99A is apparently static and does not involve rapid, small-amplitude fluctuations in structure.

To resolve the minor conformational state indicated by only a single set of crosspeaks in ^{1}H-^{15}N correlation spectra, relaxation-dispersion NMR revealed a rare (~3%) conformational state of T4L L99A[14], distinguished by two-state exchange for resonances of side chains proximal to the cavity.[15] A structure for this excited state of L99A (designated "E") was modeled from relaxation-dispersion NMR data and CS-Rosetta, constructing a new conformation unrelated to the cavity-accessible conformation; instead, F114 in the new structure occupies the cavity as a benzene mimetic, which distorts helices F and G in order to accommodate the new rotamer of the F114 side chain.[16]

A second approach to study the conformational landscape of T4L L99A incorporates high hydrostatic pressure. As described in Chapter 3, application of pressure on proteins shifts conformational equilibria towards states with reduced partial molar volumes. In contrast with structure-relaxation observed by solution NMR, by which the "E" state eliminates the cavity by reorienting F114 inwards, pressure populates a conformation of T4L L99A that eliminates the cavity by hydration (designated "G_H"). Five lines of evidence support a pressure-populated hydrated-cavity conformation rather than the structure-relaxed state: (1) high-pressure X-ray crystallography[17], (2) high-pressure intrinsic tryptophan fluorescence[7], (3) high-pressure of backbone amide ^{15}N heteronuclear single quantum coherence (HSQC) of T4L L99A encapsulated in a reverse micelle[18], (4) high-pressure side chain methyl ^{1}H-^{13}C spin relaxation[19], and (5) pressure-resolved double electron-electron resonance (PR DEER)[20]. The main

conclusions of each of these studies are summarized below:

(1) Although the L99A cavity is bordered by hydrophobic amino acid side chains, applied pressure up to 2000 bar is shown to cooperatively fill the cavity with solvent.[17] Electron density corresponding to four water molecules were identified in the high-pressure X-ray crystal structure, and molecular dynamics simulations support pressure-dependent entry of water into the L99A cavity. Cavity hydration of T4L L99A in the crystal lattice is unaccompanied by any sign of unfolding or structural collapse. Fig. 5.1 illustrates the stark contrast between the large-amplitude movement of helix F and F114 in the "E" state, shown in blue, to the conformation of the L99A ground state "G," shown in grey, and the high-pressure "G_H" state, shown in pink.

(2) Internal polarity of the L99A cavity can be monitored by intrinsic fluorescence of W138 in helix I (unpublished results, Kevin Eden), which forms part of the cavity surface. Observed red shifts in the fluorescence spectra of L99A (containing W126, W138, and W158) with increasing pressure suggest increasing hydration of the protein interior.[7] The transitions demonstrate two-state behavior, consistent with equilibrium populations between "G" and a high-energy state that is populated with pressure. Fluorescence quenching by incorporation of selenomethionine demonstrated similar compactness at high pressure to the ground state from sustained quenching efficiency with increasing pressure, supporting the formation of a folded, compact "G_H" intermediate.[7] Although the transitions were generally characterized as populating a variety of compact denatured states rather than "G_H," this study supports cavity hydration as the predominant response of T4L L99A to pressure.

(3) Pressure-induced cavity hydration shown directly by the appearance of new cross-peaks at the ^{1}H resonance of water indicate that solvent occupancy of the L99A cavity increases with

pressure.[18] To improve sensitivity of NMR nuclear Overhauser effect (NOE)-derived distances to internal waters rather than to bulk solvent, T4L L99A was encapsulated in reverse micelles prior to pressurization. Encapsulation is thought to stabilize compact states of proteins, as well.[21] However, like in the crystal lattice, suppression of conformational exchange is an important consideration, and was considered by Wand and Nucci to artificially affect the pressure sensitivity of T4L L99A.[22]

(4) Methyl side chain resonances in the vicinity of the L99A cavity from ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) spectra are sensitive to atomic packing. Signal intensities exhibited significant line-broadening with increasing pressure up to 3000 bar, indicating that the high-pressure state is made up of heterogeneous conformations in exchange on a timescale faster than ms, particularly within the L99A cavity.[19] The absence of new peaks suggests that the fluctuating structure is "G"-like. This study is unable to distinguish whether the "G_H" or "E" states are populated with pressure, though the authors remark that the data do not conflict with a model in which pressure populates the "E" state.[19]

(5) To resolve discrepancies in the interpretation of the global structure of T4L L99A at high pressure, Lerch *et al*. employ PR DEER to discriminate between unfolded states, "G_H", and "E" by measuring pairwise distance distributions between nitroxides placed on different helices at varying pressures.[20] Because movement of helix F is characteristic for the "E" state, nitroxides pairs between helices D-F and F-I in WT* and L99A were monitored by PR DEER. Even though the dataset is sparse, the distance distributions at 3000 bar compared to atmospheric pressure show only subtle shifts in the relative populations of individual distances in L99A. The "E" state is accompanied by large-amplitude movement of helix F[6,16], which was identified in an L99A construct containing the additional mutations G113A and

R119P, which stabilize the "E" state by lowering its free energy.[16] In comparison, WT* distance distributions between helices D-F and F-I at 4000 bar are essentially identical to those at atmospheric pressure and closely resemble the distributions of L99A. From this lack of conformational rearrangement, it was rationalized that the strain energy of helix F deformation must be higher than the energy of hydration in L99A, making cavity hydration in L99A the preferred response to pressure. An equivalent rationale for populating the "G_H" state at 3000 bar rather than the "E" state is described by Maeno *et al*. as differences in the exchange rates between conformational states; in this case, increasing exchange between "G" and "G_H" could be favored over exchange between "G" and "E" states, though the mechanistic interpretation by Lerch *et al*. is more structurally specific.

In this chapter, the pressure dependence of T4L WT* and L99A is investigated by SDSL EPR. Placement of a nitroxide on the solvent-exposed surface of almost all of the helices in T4L

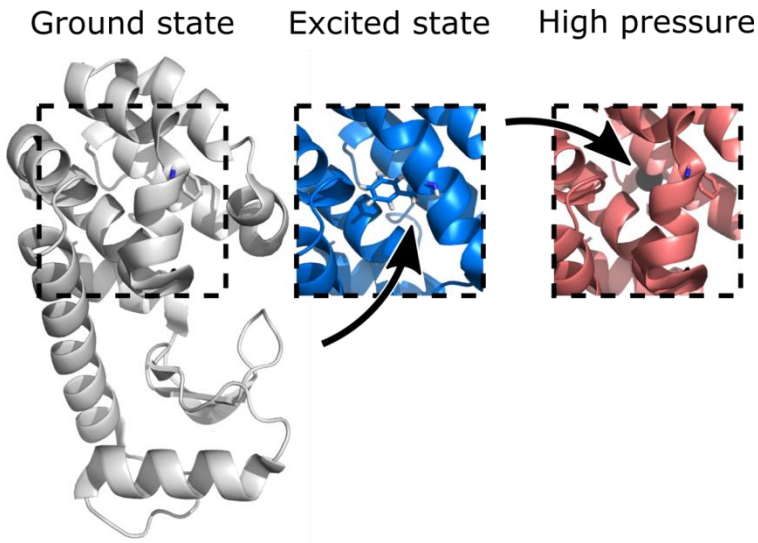

Figure 5.1 Position of F114 in the ground state, excited state, and high pressure states of T4L L99A (PDB accession codes: 3DMV, 2LCB, and 2B6W, respectively). Compared to the WT* ground state, a large rearrangement of helix F and a change in the rotamer of F114 positions the phenyl group into the hyrophobic core in the "E" state conformation are indicated by the arrow in the center example. At high pressure, electron density increases in the vicinity of the arrow in the righthand example, suggesting cavity hydration.

(helices B, C, D, E, F, G, H, I, and J) produce a coarse structural map of site-specific structure and dynamics. One aim of this work is to determine whether compressibility can be distinguished from conformational exchange using SDSL EPR. As will be shown, the response of a nitroxide attached to a protein can be categorized into three classes: (i) site-dependent changes in the internal nitroxide side chain dynamics, (ii) shifts in rotameric equilibria of the nitroxide side chain, and (iii) shifts in protein conformational equilibria. The third response is of particular biological interest in light of the number of high pressure studies on the "G_H" high-pressure conformation of T4L L99A. The results presented here reveal site-specific changes in the population of spectral components of T4 lysozyme (T4L) as a function of pressure. At high pressure, a new spectral component appears in the line shape of many sites. The pressure dependence of this component reveal its site-specific volume properties. Pairing hydrostatic pressure techniques with SDSL EPR yields a highly sensitive method by which to investigate a multiplicity of protein states. In this report, SDSL EPR and line shape analysis are used to show site-specific compressibility in T4L WT* and to distinguish compressibility from conformational exchange in T4L L99A.

5.2 SDSL of T4L WT* and L99A at atmospheric pressure

Fourteen sites were chosen to survey the structure of T4L in the WT* and L99A backgrounds: S44R1, K65R1, D72R1, R76R1, R80R1, A82R1, D89R1, A93R1, T109R1, R119R1, E128R1, V131R1, N140R1, and T151R1 (Fig. 5.2). Each position is solvent-facing and the set is intended to monitor helices B-J in the C-domain of T4L, which contains a small native cavity near L99 in the WT* background and the enlarged cavity L99A in the mutant background. These sites are designed to be noninteracting helix surface sites that do not disrupt the local fold

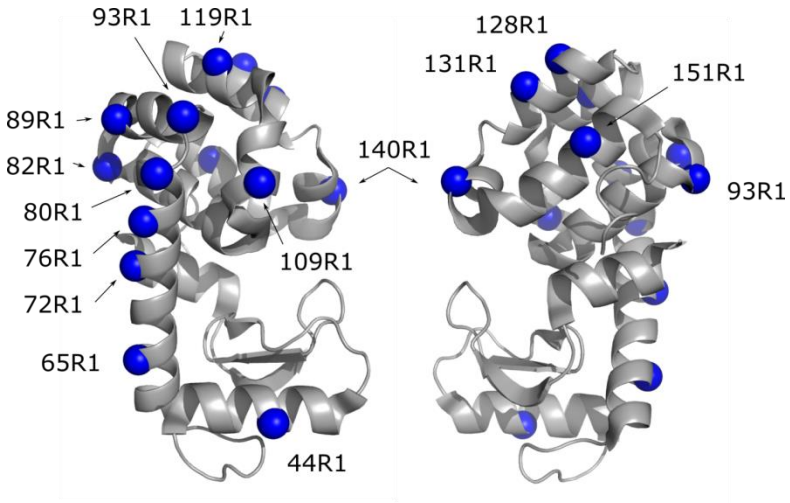

Figure 5.2 Location of nitroxides in T4L WT* and L99A in this work. The two structures are shown rotated by 180°.

of T4L and have nitroxide line shapes typical for noninteracting helix surface sites. CW line shapes at atmospheric pressure in 25% Ficoll-70, which is added to decrease the rotational diffusion of T4L, have been published elsewhere[20,23]; similar spectra in 30% sucrose are also published elsewhere.[24–27] CW spectra for the 14 nitroxide sites at ambient pressure are shown in Fig. 5.3 on the left, from which the ceramic signal has been subtracted.

In some cases, the spectra are multicomponent. The origin of the second component is either due to rotameric or conformational exchange.[28] In the case of 44R1, the spectrum is multicomponent and the origin of the immobile component is due to a local interaction with E45, shown by X-ray crystallography of spin labeled T4L and supported by line shape analysis paired with local mutagenesis.[26] Another site, 109R1, is in the short F-helix of T4L, which is thought to be in exchange in the L99A mutant between the "G" and "E" states.[15] It is multicomponent, but the immobile and mobile components are relatively mobile compared to other complex spectra in T4L. 119R1 is in the G-helix of T4L. X-ray crystallography of 119R1 shows two rotamers for this nitroxide, supporting that the two components are due to side chain rotamerization. While

one rotamer is noninteracting, the second rotamer displaces the native side chain N123.[24] The multicomponent site 128R1 is in helix H; its mobile component is predominant and a small immobile component at atmospheric pressure is observed. 93R1 is a new site designed for this study. Its line shape is relatively complex, possibly due to a local contact with the protein. Modeling a nitroxide at A93 into the crystal structure of T4L suggests that the side chain avoids local clashes with the surrounding structure.

For nearly all sites, the line shapes are nearly superimposable, indicating that the local structure of T4L is the same in both WT* and L99A backgrounds, consistent with the prediction from X-ray crystal structures of WT* and L99A. The sole exception to this trend is 93R1, which has a larger immobile component in L99A than in WT*. The local environment of the N-terminal residue of helix E (which contains L99 and L99A) is monitored by 93R1. Saturation recovery of 93R1 in the L99A background at ambient pressure is monoexponential (unpublished results, Michael Bridges), suggesting that the L99A mutation influences the equilibrium between rotameric states of 93R1 on a time scale faster than μs.

Osmolyte perturbation has been shown to identify conformational exchange from rotameric exchange by virtue of the insensitivity of rotamers to stabilizing osmolytes such as sucrose in comparison to conformational states, which are sensitive to osmolytes.[23] In the case of conformational exchange, sucrose will stabilize the immobile component of a two-component line shape. The line shape of 93R1 in WT* and L99A backgrounds are superimposable in 30% sucrose and in 25% Ficoll-70, which are matched in macroscopic viscosity.[23] This result, in combination with the SR result, indicates that the origin for the multicomponent spectra of 93R1 in both backgrounds is due to rotamer exchange. A small difference in the line shape of the mobile component of 93R1 suggests a small-amplitude structural difference between WT* and

L99A. Helix E, containing 93R1, makes a large number of contacts with helices D, F, G, and I in both WT* and L99A structures, which are nearly indistinguishable at atmospheric pressure. The sensitivity of the SDSL EPR line shape is capable of capturing such subtle shifts in local protein structure.

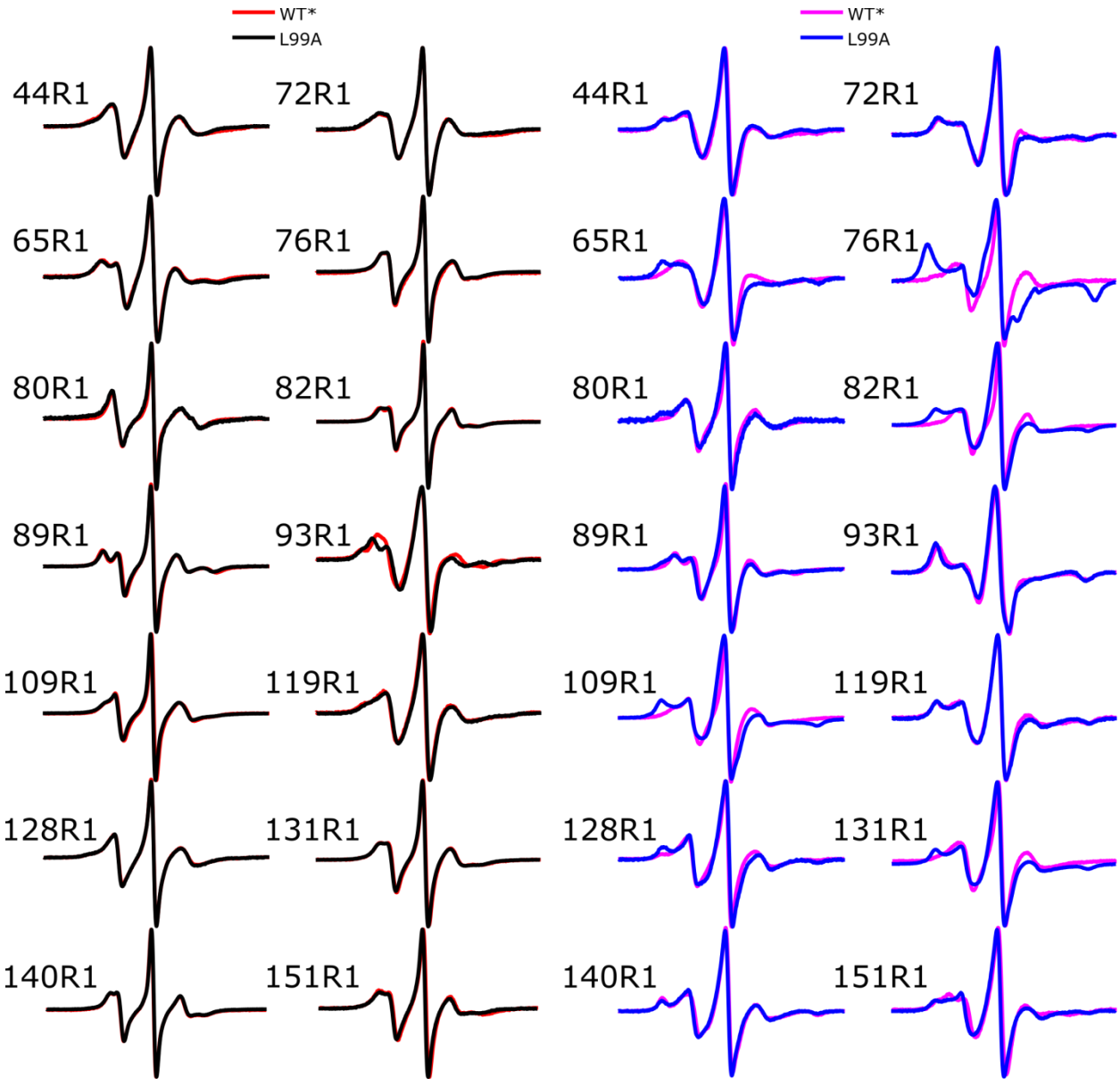

Figure 5.3 Spectra line shapes of 14 nitroxides in T4L WT* and L99A backgrounds. On the left, atmospheric pressure spectra are shown. On the right, spectra at 3000 bar are shown.

5.3 Line shapes of WT* and L99A T4L at high pressure

In contrast to the similar local structure indicated by nearly identical CW line shapes at ambient pressure, CW line shapes in WT* and L99A backgrounds show significant spectral differences at 3000 bar. In particular, 76R1, 80R1, 82R1, 109R1, 128R1, 131R1, and 150R1 have immobile components with larger fractional populations in L99A than in WT* (Fig. 5.3, right).

65R1 is an exception: at atmospheric pressure and 3000 bar, 65R1 WT* can be simulated adequately by a single component, but at 3000 bar, the line shape in the L99A background has significantly slower motion. The X-ray crystal of 65R1 supports the population of one rotamer at atmospheric pressure.[24]

All nitroxide line shapes in the L99A background except for 65R1 are at least two-component at 3000 bar, though the fractional population of the immobile component varies significantly. For example, the component populations of 89R1 L99A are 10% immobile and 90% mobile at 3000 bar, while for 82R1 L99A, they are 72% immobile and 28% mobile at 3000 bar.

Three-component spectra at high pressures are apparent for two sites: 109R1 and 151R1 in the L99A background. At 3000 bar, 109R1 L99A is best fit to a three-component system in which the new immobile component that appears at high pressure is distinct and has slower

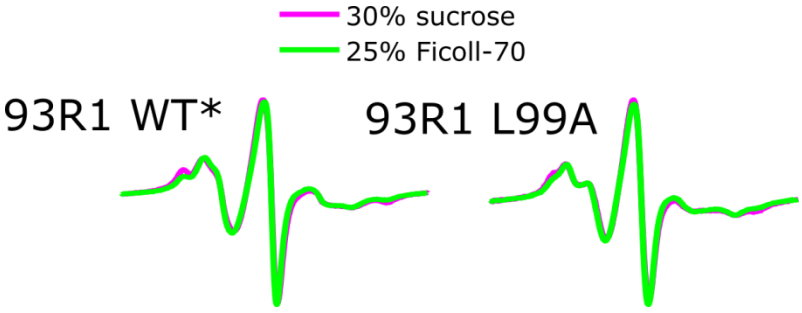

Figure 5.4 Osmolyte perturbation of 93R1 in WT* and L99A backgrounds of T4L. Macroscopic viscosity-matched conditions show a lack of an osmolyte effect on the nitroxide line shape of 93R1, suggesting the origin of the two spectral components is due to rotameric exchange.

motion than the immobile component at ambient pressure.

The following sections describe two methods by which the pressure dependence of CW EPR line shapes are analyzed for the variable pressure dataset shown in Fig. 5.5. Note that some variable pressure CW spectra have been previously published. In T4L WT* in helix D, the pressure dependence of the line shape of 82R1 revealed simple anisotropic motion constrained by the "X4/X5" model expected of a surface helix site is associated with a small ΔV^0 = -1.2 ml mol^{-1} that represents predominately internal R1 motion.[29] Also in T4L WT*, in the loop between α-helices C and D, 80R1 is a second example of a single-component line shape that fits a model in which the correlation time changes but the order remains constant. The change in partial molar volume is larger (ΔV^0 = -4.2 ml mol^{-1}), indicating that the backbone motion of the short loop segment factors into the partial molar volume.[29] In helix B of WT* T4L, pressure dependence of 44R1 demonstrates rotameric exchange and near-neighbor mutagenesis shows that the populations of these spectral components are sensitive to the strength of local contacts with side chains of varying sizes.[25] The shift between spectral components of 44R1 shows no change in compressibility between the two states, and the rotamer that interacts with nearby glutamate is favored, as it has a smaller partial molar volume (ΔV^0 = -9.4 ml mol^{-1})[29]. In both WT* and L99A backgrounds, the line shape of 72R1 in helix C at 0 bar and 2000 bar show an increase in correlation time with pressure broadens the CW spectra. In L99A, a new spectral component appears at high pressure.[20] No new component is present in the high pressure spectra of 89R1 in helix D of T4L in WT* and L99A backgrounds.[20] Lastly, in helix F of the L99A background, 109R1 a new (third) component appears at high pressure. Helix F contains F114 and its dynamics is therefore of particular interest.[20,30]

5.4 Line shape analysis of variable pressure EPR spectra

The pressure dependence of spectral populations indicates the mechanism by which they are populated. Among the 14 sites investigated here, three classes of behavior are revealed: (i) changes in the internal motion of the nitroxide, (ii) shifts in spectral populations from rotameric equilibria, and (iii) shifts in spectral populations from conformational equilibria. The first class is apparent in all line shapes, whether they are single or multicomponent: all spectra show increasing line broadening as pressure increases. Distinguishing changes that belong to the class ii and iii are of particular interest in this work.

The mechanism of the first class, modulation of the internal motion of the nitroxide, is thought to be due to a positive activation volume for fluctuation of the solvent cage around the nitroxide.[29]

Classes ii and iii are both identified by changing spectral component populations with pressure. Linear changes in the spectra as a function of pressure indicate that no structural transition occurs and that the experiment monitors simple compression within a single conformational state. As described earlier, a conformational state is made up of a number of statistical substates. Compressibility can be thought of as shifting populations towards those with smaller partial molar volumes. For small differences in the partial molar volumes between states ($\Delta\bar{V}^0 < 20$ ml mol^{-1}), pressure population of the new state may appear as a flattened sigmoid.

Larger differences in the partial molar volumes between states produce nonlinear pressure dependence in the population of the states, suggesting more clearly the existence of additional conformational states with energies close to the native state.[31] To quantify changes in the populations of spectral states as unchanging (class i), linearly changing (class ii), or nonlinearly changing (class iii) with pressure, two methods were used: singular value

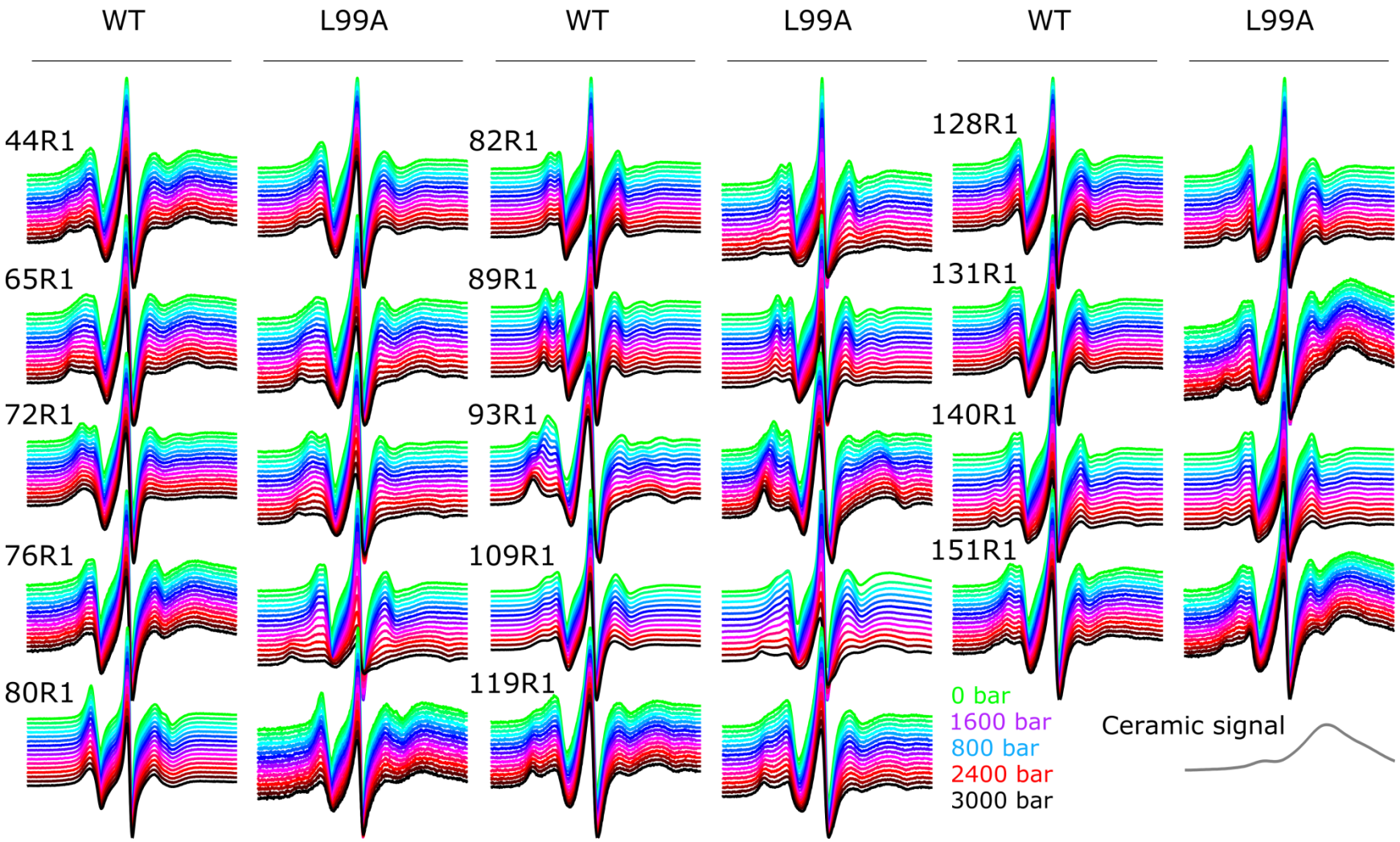

Figure 5.5 Variable pressure CW EPR spectra with the ceramic signal intact (see the variable-amplitude broad signal in the high field line), demonstrating the gradations in line shape broadening and changing spectral component populations for various sites labeled in WT* and L99A backgrounds of T4 lysozyme. At the bottom right, the isolated EPR line shape of the ceramic signal from the pressure cell is shown.

decomposition (SVD), a model-free algorithm, and spectral simulation, a model-based means to quantify the spectral line shapes. SVD and spectral simulations of multicomponent spectra provide approximate and relative populations of spectral states, respectively.

5.5 Singular value decomposition of CW EPR spectra

Singular value decomposition (SVD) is a linear algebra algorithm that factorizes a matrix M into the product of three matrices in the form of $M = USV^T$. This method benefits from being a global analysis of the pressure spectra and is model-free. Additionally, it is rapid and does not require the complex multivariable fitting procedure required for spectral simulation that can take upwards of hours to complete. SVD decomposes a data set into principal spectral components (matrix U), component amplitudes (matrix V), and the component weights (matrix S), where the number of components is user-defined. Applications of SVD to EPR have been made to identify "basis spectra," or spectral components, among datasets for line shapes that change linearly.[32,33] The SVD factorization is used in principal component analysis to evaluate the significance of each component. In variable pressure CW data, all line shapes have been found to broaden with pressure, reflecting the pressure-dependent correlation times of the nitroxide side chain motion.[29] The extent of line broadening beyond broadening due to internan side chan motion is site-specific. Line broadening in the first-derivative CW EPR spectrum is not a linear change in the line shape and therefore does not fit the SVD model well. SVD of simulated spectra are shown in Fig. 5.5. The simulated spectra represent an arbitrary two-component system in which an immobile component is populated and a mobile component is depopulated. The transition is modeled to a sigmoid and is sampled at 6 pressures. In the upper left panel of Fig. 5.5a, spectral line shapes are simulated without line broadening. In upper left panel of Fig. 5.5b, a modest

amount of line broadening accumulates linearly in each component in agreement with published findings.[29]

Difference spectra from the mean of the dataset are used to emphasize changes in the line shapes; the original spectra are shown in the inset. By considering the difference spectra, the SVD amplitudes are centered at a component amplitude of 0 about the average of the observed spectral transition. The first SVD component is assigned to the change in the spectral line shape, rather than a "basis spectrum."

The first SVD component describes the difference spectra well, both in the non-broadenend and broadened cases, shown by the weighted magnitude of the first component "basis spectra" (blue curves) in the upper right panels of Fig. 5.5a and Fig. 5.5b. The basis spectra of the reconstructed M matrix are the result of computing $U \times S$. The effect of line broadening is shown clearly in Fig. 5.5b by comparing the first component amplitude (the first column of the $V \times S$ matrix) to the populations in the simulated model. Deviation from the simulated model is also shown in the second component amplitude, which appears as a parabolic shape. By considering additional SVD components, it is possible to more closely replicate the populations of the model, but because the differences between broadened spectra are not exclusively additive, use of SVD is not ideal. However, as an initial test for the nonlinearity in the change of spectral populations, SVD is capable of identifying nonlinear transitions from linear transitions and is a good method for screening data without the effort of spectral simulation.

Shown in Figs. 5.8 and 5.9 are the SVD first component amplitudes which are scaled to integrated area-normalized atmospheric pressure spectra. Discrepancies in the low pressure range were found to be correlated with nonlinear sensitivity in the low current range of the

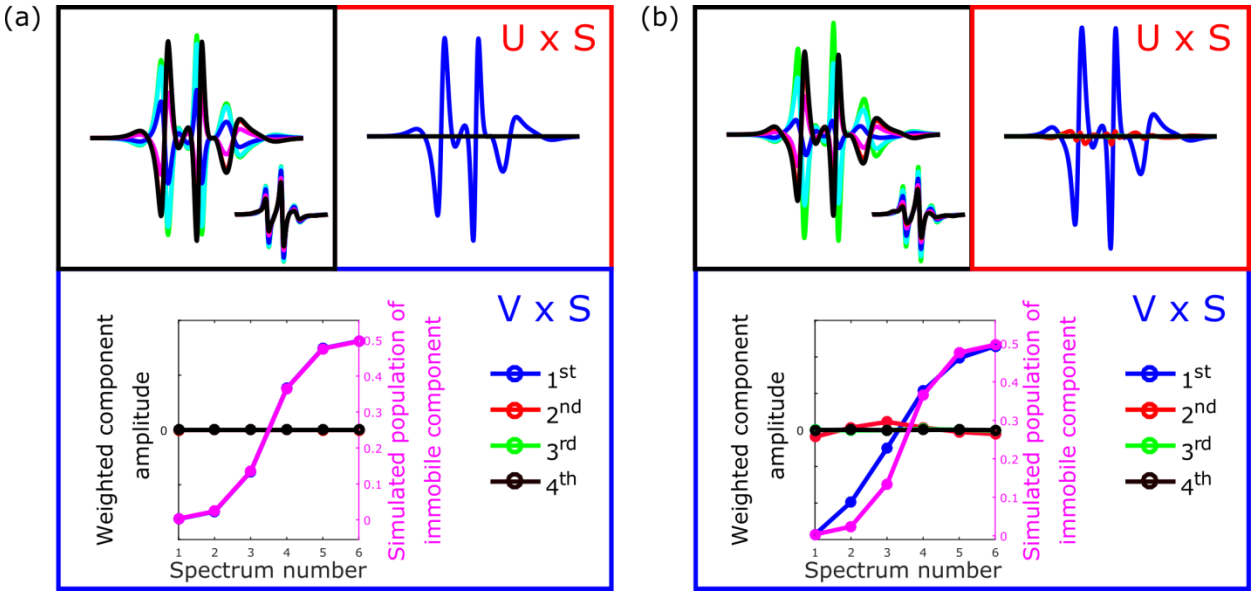

Figure 5.6 Demonstration of the effect of line broadening on the singular value decomposition results. (a) 6 two-component spectra are simulated with varying populations to model conformational exchange. (b) The same two-component spectra are modified by decreasing correlation times and the resulting nonlinear transformation is seen as higher-order SVD components in the analysis.

spectrometer detector. The detector current was adjusted to allow acquisition of variable pressure spectra without re-tuning the cavity during the course of the experiment to keep the measurements within the range of the detector. The small increase in water dielectric as a function of pressure is thought to be the cause of an increase in detector current with pressure. The most reliable SVD amplitudes are therefore above ~500 bar within the linear range of the detector. Analysis of the pressure dependence of the SVD first component amplitude therefore requires determining which transitions are sigmoidal by the shape of the pressure dependence at pressures exceeding 500 bar. For example, T4L 65R1 in the WT* and L99A backgrounds should be regarded as linear, while the shape of the curve for 76R1 in the L99A background is sigmoidal. Verification that the line shape does not change at low pressures but that the signal intensity does explains aberrations in the SVD amplitudes.

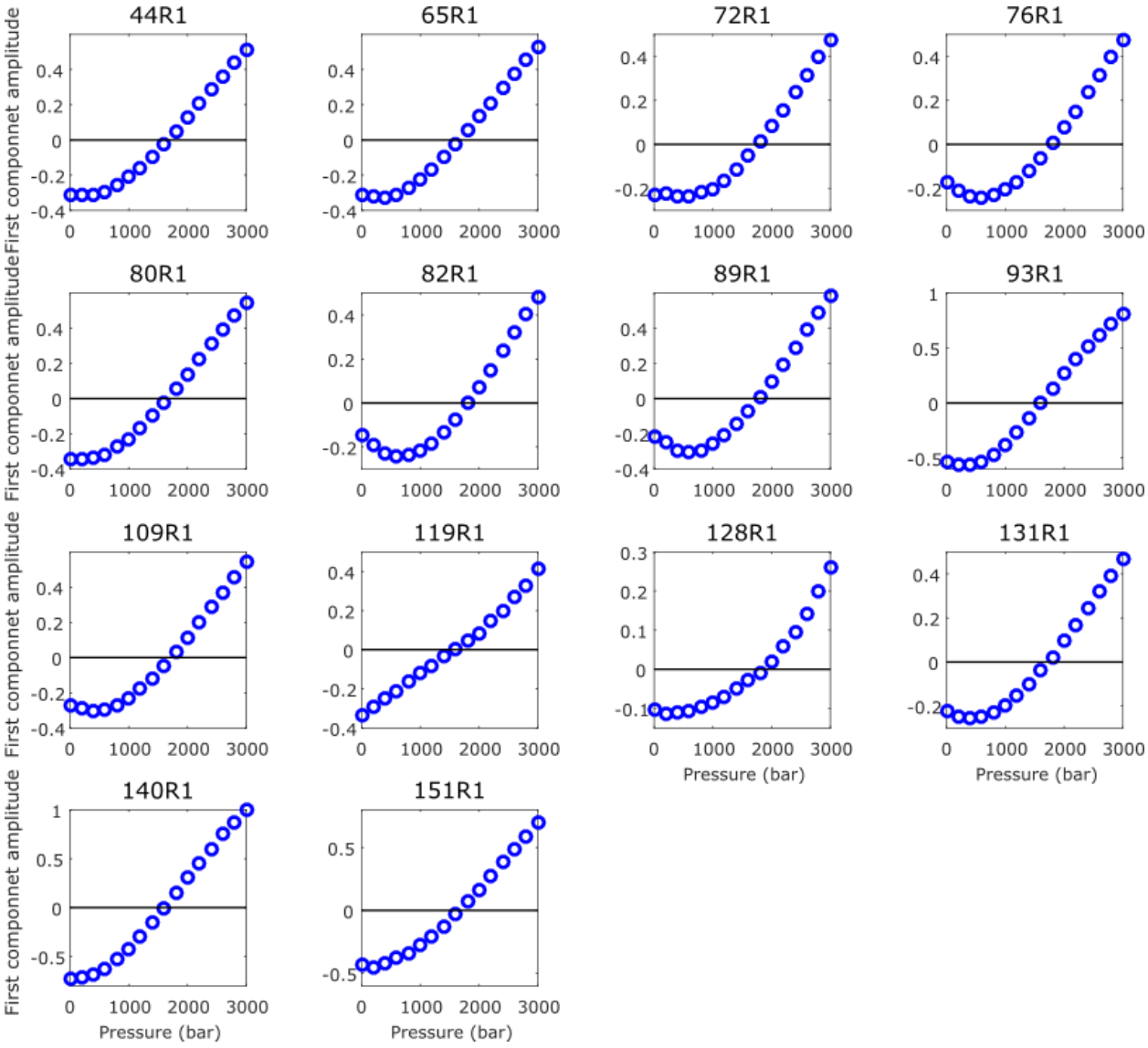

Figure 5.7 First component amplitudes as a function of pressure for WT* T4 lysozyme. The curvature of the SVD first component amplitude in the pressure regime for which line shape changes occur (between 1000 – 3000 bar) is approximately linear.

5.6 Spectral simulation of variable pressure EPR spectra

Quantification of spectral components by simulation imposes a model in order to extract population information directly. Before the spectra can be simulated, the intrinsic signal from the ceramic pressure cell must be subtracted (see Fig. 5.5 for the line shapes with ceramic background and corresponding spectra in Fig. 5.9 after subtraction). This procedure introduces

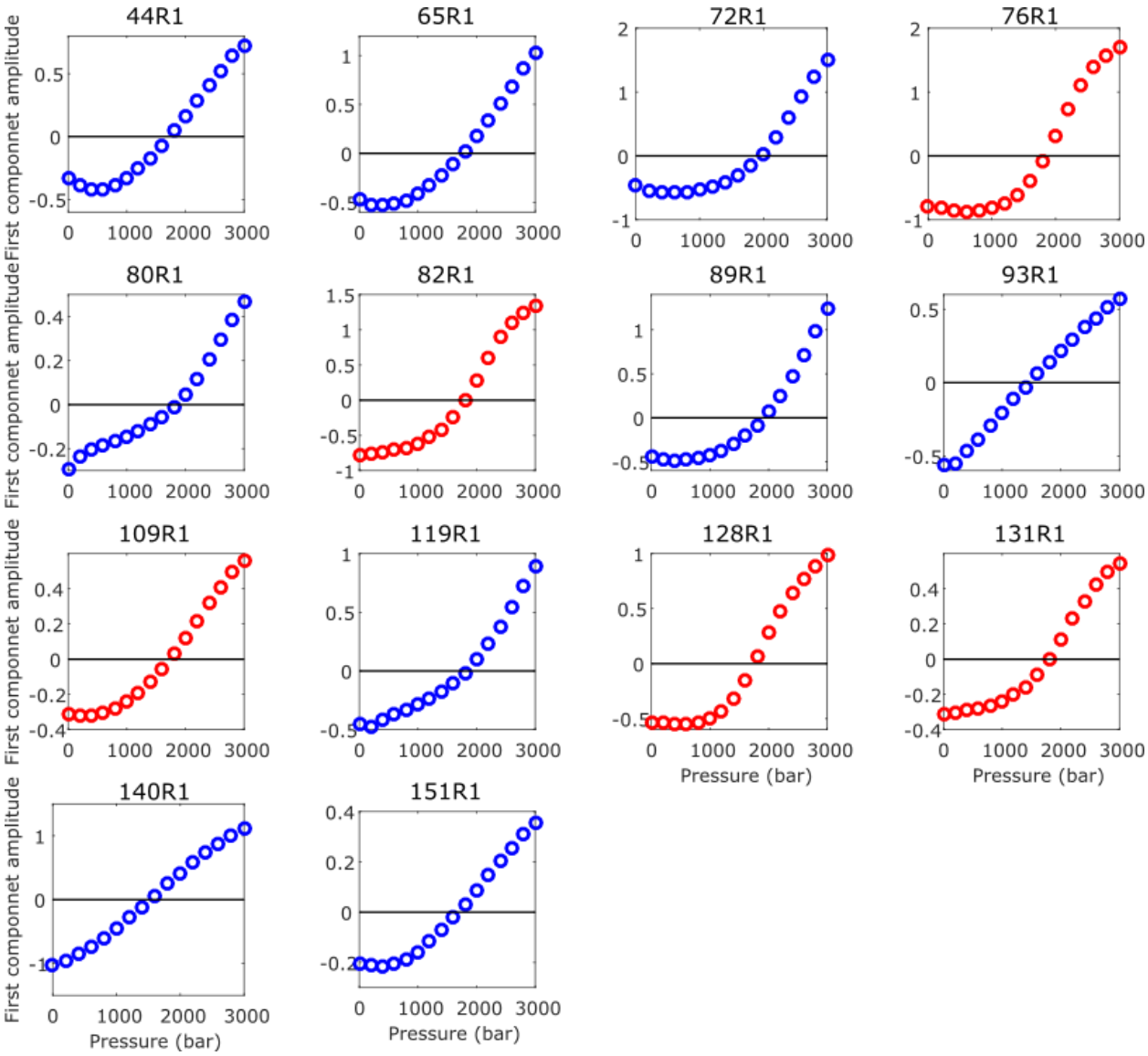

Figure 5.8 First component amplitudes as a function of pressure for the L99A cavity mutant of T4 lysozyme. The curvature of the SVD first component amplitude in the pressure regime for which line shape changes occur (between 1000 – 3000 bar) is approximately linear for the majority of sites; however, a subset of sites show nonlinearity at higher pressures (2000 – 3000 bar).

errors into the CW spectrum that affects the line shape and baselines of the resulting spectra. A fitting procedure was written in MATLAB to reduce the human error from manual subtraction and is found in the Appendix.

To minimize the number of variables in simulating the line shapes of a nitroxide at

different pressures, it was found that a MOMD model for the mobile component with a fixed order parameter and varying isotropic rates produced good fits at all pressures.[29] The immobile component was generally fit to an isotropic line shape with a slow correlation time. For high pressure spectra with small populations of the immobile component, the isotropic correlation time was fixed in order to reduce the number of variables. The fit parameters used here do not necessarily correspond to the highest-quality fit at any pressure; rather, they were determined iteratively to provide good fits to all spectra in the pressure series.

Spectral simulations for 13 sites in WT* T4L and 12 sites in T4L L99A are shown in Figs. 5.9 in steps of 1000 bar to demonstrate the quality of the fits. Fit parameters from these fits are tabluated in the Appendix (see Tables A5.1 and A5.2) and the component populations at varying pressures are shown in Figs. 5.10 and 5.11.

5.6.1 Linear transitions in the population of spectral components

In WT* T4L, all changes in line shape have either first (i) or second (ii) class pressure dependence (see Section 5.1). Class i sites are single-component and show an increase in line broadening with pressure. These sites include 65R1, 72R1, 80R1, 89R1, 109R1, and 119R1. Class ii are multicomponent with component populations which change linearly with pressure. These sites include 44R1, 76R1, 82R1, 131R1, 140R1, and 151R1. Because T4L WT* has few, small cavities, it is expected that pressure effects would be limited to simple compression, rather than population of a new conformational state.

5.6.2 Sigmoidal transitions in the population of spectral components

Despite differing interpretations of the high pressure state of T4L L99A, described above

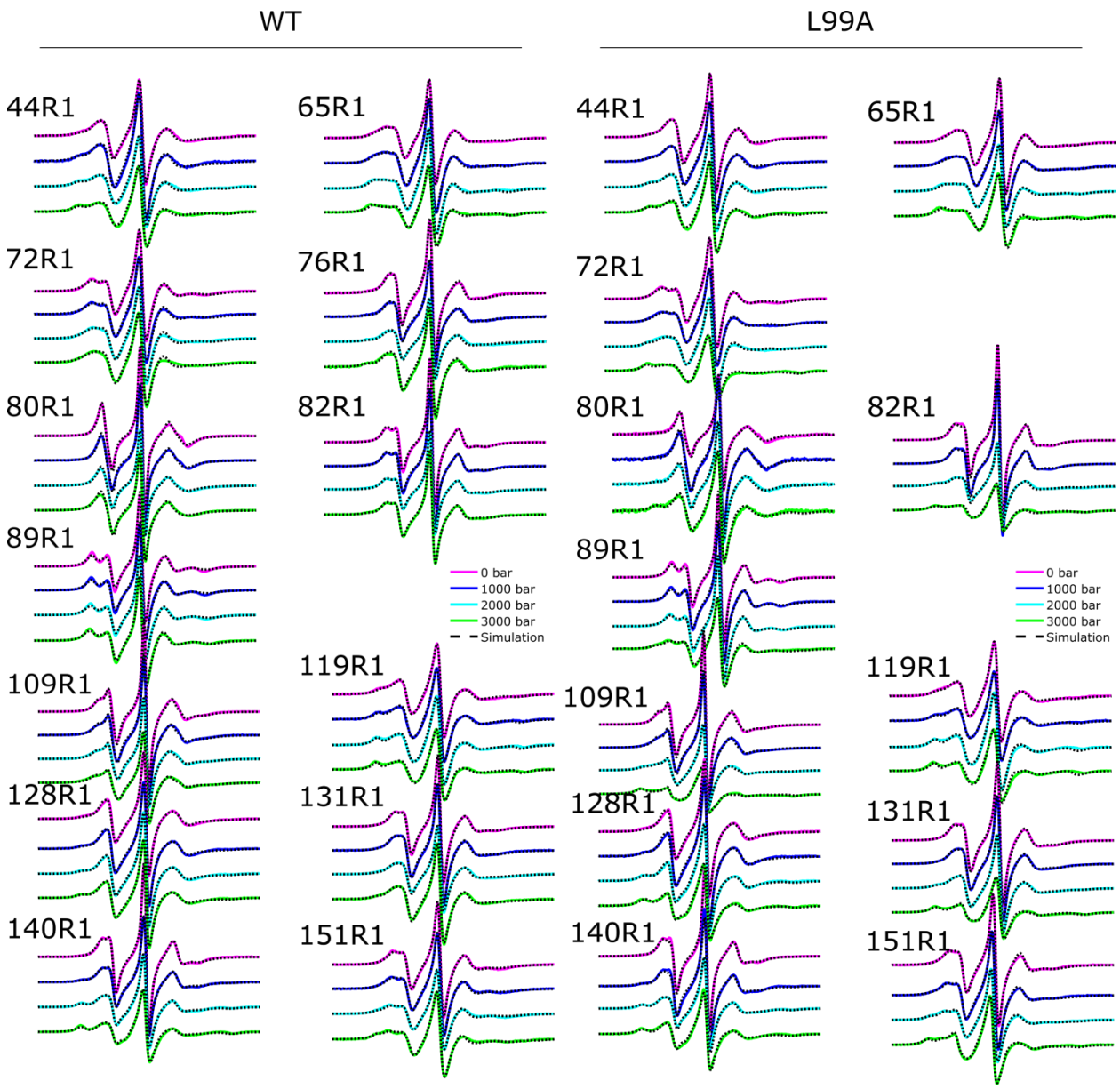

Figure 5.9 Spectral simulations of T4 lysozyme in WT* and L99A backgrounds in 25% Ficoll-70. Spectra were acquired in 200 bar steps. For simplicity, a subset of each dataset is shown in 1000 bar steps ranging from 0-3000 bar.

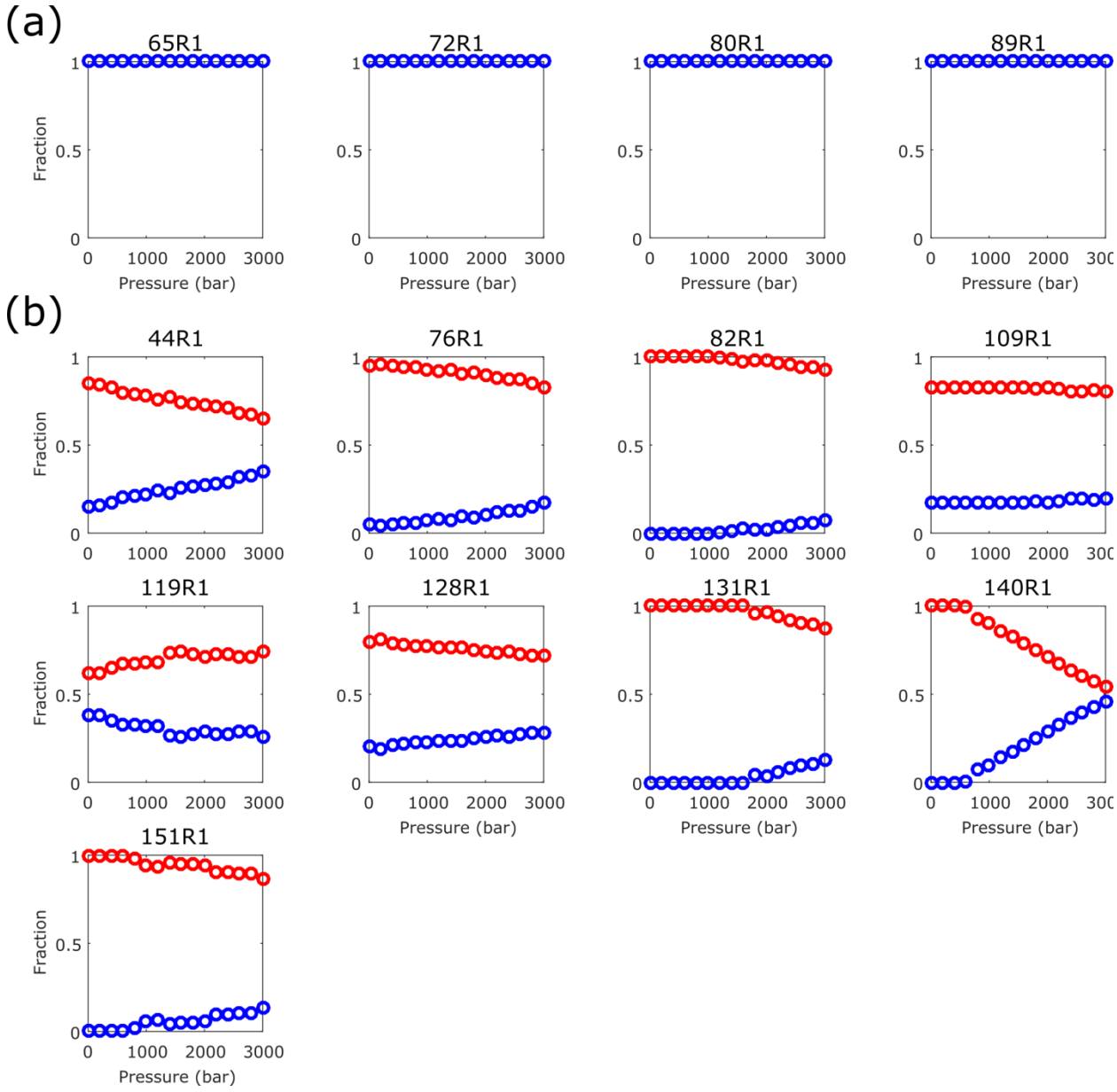

Figure 5.10 Fractional populations of CW EPR line shapes of WT* T4 lysozyme. In (a), class i pressure dependency is exemplified by a single spectral component at low to moderate (3000 bar) pressures. In (b), class ii pressure dependency of multicomponent spectra are linear, suggesting either that the line shape changes reflect intrinsic compressibility or that the line shape changes are “shallow” sigmoids which appear linear in this pressure range.

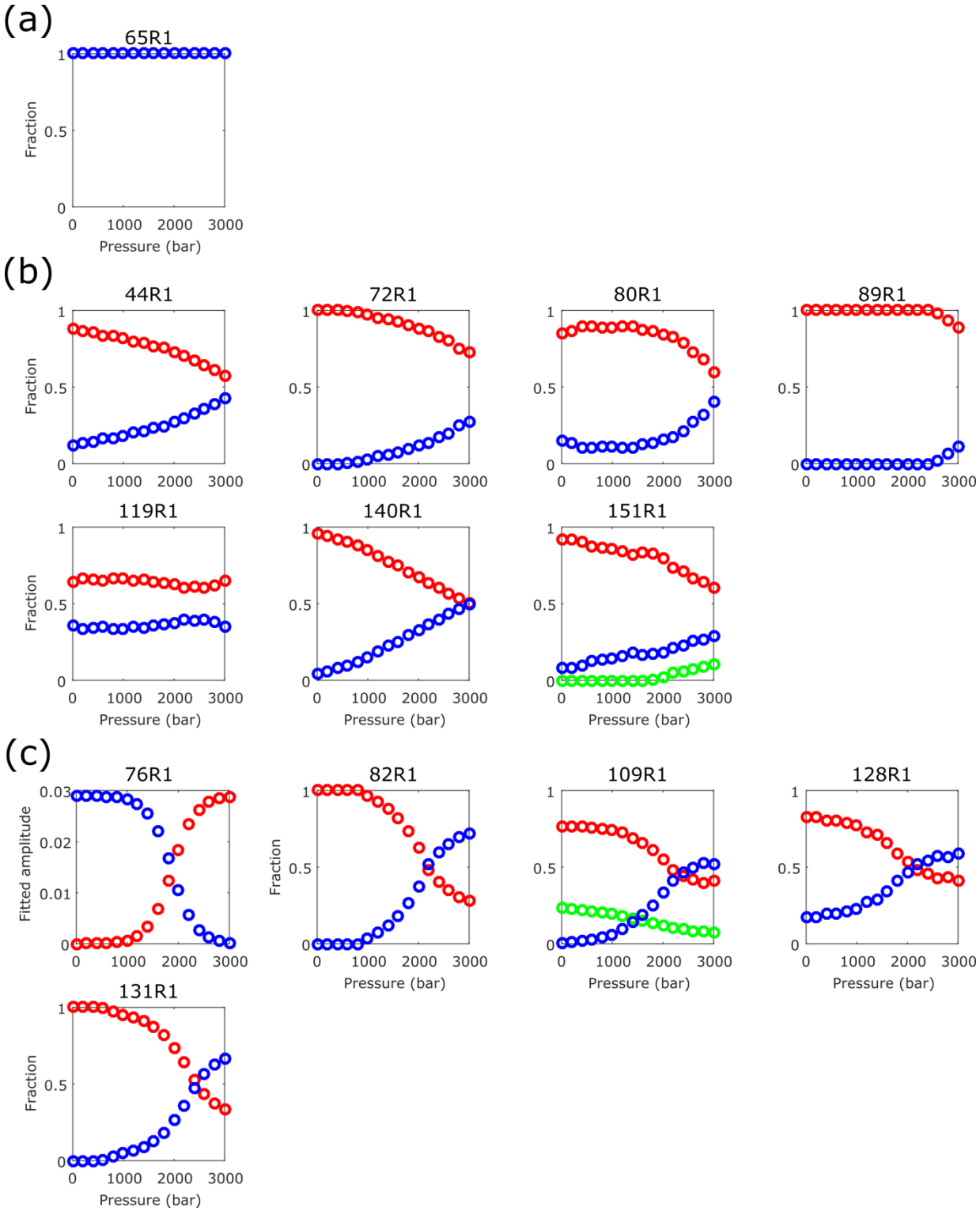

Figure 5.11 Fractional populations of CW EPR line shapes of the L99A cavity mutant of T4 lysozyme. As in Fig. 5.10, (a) shows class i pressure dependency and (b) shows class ii pressure dependency. In (c), class iii pressure dependency appears as a sigmoid, suggesting a two-state transition.

in Section 5.1, the cavity-hydrated state "G_H" is thought to have a conformation nearly identical to the native state, but with an enlarged cavity filled with upwards of 4 water molecules at 3000 bar.[17]

In L99A, only 65R1 has a class i pressure response. The line shape changes with pressure in the remaining 13 sites can be accounted for by mobility changes in the side chain and/or local helix fluctuations. Class ii sites include 44R1, 140R1, and 151R1. These sites are all class ii in the WT* background, suggesting that the pressure dependence is the same in WT* and L99A backgrounds and presumably have the same rotameric origin. Nonlinear responses include 72R1, 80R1, and 89R1 but cannot be fit to a sigmoid because the data do not reach, at minimum, the transition midpoint. In other words, it is not possible to determine from these spectra whether a second state with a large free energy difference and small molar volume change is populated at sufficiently high pressures. Lastly, sigmoidal transitions are observed from the populations determined by spectral simulation for five sites in L99A: 76R1, 82R1, 109R1, 128R1, and 131R1 (Fig. 5.12). A two-state model can be fit to the component populations of the form:

$$f_i = (J - I)\frac{1}{1+\exp(\Delta G_0/RT - P\cdot\Delta\bar{V}^0/RT)} + I \qquad (1)$$

where f_i is the fractional population of the immobile state, J and I are the extrapolated equilibrium populations at positive and negative infinite pressures, respectively, and ΔG_0 and $\Delta\bar{V}^0$ are the differences in free energy and partial molar volume of the two states, respectively.

Site-specific differences reflect heterogeneity of T4L L99A sensitivity to pressure, which is due to volume differences between the ground and excited states of different helices in T4L. *This analysis provides the first evidence for a structural change in the "G_H" state that accompanies cavity hydration*. Table 5.1 summarizes the thermodynamic free energies and volume changes from the sigmoidal fits to the population changes of 76R1, 82R1, 109R1,

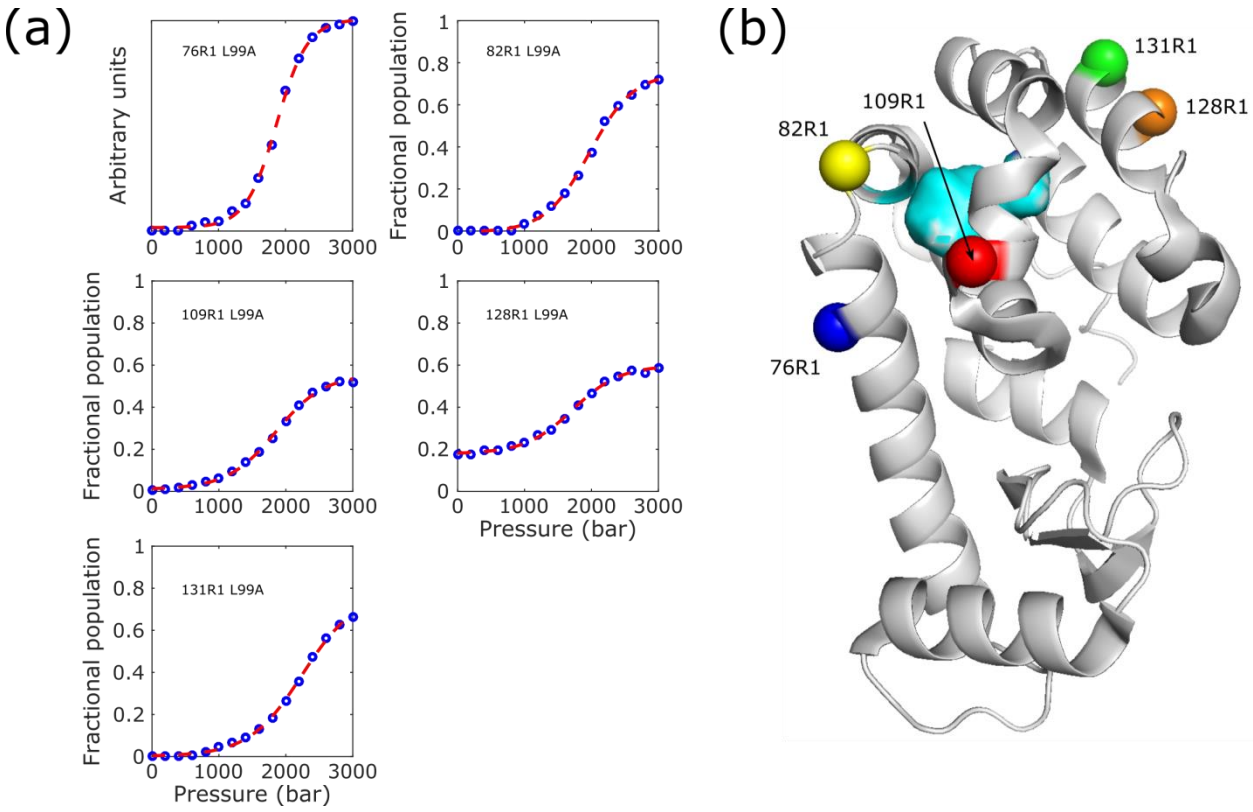

Figure 5.12 In (a), sigmoidal fits to the fractional populations of the new spectral component determined by spectral simulation. The fit function assumes the change in isothermal compressibility between 0 bar and 3 kbar states is negligible. In (b), the $\boldsymbol{C_\alpha}$ atoms of the nitroxide sites in T4L L99A are indicated by spheres, colored from red to blue with increasing difference in the partial molar volume between the native and high pressure states. The surface of the L99A cavity is shown in cyan.

128R1, and 131R1. These values are in good agreement with the ΔG_0 and $\Delta \bar{V}^0$ values for the differences in hydration of the "G" to "G_H" states determined from high-pressure tryptophan fluorescence, which were found to be 2.5 kcal mol^{-1} and -75 ml mol^{-1}, respectively[7], though the fluorescence data reflect changes in solvation rather than changes in protein structure, which is directly monitored by SDSL EPR.

Table 5.1 Thermodynamic parameters for sites in T4 lysozyme L99A

Site	ΔG_0	$\pm$ 95% CI*	$\Delta \bar{V}^0$	$\pm$ 95% CI
76R1	-5.1 kcal mol^{-1}	(-5.1, -5.8) kcal mol^{-1}	96.3 ml mol^{-1}	(81.0, 111.5) ml mol^{-1}
82R1	-3.2 kcal mol^{-1}	(-3.6, -2.7) kcal mol^{-1}	68.5 ml mol^{-1}	(58.0, 79.0) ml mol^{-1}
109R1	-2.6 kcal mol^{-1}	(-2.8, -2.4) kcal mol^{-1}	34.8 ml mol^{-1}	(26.8, 42.8) ml mol^{-1}
128R1	-3.1 kcal mol^{-1}	(-3.6, -2.7) kcal mol^{-1}	81.1 ml mol^{-1}	(69.7, 92.5) ml mol^{-1}
131R1	-3.3 kcal mol^{-1}	(-3.6, -3.0) kcal mol^{-1}	62.0 ml mol^{-1}	(54.9, 69.1) ml mol^{-1}

* The 95% confidence interval (CI) is shown here in order to demonstrate the quality of the sigmoid fit to the populations determined by spectral simulation.

5.7 Pressure-jump of T4L L99A reveals slow kinetics

To investigate the kinetics of the pressure response in T4L L99A, a coarse pressure-jump experiment was performed on each of the L99A mutants. After a 3 second jump from 0 to 3000 bar, the low field line was monitored in 10 second intervals. In the low field region of the CW spectrum, the components are well-separated and easily visible, while in the center field region the two components are intermixed and in the high field region the first derivative amplitude is small. Two sites showed slow equilibration in the spectral line shape, measurable on the sec-min timescale: 72R1 and 76R1. The remaining sites all equilibrated within the dead time of the jump from 0 to 3000 bar. In the reverse direction from 3000 to 0 bar, all sites equilibrated to the ground state within the dead time of the jump. SVD was used to identify the major change in the spectral line shape of 72R1 and 76R1, from which the first component amplitude was fit to an exponential to determine the observed rate constant in the upward-jump direction, k_{up}. These plots are shown in Fig. 5.14. Rates determined from this pressure jump experiment reflect population of the new spectral component from a pre-compressed "G" state, formed during the course of the jump.

At atmospheric pressure, the line shape of 72R1 is single-component. Immediately following the jump, a new immobile component is present, though its population is small (Fig.

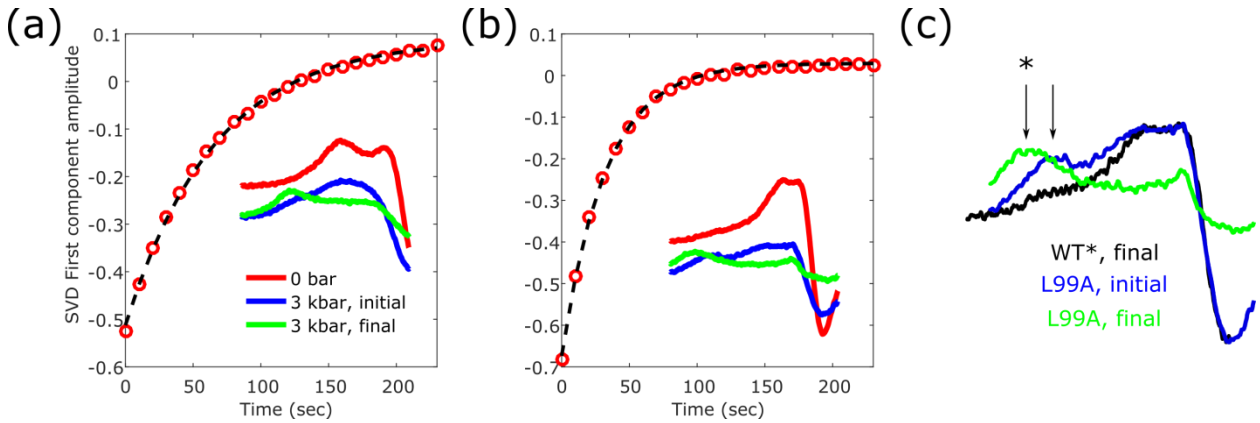

Figure 5.13 SVD first component amplitudes determined from the low field line (25 G) of 72R1 in (a) and 76R1 in (b) following a 3-second pressure jump. The initial spectra are shown in the inset in red. Below, the initial spectrum following the jump is shown in blue. The final spectrum is shown in green after a period of >4 min, which corresponds to the line shape obtained in the variable pressure datasets at equilibrium. In (c), the low field lines of 76R1 in WT* and L99A backgrounds are shown before and after a 3 second pressure jump from 0 to 3000 bar. In black, the WT* lineshape at 3000 bar does not vary with time, suggesting the transition occurred during the dead time of the jump. In blue, the line shape of 76R1 in the L99A background immediately after the jump resembles the line shape of 76R1 in the WT* background, except the population of the immobile component is greater. Over the course of ~4 minutes, the line shape of 76R1 equilibrates to a final high-pressure state. The arrows indicate the peak positions for the immobile component. The asterisk designates the final shift in the peak position of the immobile component of 76R1 in L99A.

5.13a, inset). The mobile component is significantly broadened at 3000 bar, indicating that the mechanism of line broadening occurs faster than the jump time (3 sec). The population of the immobile component grows exponentially with time. The SVD first component amplitude captures the transition well with an exponential fit to produce a rate constant $k_{up} = 0.018$ sec^{-1}.

At equilibrium, the pressure dependence of spectral populations of 76R1 show strong sigmoidal behavior, measured from the appearance of an immobile component that appears at pressures above 1000 bar. As the system equilibrates, the immobile component shifts down field. This transition has no isosbestic points, suggesting that an intermediate state is populated. However, because the SVD first component amplitude is fit well to a single exponential, the two

states are not well-resolved among the time course of this jump. Pressure-jump of 76R1 in the WT* background showed population of the immobile component occurs during the 3 sec jump. Interestingly, the position of the immobile component in L99A immediately after the jump aligns well with the position of the immobile component in WT* (Fig. 5.13c). This similarity suggests that in L99A, the initial high-pressure state resembles the compressed high-pressure state of WT*. The transition between the compressed states immediately following the jump to the equilibrium state has a rate constant $k_{up} = 0.031$ sec^{-1}. Although the rate of populating this new state cannot be determined for WT*, the fractional population of the new state is higher in L99A than in WT*. Fig. 5.13c illustrates the initial similarity.

Directional asymmetry in the upward and downward jumps is expected for processes that are path-dependent, which suggests that the transition states for G to G_H and G_H to G are different, though with the current dataset this conclusion is speculative. By executing pressure jump experiments with shorter dead times, activation volumes of the upward and downward transitions could be compared to show that the transition states may be similar to either the ground or the high-pressure states, elucidating a mechanism for water entry into the L99A cavity.

5.8 Discussion: The high pressure state of T4L WT* and L99A

Protein motions can be characterized by timescales, in which slow motions are generally large-amplitude conformational exchange events (μs-ms) and fast motions are small-amplitude fluctuations and rotameric exchange (ps-ns). Fast exchange broadening on the intrinsic EPR timescale (ps-ns at X-band) appears as a function of pressure in the data produced here, which has been modeled by simulation as an increase in the correlation time to represent components of internal nitroxide motion and fast protein motions that are sensitive to pressure.[29] Interpreting the

origin of multicomponent spectra that reflect exchange slow on the CW EPR timescale (>100 ns at X-band) requires complementary EPR experiments, including osmolyte perturbation and saturation recovery EPR. Distinguishing these slow conformational processes from rotameric exchange has been a major focus of SDSL EPR in recent years to isolate protein-specific motions rather than internal rotameric motions of the nitroxide side chain.[23,28,34] The present data show that CW EPR can distinguish between compressibility and conformational exchange by the pressure dependence of multicomponent spectra.

The compressibility of a protein conformational state is defined as the change in the volume as a function of pressure, which is a direct result of the forces that stabilize the state.[35] Compression behavior of T4L WT* and L99A is not uniform across throughout the protein presumably because packing imperfections are distributed unequally throughout the protein interior. It is found that the native cavities in WT* are compressible but do not contribute to population of a new conformational state at pressures up to 3000 bar. The arrangement of native cavities in the X-ray crystal structure of T4L WT* is identical to those in L99A.[7] Either the L99A cavity alone or dynamic cavities not described by X-ray crystallography in the L99A background may play a role in the population of a new conformational state at high pressure. The present work identifies 5 spin labeled sites on α-helix C, the C-D loop, α-helix F, and α-helix H (76R1, 82R1, 109R1, and 128R1 and 131R1, respectively) which monitor conformational exchange as a function of pressure in T4L L99A.

The nature of the conformational change in T4L L99A at high pressure is not clear. While the G_H state of T4L L99A is thought to have a native-like conformation[17], water passage into the hydrophobic L99A cavity (the hallmark of the G_H state) requires at least small-amplitude movements of the helical bundle that makes up the N-terminal domain. This rearrangement could

proceed through a new conformation or a molten-globule like state in which tertiary contacts are lost which would permit hydration of internal cavities. In a similar vein, binding of small hydrophobic molecules such as benzene into the L99A cavity could also proceed through similar solvent channels. Molecular dynamics simulation of ligand binding shows putative structural rearrangements that permit benzene binding are small in amplitude and involve helices D, F, G, and I.[36] This simulation also proposes multiple binding pathways that could be related to fluctuations that permit water access into the hydrophobic cavity. Three paths of entry are identified: the first between helices D and F, the second between helices G and I, and the third, between helices E, F, G, and H. From the present analysis, the sites showing sigmoidal (and therefore conformational) transitions are: 76R1 in helix C; 82R1 in helix D, 109R1 in helix F, and 128R1 and 131R1 in helix H. These sites show approximate agreement with conformational motion corresponding to entry points 1 and 3, suggesting that the heterogeneous pressure responses in different helices of T4L L99A can correlate with predicted points of entry of ligands and possibly solvent to the L99A cavity.

The position and volumes of cavities in the atmospheric pressure crystal structures of WT* and L99A T4L are very similar and therefore do not aid in predicting unique compressible regions of T4L in the WT* or L99A backgrounds. While native packing imperfections can be found using a probe smaller than the radius of water (1.0 Å rather than 1.4 Å) to identify small cavities, nitroxides susceptible to compression at higher pressures are not colocalized, even though these cavities are found mostly in the C-domain of both WT* and L99A proteins. They include the native ligand-binding cavity, which is the solvent-exposed cleft between the N- and C-domains (~21,000 $Å^3$), two cavities hydrated by a single water molecule (Sol 175 and Sol 208), a cavity hydrated by two water molecules (Sol 171 and 179), a hydrophobic cavity between

helices H and J (~27 Å^3), and a small hydrophobic cavity between helices E and I (~8 Å^3). The L99A mutation enlarges a hydrophobic cavity of ~40 Å^3 in WT* T4L to ~160 Å^3; the only significantly different cavity of the ambient-pressure crystal structures.[7] The conformational state that is populated at elevated pressure in T4L L99A is associated with a change in volume that requires a model more complex than elimination of unassociated native cavities. The flexibility of globular proteins may reorganize packing imperfections in a manner that X-ray crystal structures cannot depict.

It may also be important to distinguish cavities that are occupied by solvent. For example, the Sol 208 molecule that occupies a cavity in WT* and L99A backgrounds is highly conserved across many X-ray crystal structures of T4L.[7] Is this cavity eliminated by hydration at atmospheric pressure or can it be eliminated by structure-relaxation at elevated pressure? A nitroxide at 140R1 is proximal to this cavity and demonstrates class ii pressure dependence in both WT* and L99A backgrounds (i.e. reflects compressibility). Helix I is short and has loose packing with helices H and J; hence the compressibility observed at 140R1 may be due to compression of the Sol 208-occupied cavity. In comparison, 44R1 in helix B is another highly compressible site. 44R1 is distal to the L99A cavity and lies along the three-stranded β-sheet in the N-terminal domain of T4L (Fig. 5.2). The interaction surface between helix B and the β-sheet is hydrophobic and tightly packed. As discussed, the multicomponent spectrum of 44R1 is due to rotameric exchange at ambient pressure. Pressure therefore populates the rotameric state with lower partial molar volume, i.e., that which interacts E45. Because the interaction is intrahelical, the compressibility of this site is most likely unrelated to repacking of helix B with the β-sheet.

In all, direct measurement of local protein compressibility is shown to be a strength of variable pressure EPR. The high sensitivity of SDSL EPR to local structure highlights changes in

the protein that would be inaccessible to larger probes. Additionally, variable pressure EPR can distinguish between compressibility and conformational exchange, particularly by the shape of the spectral population vs. pressure curve. This work reinforces the assertion that modest pressures (< 4000 bar) do not unfold proteins[20], but shift protein populations that are already present to states that are "rare" or "invisible" at ambient pressure.

5.9 Bibliography

1. Nicholson, H., Anderson, D. E., Dao-pin, S. & Matthews, B. W. Analysis of the Interaction between Charged Side Chains and the α-Helix Dipole Using Designed Thermostable Mutants of Phage T4 Lysozyme. *Biochemistry* **30**, 9816–9828 (1991).
2. Eriksson, A. E., Baase, W. A. & Matthews, B. W. Similar Hydrophobic Replacements of Leu99 and Phe153 within the Core of T4 Lysozyme Have Different Structural and Thermodynamic Consequences. *J. Mol. Biol.* **229**, 747–769 (1993).
3. Baase, W. A., Liu, L., Tronrud, D. E. & Matthews, B. W. Lessons from the lysozyme of phage T4. *Protein Sci.* **19**, 631–641 (2010).
4. Richards, F. M. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176 (1977).
5. Connolly, M. L. Atomic size packing defects in proteins. *Int. J. Pept. Protein Res.* **28**, 360–363 (1986).
6. Eriksson, A. *et al.* Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science (80-.).* **255**, 178–183 (2006).
7. Ando, N. *et al.* Structural and Thermodynamic Characterization of T4 Lysozyme Mutants and the Contribution of Internal Cavities to Pressure Denaturation. *Biochemistry* **47**,

11097–11109 (2008).

8. Xu, J., Baase, W. A., Baldwin, E. & Matthews, B. W. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* **7**, 158–177 (1998).

9. Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* **355**, 371–373 (1992).

10. Goto, N. K., Skrynnikov, N. R., Dahlquist, F. W. & Kay, L. E. What is the average conformation of bacteriophage T4 lysozyme in solution? A domain orientation study using dipolar couplings measured by solution NMR. *J. Mol. Biol.* **308**, 745–764 (2001).

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

12. Hodsdon, M. E. & Cistola, D. P. Ligand Binding Alters the Backbone Mobility of Intestinal Fatty Acid-Binding Protein as Monitored by 15N NMR Relaxation and 1H Exchange. *Biochemistry* **36**, 2278–2290 (1997).

13. Palmer, A. G., Kroenke, C. D. & Loria, J. P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. in *Methods in Enzymology* **339**, 204–238 (Elsevier Masson SAS, 2001).

14. Hansen, D. F., Vallurupalli, P. & Kay, L. E. Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J. Biomol. NMR* **41**, 113–120 (2008).

15. Mulder, F. A. A., Mittermaier, A., Hon, B., Dahlquist, F. W. & Kay, L. E. Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 932–935 (2001).

16. Bouvignies G, Vallurupalli P, Hansen DF, Correia BE, Lange O, Bah A, Vernon RM, Dahlquist FW, Baker D, K. LE. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111–114 (2013).

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

18. Nucci, N. V, Fuglestad, B., Athanasoula, E. A. & Wand, A. J. Role of cavities and hydration in the pressure unfolding of T4 lysozyme. *Proc. Natl. Acad. Sci.* **111**, 13846–13851 (2014).

19. Maeno, A. *et al.* Cavity as a source of conformational fluctuation and high-energy state: High-pressure NMR study of a cavity-enlarged mutant of T4lysozyme. *Biophys. J.* **108**, 133–145 (2015).

20. Lerch, M. T. *et al.* Structure-relaxation mechanism for the response of T4 lysozyme cavity mutants to hydrostatic pressure. *Proc. Natl. Acad. Sci.* **112**, E2437–E2446 (2015).

21. Zhou, H. X. & Dill, K. A. Stabilization of proteins in confined spaces. *Biochemistry* **40**, 11289–11293 (2001).

22. Wand, A. J. & Nucci, N. V. Reply to Kitahara and Mulder: An ensemble view of protein stability best explains pressure effects in a T 4 lysozyme cavity mutant . *Proc. Natl. Acad. Sci.* **112**, E924–E924 (2015).

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

24. Langen, R., Oh, K. J., Cascio, D. & Hubbell, W. L. Crystal structures of spin labeled T4

lysozyme mutants: Implications for the interpretation of EPR spectra in terms of structure. *Biochemistry* **39**, 8396–8405 (2000).

25. Guo, Z., Cascio, D., Hideg, K., Kálái, T. & Hubbell, W. L. 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 (2007).

26. Guo, Z., Cascio, D., Hideg, K. & Hubbell, W. L. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci.* **17**, 228–239 (2007).

27. Fleissner, M. R. *et al.* Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci.* **106**, 21637–21642 (2010).

28. Bridges, M., Hideg, K. & Hubbell, W. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* **37**, 363–390 (2010).

29. McCoy, J. & Hubbell, W. L. High-pressure EPR reveals conformational equilibria and volumetric properties of spin-labeled proteins. *Proc. Natl. Acad. Sci.* **108**, 1331–1336 (2011).

30. Lerch, M. T., Yang, Z., Altenbach, C. & Hubbell, W. L. *High-Pressure EPR and Site-Directed Spin Labeling for Mapping Molecular Flexibility in Proteins*. *Methods in Enzymology* **564**, (Elsevier Inc., 2015).

31. Kamatari, Y. O., Kitahara, R., Yamada, H., Yokoyama, S. & Akasaka, K. High-pressure NMR spectroscopy for characterizing folding intermediates and denatured states of proteins. *Methods* **34**, 133–143 (2004).

32. Keszler, A., Kalyanaraman, B. & Hogg, N. Comparative investigation of superoxide

trapping by cyclic nitrone spin traps: The use of singular value decomposition and multiple linear regression analysis. *Free Radic. Biol. Med.* **35**, 1149–1157 (2003).

33. Lauricella, R., Allouch, A., Roubaud, V., Bouteiller, J. C. & Tuccio, B. A new kinetic approach to the evaluation of rate constants for the spin trapping of superoxide/hydroperoxyl radical by nitrones in aqueous media. *Org. Biomol. Chem.* **2**, 1304–1309 (2004).

34. Columbus, L. & Hubbell, W. L. A new spin on protein dynamics. *Trends Biochem Sci* **27**, 288–295 (2002).

35. Lee, B. Calculation of volume fluctuation for globular protein models. *Proc Natl Acad Sci U S A* **80**, 622–626 (1983).

36. Mondal, J., Ahalawat, N., Pandit, S., Kay, L. E. & Vallurupalli, P. Atomic resolution mechanism of ligand binding to a solvent inaccessible cavity in T4 lysozyme. *PLoS Comput. Biol.* **14**, 1–20 (2018).

5.10 Appendix

A1. MATLAB scripts for ceramic background subtraction from variable-pressure CW line shapes in the ceramic cell

Part 1: Data formatting

```
% Interpolate a suitable background file for the data sweep width
% Alternatively, cwsw=160 & B0=1000 or cwsw=200 & B0=800
cwsw=100;
B0=1800;
span=2048*300/cwsw;
ceramic=interp1(cell1,1:(size(cell1)-1)/(span-1):size(cell1));

%% Fit ceramic (one-by-one)!
% B0=1000;
m10=1;
options = optimset('MaxFunEvals',50000);

for i=1:size(Zeroed,2)
    [params]=fminsearch(@ceramicsub8,[B0,m10],options,ceramic,
    Data(:,i));
    fitparam(1,:)=params;
    fitted(:,i)=getGlobalx;
end
```

Part 2: Fitting the ceramic signal

```
function [sse]=ceramicsub(Param,ceramic,CW)

% Independent fit parameters
B=Param(1);
m1=Param(2);
diff=[];

data=CW(:,1);
if size(data,1)==2048
    fit2=data;
else if size(data,1)==1024
        fit2=interp1(data,1:(1024-1)/(2048-1):1024)';
    end
end

x=[1:size(fit2,1)]';

% Shift the ceramic reference by B
ceramic_shift=interp1(ceramic,B:1:(2047+B))';
fit1=ceramic_shift-ceramic_shift(1);

% Integrate the data, ceramic reference
fit2_int=cumtrapz(fit2);
```

```
fit1_int=cumtrapz(fit1);

% Apply a first order fit (include in baseline correction later on?)
fit1_int=fit1_int+x.*m1;
m0=fit2_int(end,1)/fit1_int(end,1);
fit1_int=fit1_int.*m0;

% Get the "working" spectrum (the residual)
work=fit2_int-fit1_int;

% Subtract a parabola to correct the baseline
x_LH=200;
x_RH=200;
xdim=[1:x_LH,(2049-x_RH):2048]';
ydim=work(xdim);
fitparam=polyfit(xdim,ydim,2);
fitrecon=polyval(fitparam,x);

work=work-fitrecon;

% Assemble sse fit criteria
% Type 1: 1/2 integrated area needs to occur at the x-position of the
% spectrum center max (use corresponding y-value)
work_maxpkxpos=find(work(750:1250)==max(work(750:1250)))+750;
work_maxpk=max(work(750:1250));

work_intpkxpos=interp1(cumtrapz(work(750:1250)),[750:1250],trapz(work
(750:1250))/2);
work_intpk=interp1([750:1250],work(750:1250),work_intpkxpos);

% Type 2: Integrated spectrum is above the x-axis (accumulate y-values)
% pad=cumtrapz(work)<0;
pad=work<0;

k=1;
index=length(diff);
for i=1:length(work)
    if pad(i)==1
        diff(k+index)=work(i);
        k=k+1;
    end
end

sse=sum(diff(:).^2);

setGlobalx([work;pad])

end
```

Table A5.1 Fit parameters for spectral simulations: variable pressure WT* T4 lysozyme line shapes in 25% Ficoll-70

Pressure	Amp, 0	Amp, 1	Rate, τ_{iso} 0	Rate, N 0	Rate, τ_{iso} 1	Rate, N 1	Order, 0	β_D, 0	γ_D, 0	W
44R1 WT*										
0 bar	85%	15%	1.9 ns	-	8.7 ns	-	0.32	36.0	-	0.05
1 kbar	78%	22%	2.1 ns	-	11 ns	-	0.32	36.0	-	0.05
2 kbar	73%	27%	2.4 ns	-	11 ns	-	0.32	36.0	-	0.05
3 kbar	65%	35%	2.6 ns	-	11 ns	-	0.32	36.0	-	0.05
65R1 WT*										
0 bar	100%	3.0 ns	-0.25	0.83	-	-	0.44	43.7	-	0.05
1 kbar	100%	3.6 ns	-0.25	0.83	-	-	0.44	43.7	-	0.05
2 kbar	100%	4.4 ns	-0.25	0.83	-	-	0.44	43.7	-	0.05
3 kbar	100%	5.5 ns	-0.25	0.83	-	-	0.44	43.7	-	0.05
72R1 WT*										
0 bar	100%	-	2.1 ns	-0.01	-	-	0.47	37.4	-	0.05
1 kbar	100%	-	2.3 ns	-0.01	-	-	0.47	41.2	-	0.05
2 kbar	100%	-	2.3 ns	-0.01	-	-	0.47	45.0	-	0.05
3 kbar	100%	-	2.5 ns	-0.01	-	-	0.47	48.7	-	0.05
76R1 WT*										
0 bar	95%	5%	1.4 ns	1.3	13 ns	-	0.27	32.2	-	0.05
1 kbar	92%	8%	1.5 ns	1.2	13 ns	-	0.27	32.2	-	0.05
2 kbar	90%	10%	1.7 ns	1.2	13 ns	-	0.27	32.2	-	0.05
3 kbar	82%	18%	1.8 ns	1.1	13 ns	-	0.27	32.2	-	0.05
80R1 WT*										
0 bar	100%	-	0.94 ns	-0.36	-	-	0.34	90.0	44.1	0.05
1 kbar	100%	-	1.1 ns	-0.36	-	-	0.34	90.0	44.8	0.05
2 kbar	100%	-	1.3 ns	-0.36	-	-	0.34	90.0	45.7	0.05
3 kbar	100%	-	1.4 ns	-0.36	-	-	0.34	90.0	47.4	0.05
82R1 WT*										
0 bar	100%	0%	5.6 ns	-2.3	-	-	0.31	28.4	-	0.05
1 kbar	100%	0%	4.6 ns	-1.9	-	-	0.31	28.4	-	0.05
2 kbar	98%	2%	3.2 ns	-1.3	13 ns	-	0.31	28.4	-	0.05
3 kbar	93%	7%	2.6 ns	-0.93	13 ns	-	0.31	28.4	-	0.05
89R1 WT*										
0 bar	100%	0%	2.2 ns	-1	-	-	0.46	34.4	-	0.05
1 kbar	100%	0%	2.3 ns	-0.9	-	-	0.46	34.4	-	0.05
2 kbar	100%	0%	2.5 ns	-0.79	-	-	0.46	34.4	-	0.05
3 kbar	100%	7%	2.8 ns	-0.74	-	-	0.46	34.4	-	0.05

109R1 WT*										
0 bar	83%	17%	1.4	-0.21	0.98 ns	-	0.36	23.7	-	0.05
1 kbar	83%	17%	1.7	-0.28	1.1 ns	-	0.36	23.7	-	0.05
2 kbar	82%	18%	2.3	-0.46	1.2 ns	-	0.36	23.7	-	0.05
3 kbar	81%	19%	3.4	-0.77	1.3 ns	-	0.36	23.7	-	0.05
119R1 WT*										
0 bar	62%	38%	4.3 ns	1.3	3.5 ns	-	0.62	48.9	-	0.05
1 kbar	68%	32%	4.5 ns	1.2	6.2 ns	-	0.62	48.9	-	0.05
2 kbar	71%	29%	3.5 ns	0.89	9.3 ns	-	0.62	48.9	-	0.05
3 kbar	74%	26%	5.1 ns	1.1	12 ns	-	0.62	48.9	-	0.05
128R1 WT*										
0 bar	80%	20%	0.65 ns	-0.57	6.8 ns	-	0.6	50.9	-	0.05
1 kbar	77%	23%	0.81 ns	-0.28	7.8 ns	-	0.6	50.9	-	0.05
2 kbar	74%	26%	0.92 ns	-0.16	8.7 ns	-	0.6	50.9	-	0.05
3 kbar	72%	28%	1.1 ns	-0.06	10 ns	-	0.6	50.9	-	0.05
131R1 WT*										
0 bar	100%	0%	1.8 ns	-0.26	-	-	0.36	32.4	-	0.05
1 kbar	100%	0%	1.8 ns	-0.08	-	-	0.36	32.4	-	0.05
2 kbar	96%	4%	1.8 ns	0.07	11 ns	-	0.36	32.4	-	0.05
3 kbar	87%	13%	1.9 ns	0.18	12 ns	-	0.36	32.4	-	0.05
140R1 WT*										
0 bar	100%	0%	4.1 ns	1.8	-	-	0.43	43.6	-	0.05
1 kbar	93%	7%	3.9 ns	1.6	13 ns	-	0.43	43.6	-	0.05
2 kbar	71%	29%	3.8 ns	1.5	14 ns	-	0.43	43.6	-	0.05
3 kbar	54%	46%	3.4 ns	1.3	14 ns	-	0.43	43.6	-	0.05
151R1 WT*										
0 bar	99%	1%	1.9	-0.32	0.46 ns	-	0.39	36.9	-	0.05
1 kbar	94%	6%	2.2	-0.25	6.9 ns	-	0.39	36.9	-	0.05
2 kbar	94%	6%	2.5	-0.16	100 ns	-	0.39	36.9	-	0.05
3 kbar	86%	14%	2.9	-0.08	64 ns	-	0.39	36.9	-	0.05

Table A5.2 Fit parameters for spectral simulations: variable pressure T4 lysozyme L99A line shapes in 25% Ficoll-70

Pressure	Amp, 0	Amp, 1	Amp, 2	Rate, τ_{iso} 0	Rate, N 0	Rate, N_{xy} 0	Rate, τ_{iso} 1	Rate, N 1	Rate, τ_{iso} 2	Order, 0	β_D, 0	γ_D, 0	W
44R1 L99A													
0 bar	88%	12%	-	1.7 ns	-	-	8.9 ns	-	-	0.34	37.1	-	0.05
1 kbar	82%	18%	-	1.9 ns	-	-	9.6 ns	-	-	0.34	37.7	-	0.05
2 kbar	72%	28%	-	2.1 ns	-	-	9.5 ns	-	-	0.34	38.2	-	0.05
3 kbar	57%	43%	-	2.1 ns	-	-	9.2 ns	-	-	0.34	41.1	-	0.05
65R1 L99A													
0 bar	100%	-	-	3.1 ns	-0.14	0.95	-	-	-	0.44	43.6	-	0.06
1 kbar	100%	-	-	3.7 ns	-0.14	0.95	-	-	-	0.44	43.6	-	0.09
2 kbar	100%	-	-	4.6 ns	-0.14	0.95	-	-	-	0.44	43.6	-	0.17
3 kbar	100%	-	-	6.3 ns	-0.14	0.95	-	-	-	0.44	43.6	-	0.43
72R1 L99A													
0 bar	100%	0%	-	1.7 ns	-0.23	-	21 ns	-	-	0.6	47.8	-	0.05
1 kbar	97%	3%	-	1.8 ns	-0.15	-	21 ns	-	-	0.6	50.7	-	0.05
2 kbar	88%	12%	-	2.1 ns	0.06	-	21 ns	-	-	0.6	52.0	-	0.05
3 kbar	72%	28%	-	14 ns	1.6	-	21 ns	-	-	0.6	54.5	-	0.05
80R1 L99A													
0 bar	85%	15%	-	1.3 ns	-0.02	-	6.9 ns	-	-	0.34	90.0	48.9	0.05
1 kbar	89%	11%	-	1.4 ns	-0.02	-	12 ns	-	-	0.34	90.0	49.6	0.05
2 kbar	84%	16%	-	1.6 ns	-0.02	-	13 ns	-	-	0.34	90.0	51.6	0.05
3 kbar	60%	40%		1.8 ns	-0.02	-	11 ns	-	-	0.34	90.0	44.2	0.05
82R1 L99A													
0 bar	100%	0%	-	1.6 ns	1.4	-	-	-	-	0.45	36.0	-	0.05
1 kbar	97%	3%	-	1.6 ns	1.28	-	2.3 ns	2.19	-	0.45	36.0	-	0.05
2 kbar	63%	37%	-	1.5 ns	0.82	-	4.9 ns	0.64	-	0.45	36.0	-	0.05
3 kbar	28%	72%	-	1.7 ns	0.62	-	8.0 ns	0.07	-	0.45	36.0	-	0.05
89R1 L99A													
0 bar	100%	0%	-	2.0 ns	-0.78	-	-	-	-	0.41	28.9	-	0.05
1 kbar	100%	0%	-	2.1 ns	-0.7	-	-	-	-	0.41	29.1	-	0.05
2 kbar	100%	0%	-	2.5 ns	-0.67	-	-	-	-	0.41	29.5	-	0.05
3 kbar	89%	11%	-	2.6 ns	-0.19	-	2.1 ns	-	-	0.41	31.2	-	0.05
109R1 L99A													
0 bar	76%	23%	1%	1.2 ns	-0.27	-	1.2 ns	-	1.1 ns	0.41	27.9	-	0.05
1 kbar	74%	19%	6%	1.4 ns	-0.18	-	1.2 ns	-	1.1 ns	0.41	27.9	-	0.05
2 kbar	55%	12%	33%	1.6 ns	0.24	-	1.2 ns	-	4.8 ns	0.41	27.9	-	0.05

3 kbar	41%	7%	52%	2.3 ns	0.52	-	1.2 ns	-	6.0 ns	0.41	27.9	-	0.05
119R1 L99A													
0 bar	64%	36%	-	1.8 ns	0.2	-	4.8 ns	-	-	0.46	45.6	-	0.05
1 kbar	67%	33%	-	2.1 ns	0.39	-	6.8 ns	-	-	0.46	45.1	-	0.05
2 kbar	63%	37%	-	2.8 ns	0.71	-	8.6 ns	-	-	0.46	44.7	-	0.05
3 kbar	65%	35%	-	13 ns	1.8	-	17 ns	-	-	0.46	53.6	-	0.05
128R1 L99A													
0 bar	82%	18%	-	0.60 ns	-0.68	-	7.2 ns	-	-	0.61	51.9	-	0.05
1 kbar	77%	23%	-	0.66 ns	-0.51	-	7.9 ns	-	-	0.61	51.9	-	0.05
2 kbar	53%	47%	-	0.95 ns	-0.03	-	8.0 ns	-	-	0.61	51.9	-	0.05
3 kbar	41%	59%	-	1.3 ns	0.21	-	9.1 ns	-	-	0.61	51.9	-	0.05
131R1 L99A													
0 bar	100%	0%	-	1.9 ns	-0.47	-	-	-	-	0.33	30.5	-	0.05
1 kbar	95%	5%	-	1.9 ns	-0.25	-	28 ns	-1.3	-	0.33	30.5	-	0.05
2 kbar	74%	26%	-	1.7 ns	0.02	-	26 ns	-1.3	-	0.33	30.5	-	0.05
3 kbar	33%	67%	-	1.6 ns	0.5	-	24 ns	-1.3	-	0.33	30.5	-	0.05
140R1 L99A													
0 bar	96%	4%	-	1.5 ns	1.26	-	21 ns	-	-	0.39	36.0	-	0.05
1 kbar	85%	15%	-	1.7 ns	1.18	-	17 ns	-	-	0.39	36.0	-	0.05
2 kbar	67%	33%	-	1.7 ns	0.96	-	14 ns	-	-	0.38	36.0	-	0.05
3 kbar	50%	50%	-	1.6 ns	0.63	-	11 ns	-	-	0.30	36.0	-	0.05
151R1 L99A													
0 bar	92%	8%	0%	1.3 ns	0.03	-	18 ns	-	-	0.57	45.3	-	0.05
1 kbar	86%	14%	0%	1.5 ns	0.03	-	20 ns	-	-	0.57	45.3	-	0.05
2 kbar	80%	18%	2%	1.7 ns	0.03	-	14 ns	-	1.8 ns	0.57	45.3	-	0.05
3 kbar	61%	29%	10%	1.5 ns	0.61	-	9.5 ns	-	1.7 ns	0.57	45.3	-	0.05

Chapter 6. Materials and methods

6.1 Construction, expression, and purification of T4 lysozyme mutants

Construction of all cysteine mutants have been described elsewhere with the exception of A93C.[1–4] The cysteineless "pseudo wild type" T4L construct (C54T/C97A) was a gift from F.W. Dahlquist at the University of California, Santa Barbara and was transferred into the pET11a vector in house. Point mutations were introduced by the QuikChange method (Stratagene, La Jolla, CA) and were verified by sequencing (Genewiz Inc., South Plainfield, NJ). Specific mutants were transformed into the *E. coli* BL21(DE3) expression strain (Stratagene).

Soluble T4L mutants in the WT* background were purified as described by Mchaourab *et al*.[3] T4L in the L99A background were expressed

Isolation of T4 lysozyme (T4L) from inclusion bodies was done for 93C L99A and 119C L99A as described by Liu *et al*[5]. with some modifications, as reported in López *et al*.[6] In short, after sonication of the crude extract, insoluble material was collected by centrifugation and washed with detergent. Unfolding with 4 M urea at pH 3.0 was followed by gradual dialysis in two steps, first by removing urea, and second by increasing the pH to 5.0. The sample was then purified by loading onto a Hi Trap CM FF ion exchange column (GE Healthcare, Chicago, IL) in the same manner as solubly-expressed T4L.

Sample purity of isolated T4L was greater than 95%, as judged by SDS/PAGE electrophoresis. All samples were concentrated and buffer exchanged into 20 mM MES at pH 6.8 using centrifugal concentrators (Millipore, Burlington, MA: Amicon 10,000 MWCO Filters); final protein concentrations were determined by absorption at 280 nm (ϵ = 24,750 L mol^{-1} cm^{-1}).

6.2 Construction, expression, and purification of intestinal fatty acid binding protein mutants

In a similar manner to T4 lysozyme, point mutations were introduced by the QuikChange method (Stratagene, La Jolla, CA) into the wild-type rat intestinal fatty acid binding protein (rat I-FABP) gene in a pET11d plasmid—a gift of Alan M. Kleinfeld (Torrey Pines Institute for Molecular Sciences, San Diego, CA)—verified by sequencing (Genewiz Inc., South Plainfield, NJ). All mutants were co-transformed into *E. coli* BL21(DE3) expression cells with the pEVOL-pAzF plasmid—a gift of Peter Schultz (The Scripps Research Institute, La Jolla, CA), which encodes a tRNA with specificity towards the amber codon TAG.

Induction at a cell density of 0.6-0.8 with isopropyl β-D-1-thiogalactopyranoside (IPTG) was followed after 20 min by induction with 2% arabinose (Sigma Aldrich, St. Louis, MO) and addition of 0.2 g (1 M in 1 L culture) para-azido-phenylalanine (SynChem Inc, Elk Grove Village, IL), after which the temperature was reduced to room temperature and the expression was allowed to proceed for 16 hr. The native rat I-FABP sequence has a dual TAG-TAA stop codon sequence, to which a para-azido-phenylalanine is incorporated at the C-terminus.

Purification of soluble I-FABP has been described elsewhere.[7] Briefly, after sonication, ammonium sulfate precipitation is followed by cation and anion exchange at the pI using HiTrap SP HP and HiTrap Q FF columns (GE Healthcare, Chicago, IL)in tandem at pH 7.3 and size exclusion using a 24 ml Superdex 200 column (GE Healthcare, Chicago, IL).

Purification of insoluble I-FABP from inclusion bodies was done for 34A and 51A mutants by solubilization of the cell mass into 8 M urea, 20 mM potassium phosphate, pH 7.3 after 5 min sonication on ice and incubation at 4 °C. The gelatinous cell membranes were separated by centrifugation and the unfolded protein was dialyzed against 2 L of refolding buffer

(20 mM potassium phosphate, pH 7.3) at 4 °C overnight. The refolded sample is cleared by centrifugation and is purified by application over a 24 ml Superdex 200 column. Final protein concentrations were determined by absorption at 280 nm with an extinction coefficient of $\epsilon =$ 18,700 L mol^{-1} cm^{-1}, which was validated experimentally by Kirk *et al*.[8]

6.3 Spin labeling of T4 lysozyme and intestinal fatty acid binding protein mutants

Cysteines of T4L and I-FABP were reduced by incubation in 2 mM dithiothreitol (DTT) for 30 min at room temperature. Immediately before introducing the spin label reagent, samples were passed over two tandem 5 ml HiTrap Desalting columns (GE Healthcare, Chicago, IL). Spin labeling of T4L was carried out at pH 6.8 in buffer (50 mM MOPS ((3-(N-morpholino)propanesulfonic acid)), 25 mM NaCl, at pH 6.8) and I-FABP at pH 7.3 (20 mM potassium phosphate, 0.1 mM EDTA, at pH 7.3) with a 3-fold molar excess of MTSSL ((S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate) diluted from a 200 mM acetonitrile stock and incubated at 4 °C overnight. Excess spin label was removed by passage over two 5 ml HiTrap Desalting columns. Samples were concentrated with centrifugal concentrators and diluted into 25% Ficoll-70 (Sigma Aldrich, St. Louis, MO).

6.4 Site-specific immobilization and delipidation of intestinal fatty acid binding mutants

I-FABP was site-specifically biotinylated using a Click chemistry scheme developed by López *et al*.[9] in which WS Biotin DBCO (a water soluble dibenzocyclooctyne-amine coupled to biotin via a sulfonated 13-atom linker, Click Chemistry Tools, Scottsdale, AZ) is diluted to a molar equivalent with the quantity of I-FABP (generally 5 mg of protein) from a 10 mM stock in dimethylsulfoxide. Modification at 4 °C overnight was followed by passage over two 5 ml

HiTrap Desalting columns, then incubation at 4 °C overnight on 100 ul of high-capacity Streptavidin agarose beads (Thermo Scientific, Waltham, MA). The unbound protein which is enriched in I-FABP that lacked the para-azido-phenylalanine modification was removed from the supernatant, and the Streptavidin-immobilized I-FABP was delipidated by buffer washes in 8 M urea, followed by refolding in 20 mM potassium phosphate, 0.1 mM EDTA, at pH 7.3.

Buffers at varying pH values include: 50 mM glycine at pH 2.2, pH 2.8, and pH 3.6 prepared from glycine-HCl; 50 mM acetate at pH 4.2 and pH 4.8; 50 mM formic acid at pH 3.6 and at pH 4.2. 50mM bis-Tris methane, 13 mM NaCl at pH 6.1 and 50 mM bis-Tris methane, 31 mM NaCl at pH 6.8. These buffers are approximately matched in ionic strength to 20 mM potassium phosphate, 0.1 mM EDTA, at pH 7.3. Ionic strength was measured by conductivity and approximate concentrations of NaCl were determined from a polynomial fit to the conductivity of NaCl at varying weight percentages at 25 °C.[10]

6.5 Continuous wave EPR

Atmospheric pressure continuous wave (CW) EPR spectra were recorded at room temperature at X-band on a Varian E-109 spectrometer fitted with a two-loop one-gap resonator (Medical Advances, Milwaukee, WI, USA)[11] operating with 2 mW incident power. Samples were loaded by capillary action into fused quartz capillaries and sealed with Critoseal (McCormick Scientific, St. Louis, MO).

Variable pressure CW EPR spectra were recorded at room temperature at X-band on a Varian E-109 spectrometer fitted with a 5-loop 4-gap "Rising Sun" resonator operating with 5 mW incident power. Protein samples of 20 μl were loaded into a pressure-resistant ceramic cell developed by Pressure Bioscience, Inc (PBI, Easton, MA) and Michael Lerch. A pressure

intensifier (model HUB440) and control software written by Christian Altenbach were used for equilibrium and pressure-jump experiments. Atmospheric pressure spectra acquired before and after pressurization were collected to verify reversibility; all data in this work are reversible.

6.6 Saturation recovery EPR

Saturation recovery (SR) EPR data were recorded on a Elexsys E580 pulsed EPR spectrometer (Bruker, Billerica, MA) fitted with a Stanford Research Instruments amplifier (Part #SR445A) as reported previously.[12] Temperature and atmosphere control were controlled by a N_2 temperature control unit (model BVT3000, Bruker, Billerica, MA). Pump pulses at 250 mW and 500 ns duration were provided by the electron-electron double resonance source at the maximum absorbance line of the $m_I = 0$ manifold of the nitroxide spectrum. The CW observe power was 100 μW at the same frequency as the pump pulse. The defense pulse length was 280 ns. Each measurement was independently repeated 3-5 times with 2.10 million accumulations each, half on- and half off-resonance with a 1 Hz field step of -10 G upfield.

UPEN fitting was performed by a LabVIEW package written by Michael Bridges (anticipated availability on the Hubbell Lab web site).

6.7 Double electron-electron resonance EPR

Protein samples of 12 μl were loaded into fused-quartz capillaries and snap-frozen in liquid nitrogen.

Four-pulse double electron-electron resonance (DEER) at Q-band was conducted on an Elexsys E580 pulsed EPR spectrometer fitted with an ER510QT-2 cavity resonator (Bruker, Billerica, MA). Samples were maintained at 50 K with a recirculating helium cry-cooling system

(ColdEdge Technologies, Allentown, PA). Observe pulse lengths were ~18 ns (90°) and ~36 ns (180°) as optimized by an echo nutation experiment. A pump pulse consisting of a linear frequency-swept pulse of 50 MHz half-width generated by an arbitrary waveform generator (Bruker, Billerica, MA) was applied 70 MHz below the observer frequency. All pulses were amplified with an external 150-watt TWT amplifier (Applied Engineering Systems, Fort Worth, TX).

Background correction and model-free fitting of dipolar evolution data was performed using LongDistances, a LabVIEW program written by Christian Altenbach available at https://sites.google.com/site/altenbach/labview-programs/.

6.8 Bibliography

1. Guo, Z., Cascio, D., Hideg, K., Kálái, T. & Hubbell, W. L. 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 (2007).
2. Guo, Z., Cascio, D., Hideg, K. & Hubbell, W. L. Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci.* **17**, 228–239 (2007).
3. Mchaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35**, 7692–7704 (1996).
4. Koteiche, H. A. & McHaourab, H. S. Mechanism of chaperone function in small heat-shock proteins. Phosphorylation-induced activation of two-mode binding in αb-crystallin. *J. Biol. Chem.* **278**, 10361–10367 (2003).

5. Liu, L., Baase, W. A., Michael, M. M. & Matthews, B. W. Use of stabilizing mutations to engineer a charged group within a ligand-binding hydrophobic cavity in T4 lysozyme. *Biochemistry* **48**, 8842–8851 (2009).

6. Lopez, C. J., Yang, Z., Altenbach, C. & Hubbell, W. L. Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc. Natl. Acad. Sci.* **110**, E4306–E4315 (2013).

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

8. Kirk, W. R., Kurian, E. & Prendergast, F. G. Characterization of the sources of protein-ligand affinity: 1-Sulfonato-8-(1 ′)anilinonaphthalene binding to intestinal fatty acid binding protein. *Biophys. J.* **70**, 69–83 (1996).

9. López, C. J., Fleissner, M. R., Brooks, E. K. & Hubbell, W. L. Stationary-phase EPR for exploring protein structure, conformation, and dynamics in spin-labeled proteins. *Biochemistry* **53**, 7067–7075 (2014).

10. Shreiner, R. H. & Pratt, K. W. Primary Standards and Standard Reference Materials for Electrolytic Conductivity. in *NIST Special Publication* 1–24 (2004).

11. Hubbell, W. L., Froncisz, W. & Hyde, J. S. Continuous and stopped flow EPR spectrometer based on a loop gap resonator. *Rev. Sci. Instrum.* **58**, 1879–1886 (1987).

12. Bridges, M., Hideg, K. & Hubbell, W. Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Appl. Magn. Reson.* **37**, 363–390 (2010).