Interplay of Local Structure, Dynamics, and Self-Association in HSPB1

Amanda F. Clouser

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

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Rachel Klevit, Chair

Michael Ailion

Champak Chatterjee

Program Authorized to Offer Degree:

Biochemistry

University of Washington

Abstract

Interplay of Local Structure, Dynamics, and Self-Association in HSPB1

Amanda F. Clouser

Chair of the supervisory committee:

Rachel Klevit

Department of Biochemistry

Small heat shock proteins (sHSPs) comprise a class of ATP-independent chaperones that prevent aggregation of destabilized proteins in the cell. The structural mechanisms by which sHSPs complete their function remain enigmatic due to inherent properties that make them refractory to traditional approaches in structural biology. While sHSP monomers are small (on average 20 kDa), for certain sHSPs these monomers assemble into large homo-oligomers (~500 kDa) of variable size and shape. Structural heterogeneity is observed not only on a global scale but also at the local level. Observed heterogeneity is thought to stem from the predominantly disordered N-terminal region (NTR) of sHSPs. There is evidence for numerous sites on sHSPs for interacting with client proteins, although these interactions are often transient. The combination of transient interactions with clients, competing sHSP-client and sHSP-sHSP interactions, and presentation of different binding sites in variable states of an oligomer make it particularly challenging to understand structure-function relationships in sHSPs.

The work presented in this thesis correlates local and global structural properties of the polydisperse human sHSP HSPB1. Hydrogen-deuterium exchange mass spectrometry (HDXMS) was used to identify solvent protected regions in the NTR of HSPB1 oligomers, providing the first full look at the NTR of an oligomeric human sHSP. Additionally, specific regions of local structural heterogeneity were identified in the NTR. Utilizing known mutations that perturb inter-protomer interactions in oligomers, a dimeric, full-length form of HSPB1 was obtained to represent a dispersed subunit of the system. The dimeric form of HSPB1 shows altered local structure in the NTR by HDXMS relative to WT oligomers, indicating distinct structural changes among different states of HSPB1. The dimeric construct is also amenable to traditional NMR approaches, allowing for development of a comprehensive model of full-length HSPB1 dimers, which still show structural heterogeneity. Disease-associated mutations in the NTR of HSPB1 were also characterized using HDXMS and a suite of other biophysical techniques. Specific regions in the NTR have altered protection among these disease mutants. These disease mutations were introduced to the dimeric construct to probe each mutation's effect on oligomerization propensity. Interestingly, two of the disease mutations promote oligomerization while the other two disease mutations retain some dimer-like characteristics in the dimer-mutant context. HDXMS analysis of the disease mutants in the dimeric construct reveals distinct local structural effects for each mutant. It is most remarkable that mutations in the same region of HSPB1 have different local mechanistic effects, and that the subsequent global effects on oligomeric size or chaperone activity can appear similar for certain mutants. The methods explored in this thesis for characterizing heterogeneous structure in HSPB1 can be implemented to other sHSPs and similarly complex protein systems.

TABLE OF CONTENTS

ABSTRACT iii

TABLE OF CONTENTS v

LIST OF FIGURES AND TABLES ix

1. INTRODUCTION 1

1.1 Small heat shock proteins 1

1.1.1 Biological roles 1

1.1.2 General structural properties and implications for function 3

1.1.3 Reported structures and alternative oligomeric models 7

1.2 HSPB1 9

1.2.1 Overview of HSPB1 relative to other human sHSPs 9

1.2.2 Biological roles, disease-associated mutations, and stress-induced phosphorylation 10

1.2.3 α-crystallin domain structures 11

1.2.4 Structural features of the full-length protein 12

1.3 Structural characterization of heterogeneous, partially disordered proteins 13

1.3.1 Nuclear magnetic resonance spectroscopy 14

1.3.2 Hydrogen-deuterium exchange mass spectrometry 16

1.4 Scope of this thesis 17

2. METHODS 27

2.1 Mutagenesis, expression, and purification of HSPB1 27

2.1.1 Vectors and mutagenesis 27

2.1.2 Protein expression 27

2.1.3 Purification 29

2.2 Functional (chaperone) assays 31
2.3 Global biophysical characterization 31
2.3.1 Analytical size-exclusion chromatography 31
2.3.2 SEC coupled with multi-angle light scattering 32
2.3.3 Circular dichroism spectroscopy 32
2.3.4 Negative-stain electron microscopy 33
2.3.5 Fluorescence-based subunit exchange 33
2.4 Hydrogen-deuterium exchange mass spectrometry 34
2.4.1 Sample preparation 34
2.4.2 Mass spectrometry 35
2.4.3 Data analysis 35
2.5 Nuclear magnetic resonance spectroscopy 36
2.5.1 Triple resonance experiments for assignments 36
2.5.2 Intensity and CSP analysis of 2D spectra 37
2.5.3 Paramagnetic relaxation enhancement 37
2.5.4 Methyl-TROSY 37
3. HSPB1 OLIGOMERIZATION IS MODULATED BY THREE TYPES OF INTERDEPENDENT INTERMOLECULAR INTERACTIONS 45
3.1 Phosphorylation-mimicking mutations in the NTR alter local structure and lower oligomeric propensity 45
3.1.1 WT and the triple phosphorylation mimic (“D3”) 45
3.1.2 Single phosphorylation mimics 48
3.1.3 Double phosphorylation mimics 49

3.1.4 Mixing different phosphorylation mimics 50
3.2 Changes in redox state of the ACD alter oligomeric properties 52
3.3 Disruption of the CTR IXI motif is associated with larger, more polydisperse oligomers 53
3.4 Combination of phosphorylation-mimicking and CTR mutations 54
3.5 Discussion of the interplay of inter-protomer interactions in HSPB1 oligomerization 57
3.5.1 Perturbation of NTR-NTR interactions 57
3.5.2 Perturbation of the inter-subunit ACD-CTR interaction 61
3.5.3 Combinatorial effects of perturbing NTR-NTR and ACD-ACD interactions 64
4. STRUCTURAL CHARACTERIZATION AND COMPARISON OF WT OLIGOMERS AND D3GXG DIMER 77
4.1 Comparison of solvent accessibility of regions in WT oligomers and D3GXG dimers 78
4.1.1 Technical considerations 78
4.1.2 ACD and CTR 79
4.1.3 NTR 80
4.2 Stability of structural features in WT oligomers and D3GXG dimers 82
4.2.1 Thermal melt analysis of WT and D3GXG 82
4.2.2 Characterization of N-terminal peptide proposed to form β-hairpin 83
4.3 Structural analysis of D3GXG dimer by NMR 85
4.3.1 Technical considerations 85
4.3.2 Assignment of the NMR spectrum of the full-length dimer 87
4.3.3 Comparison with ACD-only construct 89
4.3.4 Insight into the NTR 90

4.3.5 Assignment of the CTR 91
4.3.6 Probing long-range interactions by PRE 91
4.4 NMR of HSPB1 oligomers 94
4.5 Discussion of structural analysis of WT oligomers and D3GXG dimers 97
4.5.1 Structural changes in the NTR between WT oligomers and D3GXG dimers 97
4.5.2 Model of the D3GXG dimer 98
4.5.3 Implications of NTR-ACD contacts for disease-associated mutants in the ACD 101
5. MECHANISMS BY WHICH DISEASE-ASSOCIATED MUTATIONS IN THE NTR AFFECT HSPB1 OLIGOMERIZATION 124
5.1 Disease-associated mutants have substantial changes in global structure and function ..124
5.2 Disease-associated mutants have subtle changes in local structure and dynamics 128
5.3 Disease-associated mutations cause disparate global effects in a dimer-prone mutant ...132
5.4 Sites of local changes in structure and dynamics in disease-associated mutations in D3GXG 134
5.5 Discussion of structure-function relationship in disease-associated mutants of HSPB1..136
6. CONCLUSIONS AND FUTURE DIRECTIONS 160
6.1 Summary of novel findings on HSPB1 structure 160
6.2 Future directions 162
6.3 Functional questions in the sHSP field 165
6.4 A technical roadmap for characterizing sHSPs 168
REFERENCES 171
APPENDIX A- HDXMS peptides.....176
APPENDIX B- NMR assignments 189

LIST OF FIGURES AND TABLES

Figure 1.1- Domain architecture of sHSPs and sequence alignment of better-characterized human sHSPs 20

Figure 1.2- sHSP ACD structure and interacting regions among protomers 21

Figure 1.3- Multiple conformations of the NTR in human sHSP structures 22

Figure 1.4- Sequence alignment of HSPB1 orthologs, residues of interest, and known and predicted secondary structure 23

Figure 1.5- Comparison of the orientation of the β2 region in HSPB1 ACD-only structures 24

Figure 1.6- Observation of conformational heterogeneity by solution NMR 25

Figure 1.7- Observation of conformational heterogeneity by hydrogen-deuterium exchange mass spectrometry 26

Table 2.1- HSPB1 constructs used in this thesis 39

Figure 2.1- Schematic of procedure for making methyl-labeled perdeuterated protein 40

Figure 2.2- Properties of the fluorophore used for subunit exchange experiments 41

Figure 2.3- Schematic for HDXMS experiment and analysis of bimodal peptides 42

Figure 2.4- MTSL spin label used for NMR-PRE experiments 44

Figure 3.1- SEC and CD profiles of WT and D3 HSPB1 oligomers and ACD-only 66

Figure 3.2- Concentration dependence of SEC and CD profiles of single phosphorylation mimics of HSPB1 67

Figure 3.3- Concentration dependence of SEC and CD profiles of double phosphorylation mimics of HSPB1 68

Figure 3.4- SEC profiles of mixtures of single phosphorylation mimics of HSPB1 69

Figure 3.5- Redox effects on SEC profiles of intermediate phosphorylation mimics of HSPB1 70

Figure 3.6- Characterization of the GXG mutant of HSPB1 by SEC, CD, and negative-stain EM 71
Figure 3.7- SEC, MALS, and CD profiles of the D3GXG mutant of HSPB1 73
Figure 3.8- SEC and CD profiles of single phosphorylation mimics with the GXG mutation in HSPB1 74
Figure 3.9- Chaperone assay of WT, D3, GXG, and D3/GXG forms of HSPB1....75
Figure 3.10- Summary of effects of perturbing inter-protomer interactions in HSPB1 76
Figure 4.1- Comparison of solvent protected regions in HSPB1 WT oligomers and D3GXG dimers using HDXMS 103
Figure 4.2- Far UV CD thermal melts of HSPB1 WT and D3GXG proteins 104
Figure 4.3- Analysis of HSPB1 NTR peptide consisting of residues 40-53....105
Figure 4.4- Comparison of D3/176del and D3GXG ^{1}H-^{15}N HSQC-TROSY spectra 106
Table 4.1- Expected and observed resonances in D3/176del ^{1}H-^{15}N HSQC-TROSY spectra....107
Figure 4.5- Comparison of ACD-only with D3/176del assignments, chemical shifts, and heterogeneity....108
Figure 4.6- Two distinct sets of peaks observed for a region of the NTR in the D3/176del ^{1}H-^{15}N HSQC-TROSY spectrum....112
Figure 4.7- Interactions of the NTR with the ACD as seen with PRE probes....114
Figure 4.8- ^{1}H-^{13}C HMQC spectra of δ-methyl Ile-labeled WT and D3 oligomers....118
Figure 4.9- ^{1}H-^{13}C HMQC of δ-methyl Leu/γ-methyl Val labeled WT oligomers 119
Figure 4.10- Model of structural features observed in D3GXG 120
Figure 4.11- Interactions of the NTR with the ACD in HSPB6 dimers 123
Figure 5.1 - Chaperone assay of NTR disease mutants and WT HSPB1 145

Figure 5.2 - SEC and SEC-MALS profiles of NTR disease mutants and WT HSPB1 146

Table 5.1- Molar masses and corresponding number of subunits calculated for WT and NTR disease mutant oligomers based on SEC-MALS 147

Figure 5.3 - Far-UV CD profiles of WT and NTR disease mutants of HSPB1.....148

Figure 5.4 - Subunit exchange kinetics of NTR disease mutants and WT HSPB1 via homo-FRET 149

Table 5.2- Calculated subunit exchange rates for NTR disease mutants and WT HSPB1 at different concentrations and temperatures via homo-FRET 150

Figure 5.5 - Comparison of solvent protected regions in NTR disease mutants WT HSPB1 using HDXMS 151

Figure 5.6- Chaperone assay of NTR disease mutants in the D3GXG context of HSPB1.....153

Figure 5.7- SEC and CD profiles of NTR disease mutants in the D3GXG context of HSPB1...154

Figure 5.8- Comparison of solvent protected regions in NTR disease mutants in the D3GXG context of HSPB1 using HDXMS 156

Figure 5.9- Comparison of G84R/D3GXG and D3/176del HSPB1 dimers by ^{1}H-^{15}N HSQC-TROSY spectra 158

Table 5.3- Summary of functional, structural, and dynamics changes in NTR disease-associated mutants of HSPB1 159

ACKNOWLEDGEMENTS

They say it takes a village to raise a child, and it certainly takes a broad scientific community to raise a scientist. For completion of this interdisciplinary thesis, I must thank people in many departments across the University of Washington campus. First and foremost, my adviser and mentor Prof. Rachel Klevit has provided a creative environment with what felt like limitless opportunity. I thank her for giving me the chance to pursue such an exploratory and challenging project and providing essential suggestions and guidance along the way. Most importantly, she encouraged me to pursue interests beyond pure research, which allowed me to realize my love for biophysical instrumentation and helping others use these tools.

I thank Dr. Ponni Rajagopal for being a superb mentor and for teaching me most of what I know about NMR and sHSPs. While every member of the Klevit lab has aided in my development in some way, I especially thank a few members of the sHSP project- Chris Woods, Hannah Baughman, and Andrew Borst. These fellow graduate students have provided invaluable feedback on my project and our sHSP projects as a whole. Most importantly, they made work fun and helped generate a collaborative environment. I must also thank the undergraduate students I mentored, Bobby Shih and Thanh Chu, as they maybe taught me more than I taught them.

As for the broader scientific community, I thank my committee members- Professors Michael Ailion (Department of Biochemistry), David Baker (Biochemistry), Champak Chatterjee (Chemistry), and Kelly Lee (Medicinal Chemistry). Their support and diverse suggestions over the years have been vital to my success. Prof. Miklos Guttman (Medicinal Chemistry) taught me everything I know about HDXMS, first as my mentor in my very first rotation, through my final experiments in graduate school. His selflessness in sharing his

expertise with scientists across campus has brought new life to our structural biology community.

The Biophysical Physics, Structure & Design Program that accepted me and the Department of Biochemistry that became my home department have provided countless opportunities and truly wonderful, collaborative, supportive environments. I could not have been part of a better starting class of graduate students in these two programs, as we enjoyed the highs and suffered the lows together. Without the compassion, patience, and persistence of our graduate coordinator, Erin Kirschner, I do not know how I would have made it across the finish line. Financially, I received support from the Hurd Fellowship and the Molecular Biophysics Training Program, which was a diverse group of biophysicists that I loved being a part of.

Finally, I thank my family for their unconditional love and support through the years- my husband Clancey, our three pets, my brother, and my parents. My parents, an astronomer and a geologist, were my first scientific mentors. They always encouraged me to use my brain to its full capacity through math, science, and music. They went so far as to invent the gift-granting "math fairy" when I was young, and although I was initially devastated to find out she was not real, to this day I cannot help but feel the same sense of accomplishment anytime I solve a math problem. Wading through creeks looking for fossils and staring through telescopes on cold, clear nights taught me to always be searching for something, to make discoveries, and to question and appreciate the world around us.

1. INTRODUCTION

1.1 Small heat shock proteins

1.1.1 Biological roles

For proper function, proteins must fold into their native 3-dimensional structure. Many proteins can adopt a variety of functionally relevant structural conformations, and certain "structures" are often highly flexible disordered states. Nature has evolved careful regulation of various structural states of proteins, often fine-tuning their functions. However, proteins sometimes do not adopt their ideal conformations and will "misfold". Misfolding occurs stochastically but is aggravated by certain stressors, such as increased temperature, oxidative stress, acidosis, or mutation of the protein sequence. Misfolded proteins generally lose their native function and can be prone to aggregation. As misfolded proteins associate into aggregates they can also draw in other proteins, interfering with multiple cellular pathways. When misfolding and aggregation of proteins overwhelms a system, cell death and disease can ensue. Therefore nature has developed a class of proteins, termed chaperones, which help other proteins refold or prevent aggregation. The function of chaperones is tightly regulated to maintain proteostasis. Levels of control range from tissue-specific and stress-induced transcriptional regulation to post-translational modification. Chaperones themselves are subject to mutation, which can alter their ability to maintain proteostasis. Many mechanistic questions associated with chaperone structure-function are only partially answered, such as: how do chaperones interact with misfolded/aggregated proteins, how is chaperone function regulated, and how do chaperones function differently when mutated.

The chaperone examined in this thesis is a member of the small heat shock proteins (sHSPs). sHSPs are shorter in sequence length than most other classes of chaperone (~150-200

residues in human sHSPs), hence the "small" part of the name. The "heat shock" part of the name is primarily based on their discovery. Cells subjected to heat shock were found to upregulate certain proteins, including some of the sHSPs, especially HSPB1 (Ritossa 1996). The canonical function of sHSPs is to maintain aggregation-prone proteins ("clients") in a soluble state ("chaperone" activity). Interaction with clients and prevention of aggregation is ATP-independent for sHSPs, contrary to the other major classes of chaperones. Chaperone activity *in vitro* is determined by measuring the ability of a sHSP to delay aggregation of a model client protein that has been destabilized. Certain sHSPs are also known to interact with proteins outside of a chaperoning context, such as cytoskeletal proteins and proteins in the apoptosis pathway (Gusev et al. 2002; Lanneau et al. 2008). sHSPs are found in all kingdoms of life, with multiple paralogous sHSPs present in many organisms. Within multicellular organisms, some sHSPs are differentially expressed among tissue types while others are more ubiquitous.

There are ten human sHSPs (named HSPB1-10), several of which are stress-regulated and/or implicated in various diseases. Many of the human sHSPs are phosphorylated in response to stress. Phosphorylation-mimicking mutants (phosphorylation site is mutated to Asp or Glu) of certain sHSPs have increased activity *in vitro* and confer increased cell viability when expressed *in vivo* (Rogalla et al 1999). Mutant forms of sHSPs are associated with different diseases, sometimes dependent upon their tissue expression (Datskevich et al. 2012). For example, HSPB4 is expressed almost exclusively in the eye lens, and mutant forms of HSPB4 are reported in some patients with early-onset cataracts. Other mutant forms of human sHSPs are implicated in different types of neuropathies and myopathies. Most mutations reported in human patients affected by these diseases are missense mutations and occur in only one copy of the gene, thus the effects are generally autosomal dominant. Certain mutations have been introduced into

model organisms to demonstrate causative effects (Xi et al. 2008; Zhang et al. 2014). Some human sHSPs have also been found upregulated in cancer cells that have become resistant to chemotherapy (Arrigo et al. 2007; Vargas-Roig et al. 1998). Therefore sHSPs are a medically relevant topic of research, however much less is known about the structure-function relationship in sHSPs compared to other chaperone types.

1.1.2 General structural properties and implications for function

The primary structure of sHSPs can be separated into three regions as shown in Figure 1.1a: the N-terminal region (NTR), the α-crystallin domain (ACD), and the C-terminal region (CTR). Disease-associated mutations are reported for all three regions and will be discussed with respect to results in subsequent chapters. It is well documented that many sHSPs form homo-oligomers (and even hetero-oligomers with other sHSPs) via interactions involving one or more of these regions. Oligomers range in size and can be monodisperse (one type of multimer exists) or polydisperse (a range of multimers coexist) depending upon the sHSP. The three regions of a sHSP are involved in a minimum of three types of inter-protomer interaction within an oligomer. Below is a description of general sequence and structural properties of each of the three regions, their involvement in homo-oligomer formation, and their role in binding of client proteins.

(1) α-crystallin domain (ACD)- sHSPs are defined by their ACD, which is highly conserved both in sequence and structure (Figure 1.1b). The ACD is around 90 residues in length and adopts a β-sandwich fold consisting of six β-strands (Figure 1.2a). In almost all reported structures of sHSPs, a dimer is formed via interaction of two ACDs (Figure 1.2b, type 1 interaction, "ACD-ACD"). The β6+7 strands between two protomers are aligned anti-parallel in the dimer, forming a continuous β-sheet. The ACD can be truncated and expressed on its own

("ACD-only"), forming the same topology as in the oligomeric context. The strand nomenclature of the ACD is based upon older structural characterization of non-vertebrate sHSPs in which at least nine β-strands are observed. In some non-vertebrate sHSPs, the β6 and β7 strands are distinct, with a long connecting loop. The β6 strand then swaps with the β6 strand of the neighboring protomer, forming a dimer with a different interface from vertebrate ACD dimers. In vertebrate sHSPs, the loop between the β6 and β7 strands is missing, resulting in a single β-strand in reported structures ("β6+7"). The β4 and β8 strands form the "β4/β8" groove at the edge of the ACD opposite the ACD-ACD interface. The β4/β8 groove is a binding site for an "IXI motif", which is a 3-residue motif with isoleucines (or valines) at the first and third positions, in a disordered region of a protein. Therefore, the groove can interact with both the CTR of sHSPs that contain an IXI motif (Figure 1.2b, type 2 interaction, "ACD-CTR") and client proteins with an IXI motif in a natively or destabilization-induced (unfolded/misfolded) disordered region.

(2) C-terminal region (CTR)- The flexible, short CTR follows the ACD and is highly variable in sequence and length (~10-30 residues). The CTR is rich in charged and polar residues and is predicted to be disordered. Many sHSPs contain an IXI motif (I/V-X-I/V) in their CTR that can bind the β4/β8 groove in the ACD of a neighboring protomer (Figure 1.2b, type 2 interaction, "ACD-CTR"). Additionally, peptides that mimic the CTR and include the IXI motif will bind the β4/β8 groove. However, only a small fraction of CTRs in an oligomer is thought to be bound to neighboring ACDs, based on observation of bound and unbound states by nuclear magnetic resonance spectroscopy (Baldwin et al. 2012). Of the sHSPs that have some structural characterization, ones lacking the IXI motif in their CTR generally do not form large oligomers.

(3) N-terminal region (NTR)- The NTR precedes the ACD and is also highly variable in sequence and length. This region ranges in length from ~50-100 residues in eukaryotes and is generally shorter than 50 residues in bacteria and archaea. The NTR is enriched in nonpolar residues, particularly aromatics and prolines. However, the NTR is predicted to be primarily disordered based on secondary structure algorithms and is mostly not resolved in reported structures. Only one short region in the NTR is conserved among human sHSPs (26-SRLFDQAFG-34 in HSPB1, Figure 1.1b). Although the NTR is not well resolved in reported structures, many models predict it to interact with neighboring NTRs in sHSPs that form large oligomers (Figure 1.2b, type 3 interaction, "NTR-NTR"). Some sHSPs have an IXI motif in their NTR, suggesting that the NTR might also bind the β4/β8 groove of the ACD, and this has been structurally shown in one case. Many sHSPs are phosphorylated in their NTR, which is often stress induced. Phosphorylation or phosphorylation-mimicking mutations in the NTR generally causes dissociation of sHSPs that form large oligomers into smaller species. For many sHSPs, phosphorylated forms or phosphorylation-mimicking mutants of the protein are more active chaperones. Many studies have also implicated the NTR in client interactions.

Most sHSPs self-associate to form a broad range of species, from dimers to >50mers. sHSPs that form only dimers associate via the ACD-ACD interaction. Larger oligomers are formed via the ACD-CTR and NTR-NTR interactions in addition to the ACD-ACD interaction. sHSPs that form large oligomers are classified as either monodisperse or polydisperse (a range of multimers coexist). The oligomeric non-vertebrate sHSPs tend to be monodisperse, forming discrete species like 12-mers or 24-mers. Oligomeric vertebrate sHSPs tend to be polydisperse, forming a range of species as broad as 10-50mers. Subunits exchange between oligomers

rapidly, although it is unknown what the exchanging species are, which could range from monomers to small oligomeric species.

Due to the polydisperse and dynamic behavior of many vertebrate sHSPs, it is also challenging to characterize interactions with client proteins and make correlations with function. Transient interactions have been shown between client proteins and both the β4/β8 groove of sHSP ACDs and parts of the NTR. In the truncated ACD-only form of a sHSP, the β4/β8 groove is maintained and can interact with model client proteins. However, the truncated form generally has little to no chaperone activity relative to the full-length sHSP. Therefore, binding (particularly to the β4/β8 groove) is not necessarily associated with chaperone activity. The NTR is implicated as functionally important for chaperone activity, at least for certain clients. NTR-client interactions are particularly challenging to capture due to the disorder of the NTR and the presence of multiple, weak, non-specific interactions with the client. It is likely that different clients interact with sHSPs by different modes as there is variability in the putative interacting regions of clients. Therefore, the function of sHSPs is better described by a range of possible interactions with clients and regulation of different interactions.

For large, polydisperse oligomers, it is generally thought that control over the oligomeric state of sHSPs is key to regulating their function. There are opposing ideas on whether "small" oligomers are more or less active than "large" oligomers. In most cases where phosphorylation (or mimicking mutations) reduces the average oligomeric size of a sHSP, chaperone activity also increases. However, this correlation of quaternary size and function is not maintained for many disease-associated mutants and other mutated forms of sHSPs that have been characterized. As there are overlapping regions involved in both self- and client-association (NTR and β4/β8 groove), a simple competition model would suggest that smaller oligomers would have more

available binding sites for client proteins. However, recruitment of a client protein into a larger oligomer provides the opportunity for avidity, as a client protein can interact with multiple sHSP subunits. Therefore, additional properties of sHSPs must be probed to better understand regulation of their function. Specifically, information on changes in local structure and dynamics (within a protomer) among oligomeric forms is lacking and would aid in a more comprehensive view of sHSPs, as is presented in this thesis.

1.1.3 Reported structures and alternative oligomeric models

There are numerous crystallographic and NMR structures reported for truncated, ACD-only forms of sHSPs (Bagneris et al. 2009; Baranova et al. 2011; Hochberg et al. 2014; Laganowsky et al. 2010; Rajagopal et al. 2015a and b; Weeks et al. 2013). Structures of vertebrate sHSPs adopt the same dimeric β-sandwich topology shown in Figure 1.2a, with the β6+7 strands aligned antiparallel. In the context of vertebrate oligomeric models, the β-sandwich structure of the ACD is conserved. However, among all of these structures, the curvature of the ACD dimer and precise alignment ("register") of the β6+7 strands are variable. Therefore many of the variables among these structures (length of construct, buffer conditions, mutations, etc.) can subtly influence the surfaces of the ACD dimer involved in self- and client-association (e.g. the β4/β8 groove).

There is little structural information on the NTRs of sHSPs in either small or large oligomeric forms. As the NTR is mostly disordered and even ordered regions adopt multiple conformations, this part of the protein remains refractory to routine structural approaches. In the context of an oligomer, many structures and models propose NTR interactions between neighboring protomers. There are atomic level structures and models for four non-vertebrate, monodisperse sHSP oligomers, all from different species (Kennaway et al. 2005; Kim et al.

1998; Van Montfort et al. 2001; White et al. 2006). The crystal structure of archaeal Hsp16.5 (monodisperse 24-mers) reveals a hollow ball but the NTRs are not resolved. A more recent cryo-EM model of this protein shows density inside the oligomer attributed to NTRs. Similarly, discrete 24-mers of yeast Hsp26 form hollow spheres in cryo-EM models with NTRs interacting. Tuberculosis Acr1 forms discrete, hollow spherical 12-mers, and the NTRs are not resolved in negative-stain EM models but are proposed to be directed inwards and highly disordered. In the crystal structure of wheat Hsp16.9 barrel-shaped 24-mers, half of the N-termini are resolved, helical, and form important intermolecular interactions with each other. Therefore, even in monodisperse sHSP oligomers, it is challenging to elicit any definitive "structure" (or map of ordered/disordered regions) of the NTR.

There are structural oligomeric models for only one vertebrate, polydisperse sHSP. For HSPB5 (also called αB-crystallin) there are two oligomeric models of tetrahedral 24-mers, which only represents one oligomeric species in a large range (Braun et al. 2011; Jehle et al. 2011). Both models have similar architecture of protomers: ACD-ACD interactions form dimers, ACD-CTR interactions link three dimers into a hexamer, and NTR-NTR interactions bring together four hexamers into the tetrahedral 24-mer. However, the models differ in their placement of the NTR. The model put forth by the Klevit lab (Figure 1.3a) is based on solid-state nuclear magnetic resonance (NMR) spectroscopy, small-angle X-ray scattering (SAXS), and negative-stain electron microscopy (EM). NMR provided numerous restraints for the ACD (enough for ACD structure determination) and sparse restraints for the NTR. Modeled interacting regions between NTRs were confirmed by chemical cross-linking. Even from the sparse restraints for the NTR, conformational heterogeneity was evident. The resulting model of local NMR restraints fit into SAXS and EM densities places the NTRs oriented toward the center

of the oligomer. Two distinguishable conformations of the NTR are modeled in the 24-mer, although it is unknown how different conformations might be partitioned among different oligomeric states. The other model uses the same NMR restraints for the ACD, but models the entirety of the sequence into a higher resolution cryo-EM density. This model places the NTRs oriented away from the center of the oligomer, adopting different local structures, and forming contacts via different regions of the NTR. In both cases, limited regions of secondary structure are modeled in the NTR, most of which are helical and encompass the one conserved region in the NTR (residues 28-36 in HSPB5). Even within both models the NTRs are in two distinct conformations. Specifically, the β2 strand forms antiparallel to its proceeding β3 strand in half of the protomers (Figure 1.3b). In the other half of the protomers, this β2 region is modeled to extend in the opposite direction away from the ACD. The β2 strand (as part of the β-sandwich) is sometimes observed in crystal and NMR structures of truncated ACD-only constructs. Given that ACD-only constructs are often truncated very near or in the β2 strand, it is difficult to extrapolate the structure of this region in the oligomeric counterpart of each sHSP. As the sequences of NTRs are not very conserved among sHSPs, it is also imprudent to extrapolate to other sHSPs much NTR structural information from the HSPB5 oligomeric models.

1.2 HSPB1

1.2.1 Overview of HSPB1 relative to other human sHSPs

Human HSPB1 is the subject of this thesis. HSPB1 and many mutant forms are implicated in human disease, therefore there is a wealth of biological data. Of the human sHSPs, HSPB1 has the most reported disease-associated mutations (Datskevich et al. 2012; Efgrafov et al 2004; Houlden et al. 2008; Rossor et al. 2012). However, compared to the most structurally-studied human sHSP, HSPB5, there are relatively few structural studies of HSPB1.

Understanding the structure-function relationship for HSPB1 with its putative clients requires a stronger structural foundation of the protein itself, disease-associated mutants, and regulation of its oligomeric size. Of the human sHSPs, HSPB1 has the longest NTR (~90 residues), and consequently the longest sequence (205 residues). The NTR has a proline-rich insertion not present in the other human sHSPs (Figure 1.1b, residues 55-70). The CTR of HSPB1 contains an extended IXI motif (Figure 1.4, 179-ITIPV-183), and there is no IXI motif in its NTR. There is one cysteine residue in the sequence, which occurs at position 137, in the middle of the β6+7 strand. Unique from the other sHSPs, a disulfide bond can form across the ACD-ACD dimer interface under non-reducing conditions. HSPB1 forms large, polydisperse oligomers much like HSPB5, for which there are oligomeric structural models, but there is little sequence homology between the NTRs of these two proteins.

1.2.2 Biological roles, disease-associated mutations, and stress-induced phosphorylation

HSPB1 has been implicated in multiple biological roles in diverse cellular pathways, diseases, and tissues. It shows chaperone activity with model client proteins in assays *in vitro*. Outside of the canonical chaperone role, HSPB1 is also implicated in the apoptosis pathway and interacts with cytoskeleton proteins (Gusev et al. 2002; Lanneau et al. 2008). Upregulation of HSPB1 has been observed in certain types of cancer and has therefore become a target for drug development. Several mutations in HSPB1 are associated with severe neurological disorders-Charcot-Marie-Tooth disease (CMT) and distal hereditary motor neuropathy (dHMN). Predominantly single point mutations have been observed, but both truncation and frame-shift mutations have been reported in affected patients. Disease-associated mutations are found in all three regions of the protein, and almost all mutations appear to be autosomal dominant. Limited biochemical analysis has been performed on only some of these mutants, focusing primarily on

oligomeric size, chaperone activity, and other basic properties. There are four known disease-associated mutations in the NTR of HSPB1, which are the subject of Chapter 5- G34R, P39L, E41K, and G84R (Figure 1.4). A better understanding of how these mutations alter HSPB1 structure and function could aid our understanding of its native activity. Basic biochemical characterization has been reported for these mutants (Muranova et al. 2015, Nefedova et al. 2013b), but structural mechanisms for their impact on function is lacking.

HSPB1 has three phosphorylation sites in the NTR, which are conserved among orthologs (Figure 1.4). HSPB1 is substantially phosphorylated within minutes after cells are subjected to stress (heat, oxidative stress) by stress-related kinases. Phosphorylated states of HSPB1 or phosphorylation-mimicking mutants generally increase holdase activity *in vitro*. Expression of phosphorylation-mimicking mutants *in vivo* increases cell viability after being subjected to stress (Rogalla et al. 1999). Along with these biological studies, numerous reports demonstrate that phosphorylated HSPB1 or phosphorylation-mimicking mutants forms much smaller oligomers (Hayes et al. 2009; Jovcevski et al. 2015; Landry et al. 1991; McDonald et al. 2012). However, little is known about the structural mechanism for dissociation into smaller oligomers upon phosphorylation.

1.2.3 α-crystallin domain structures

Three structures have been reported for the truncated ACD-only form of HSPB1. The first reported crystal structure shows a protomer structure similar to that of other sHSP ACDs (Baranova et al. 2011). However, in the crystal the protomers formed a hexameric ring via non-canonical contacts. The authors suggested the hexameric ring was likely due to mutations (engineered to enhance crystallization) within the ACD and crystallization conditions. When the WT ACD was analyzed in solution by SAXS, the protomer structure fit a more canonical ACD-

ACD dimer envelope. The second crystal structure shows a canonical dimer via antiparallel alignment of the β6+7 strands (Figure 1.5, PDB ID: 4MJH, Hochberg et al. 2014). A peptide mimicking the CTR of HSPB1 (including the IXI motif) is bound to the ACD β4/β8 groove. In this structure, the β2 strand is evident and antiparallel to the β3 strand of the same protomer, analogous to HSPB5 ACD-only structures and half of the states reported in the HSPB5 oligomeric models. The Klevit lab solution NMR structure of HSPB1 shows the same canonical dimer structure (Figure 1.5, PDB ID: 2N3J, Rajagopal et al. 2015a). However, the β2 strand is not formed in this structure, even though the sequence of the construct included the same β2 region as in the crystal structure, with additional preceding residues. The β2 strand is predicted to form via secondary structure analysis (PSIPRED, Figure 1.4, Liam et al. 2000). These differing observations suggest that it has propensity to form under some conditions, but this state might not be fully populated. The relative orientation of the β2 region bridging the NTR and ACD is likely associated with very different conformations and arrangement of the NTR.

1.2.4 Structural features of the full-length protein

There are no reported structures of full-length HSPB1, but many studies have probed properties of the NTR. Secondary structure of the NTR is predicted to be mostly random coil with some α-helical regions (PSIPRED, Figure 1.4). Most structural studies include large truncations in the NTR or cysteine mutations to attach probes, which likely perturb the local structure and/or dynamics of the region. Truncations in the NTR generally limit the oligomeric size but affect chaperone activity and subunit exchange depending upon the mutated region. Electron-paramagnetic resonance (EPR) was used to probe the relative flexibility of different regions of the NTR. Overall, all probed sites in the NTR in WT oligomers were found to be relatively rigid, presumably involved in inter-protomer contacts. In a triple phosphorylation-

mimicking mutant ("D3"), the same sites in the NTR were found to be far more flexible, correlating with dissociation into smaller oligomers where the NTRs are less restrained. However, attachment of the EPR spin-label probe required single cysteine substitutions throughout the NTR. Many mutations led to changes (increase or decrease) in the oligomeric propensity. Altogether, these results suggest that much of the NTR in a WT-like HSPB1 oligomer is likely fairly restrained, but restrained regions do not necessarily adopt conventional secondary structure.

1.3 Structural characterization of heterogeneous, partially disordered proteins and complexes

The lack of atomic-level structural details on sHSPs is due to several properties that have made them intractable to traditional structural techniques- most vertebrate sHSPs have (1) intrinsically disordered regions (NTR and CTR), (2) variable oligomeric states (polydispersity), and (3) conformational heterogeneity among protomers (e.g. different modeled states of the NTR in the HSPB5 models). Structural techniques have been optimized for characterizing proteins that are highly ordered and monodisperse. X-ray crystallography, which can yield very high-resolution structures, is restricted to proteins that have limited flexible/disordered regions and sometimes requires truncation of flexible regions that might be of functional importance. While it is possible to crystallize protein complexes, a high level of monodispersity (one oligomeric state) is generally required. Therefore this method has been useful for examining ACD-only constructs of sHSPs that form monodisperse dimers and full-length versions of non-vertebrate sHSPs that form large but monodisperse oligomers but is intractable for polydisperse human sHSPs such as HSPB1. The field of cryo-EM has seen a rapid increase in resolution of published structures and is amenable to large proteins and complexes. However, disordered regions still

cannot be resolved by this method, and structural heterogeneity complicates analysis and limits resolution. Alternate techniques, including NMR spectroscopy and hydrogen-deuterium exchange mass spectrometry, can yield structural information that is inaccessible using techniques that rely on highly homogenous samples.

1.3.1 Nuclear magnetic resonance spectroscopy

Unlike other traditional structural techniques, nuclear magnetic resonance (NMR) spectroscopy is capable of characterizing disordered proteins/regions and conformational heterogeneity. NMR offers a diverse range of experiments for probing protein structure, dynamics, and properties of interacting regions with other proteins or ligands. Specific peaks (or "resonances") in an NMR spectrum arise from chemically distinct nuclei in a molecule. In protein NMR, it is most common to probe protons (abundant, NMR-active) coupled to isotopically labeled nitrogen atoms (^{15}N), yielding a two dimensional spectrum where each peak corresponds to one H-N group in a protein, predominantly the backbone amides. Therefore it is possible to probe the distinct chemical environments of individual residues in a protein. Additional nuclei can be observed and the resulting multi-dimensional experiments are listed in Chapter 2.

Peaks corresponding to nuclei in disordered regions of a protein tend to be sharp due to their flexibility, which results in longer relaxation times. Therefore disordered residues often have a higher "apparent intensity" compared to residues in structured regions of a protein. However, the degenerate chemical environments of disordered residues can lead to overlap of peaks, limiting spectral resolution. If a residue in a protein exists in two chemically distinct states that exchange slowly relative to the timescale of the NMR experiments (milliseconds and longer), two distinct peaks will be observed for that residue (Figure 1.6). Therefore it is possible

to observe and characterize multiple conformations. However, if the exchange between states is very fast, a single averaged peak will be observed. One complication is that conformational exchange on a timescale similar to the experiment ("intermediate exchange") will result in peak broadening, sometimes to the point of being unobservable. Similarly, peaks might be unobservable if the signal is partitioned among too many chemically distinct states. The exchange rate between conformational states can be measured by collecting a series of NMR spectra.

For traditional structural experiments to be carried out effectively in solution, the protein must be relatively small (< 20kDa for a streamlined approach). The NMR signal of large proteins relaxes very quickly, leading to very broad peaks (appear "low intensity"). Certain experimental schemes, discussed in the following chapters, can be used to examine the structure of proteins in solution that are much larger (up to ~80 kDa), but restraints for structural modeling are far more limited. For even larger proteins, solid-state NMR can be used for structure determination but this is generally an expensive, laborious route carried out on only highly specialized NMR spectrometers. Finally, minimal residue-level structural information can still be obtained for large proteins in solution via the methyl-TROSY method (Ruschak and Kay 2010). However, this approach is also expensive and is limited to flexible methyl groups in particular side-chains, which might not be of interest. NMR is valuable for pursuing certain questions about sHSP structure and behavior. For example, NMR can be used to report on the local dynamics throughout an ACD-only dimer. Also, it is possible to monitor residues affected in an ACD-only dimer as it interacts with a model client protein. Strategies for studying a full-length sHSP in a small oligomeric form (dimer) by NMR is a major topic of this thesis.

1.3.2 Hydrogen-deuterium exchange mass spectrometry

Hydrogen deuterium exchange mass spectrometry (HDXMS), although not typically included in the list of techniques for protein structure determination, is proving to be an invaluable alternative for a wide variety of structural questions unanswerable by traditional techniques. HDXMS measures dynamics of regions of a protein, defined by the digestion of the protein into small peptides after the exchange step. Therefore, the resolution is limited to the lengths of these peptides and no direct long-range information is obtained. However, in principle there is no limitation on the type of protein, size, or solution conditions. Additionally, structural changes can easily be tracked as a function of time following a perturbation (addition of a substrate, change in pH, etc.). It is far less feasible to obtain a crystal structure or prepare NMR samples for numerous points in a reaction. Therefore, HDXMS is ideal for answering questions about structural changes in complicated systems, especially for systems where a structure is currently unobtainable.

As with NMR, HDXMS is useful for characterizing disordered and heterogeneous structure, and there is no limitation on protein size. A protein sample is incubated in deuterated solvent and allowed to exchange for set periods of time before the reaction is quenched and the protein is digested into peptide fragments. Protons from backbone amides are able to exchange with deuterium at a rate corresponding to their propensity to be hydrogen-bonded or buried, and therefore protected from solvent. The resulting deuteration profiles of each peptide can be examined by mass spectrometry, where the isotopic distribution of the peptide mass shifts with more deuterium incorporated. Disordered regions of proteins usually readily exchange with the deuterated solvent. If two conformations are present in a protein with different levels of protection and the exchange rate between these states is slower than the first time point, then two

isotopic distributions can be observed (Figure 1.7). Assuming there is not substantial spectral overlap of the two conformations, it is possible to track both the intrinsic dynamics of each conformation (deuteration level) and the exchange rate between the two states (change in proportions of each conformation).

1.4 Scope of this thesis

The overarching goals of this thesis are interdependent- (1) to structurally characterize different oligomeric states of HSPB1 in order to better understand its regulation and (2) to explore more accessible and comprehensive methods for obtaining structural information about polydisperse, heterogeneous, and partially disordered proteins and complexes.

The results presented here are progressive in that it was necessary to first probe and perturb the known types of interactions holding together HSPB1 oligomers before developing methods to further characterize them structurally. Briefly, mutations to each of the three regions (NTR, ACD, and CTR) typically result in changes to the oligomeric states of HSPB1. Most point mutations reported in the literature and here result in formation of larger oligomers, with the exception of phosphorylation-mimicking mutations in the NTR. There is substantial interplay among these interactions, as perturbations to two types of interaction do not necessarily result in a sum of the two separate perturbations.

Mutations in the NTR and CTR were combined to limit interactions involving these regions to generate a highly active, monodisperse dimer, termed “D3GXG”. It combines three mutations in the NTR that mimic a stress-induced phosphorylated state with three mutations in the CTR that prevent interactions with the ACD. This dimeric mutant is a model construct for studying the basic building block of phosphorylated HSPB1 oligomers while avoiding the experimental limitations of polydisperse samples with high molecular weight species. It is

unknown how different the structures of phosphorylated and non-phosphorylated full-length dimers are. Nevertheless, the changing structural features observed in D3GXG are similar to those observed in intermediate phosphorylation mimics, specifically at low concentrations and smaller average oligomeric size. Therefore, local structural changes occur upon dissociation into smaller oligomeric species rather than resulting directly from specific mutations.

As D3GXG is amenable (relative to larger oligomers) to conventional NMR experiments, it was possible to assign NMR resonances to specific residues in parts of the NTR and CTR and most of the ACD. This process was heavily aided by previously published assignments of the ACD-only construct. Assigned regions of the NTR are highly disordered but also adopt multiple conformations. While much of the ACD appears structurally similar between the truncated (ACD-only) and full-length (D3GXG) contexts, certain regions are clearly different in D3GXG and demonstrate conformational heterogeneity. Placement of paramagnetic relaxation enhancement (PRE) probes throughout the NTR of D3GXG aided in determining relative positioning of regions of the NTR with respect to the ACD.

NMR of the large WT oligomers is limited to the methyl-TROSY approach in which one observes only the flexible methyl groups of particular side-chains. Spectra were obtained of Ile methyl-labeled WT and D3 oligomers. There are only six isoleucines in HSPB1, however more than six peaks were observed in the WT spectrum, indicating conformational heterogeneity. The spectrum of D3 oligomers also showed heterogeneity, but many peaks did not overlay with WT, indicating different conformations present. Due to the high cost and experimental difficulty of the methyl-TROSY technique, limited structural data was obtained for large HSPB1 oligomers by NMR spectroscopy.

HDXMS proved to be an invaluable tool for probing all regions of full-length HSPB1 in any oligomeric state. Regions of protection were observed in the NTR of both D3GXG dimers and WT oligomers. Multiple regions within the NTR also showed conformational heterogeneity in WT oligomers. Moderate but distinct changes in the NTR could then be demonstrated between these two states of HSPB1. Coupled with NMR results, the HDXMS profile of D3GXG has provided a working model for the general structure of a dispersed dimer of HSPB1.

Four NTR disease-associated mutants (G34R, P39L, E41K, G84R) were studied using many of these methods. Briefly, a combination of analyses of secondary structure (circular dichroism spectroscopy), quaternary structure, and regional dynamics (HDXMS) of disease mutants in both their native (no other mutations) and the D3GXG context revealed specific regions affected by each mutation. Additionally, a simple fluorescence-based assay was developed to measure the kinetics of subunits exchanging among oligomers and compare the global dynamics of these disease mutants. Therefore, it was possible to gain insights into the structural mechanisms by which these disease mutations lead to altered chaperone activity.

Altogether this thesis defines structural states of the polydisperse human sHSP HSPB1: in WT, "phosphorylated", and "diseased" contexts. This work ideally provides a roadmap for how to study other sHSPs and similarly complicated protein systems. Given the plasticity of sHSPs and similar proteins, it is misguided to try to draw mechanistic inferences from classical single-state structures. Therefore, it is necessary to embrace a combination of biophysical techniques that can report on both disordered and ordered regions, multiple conformations, and the dynamics of exchanges among states in order to truly define the "structure" of the challenging protein systems that remain to be studied.

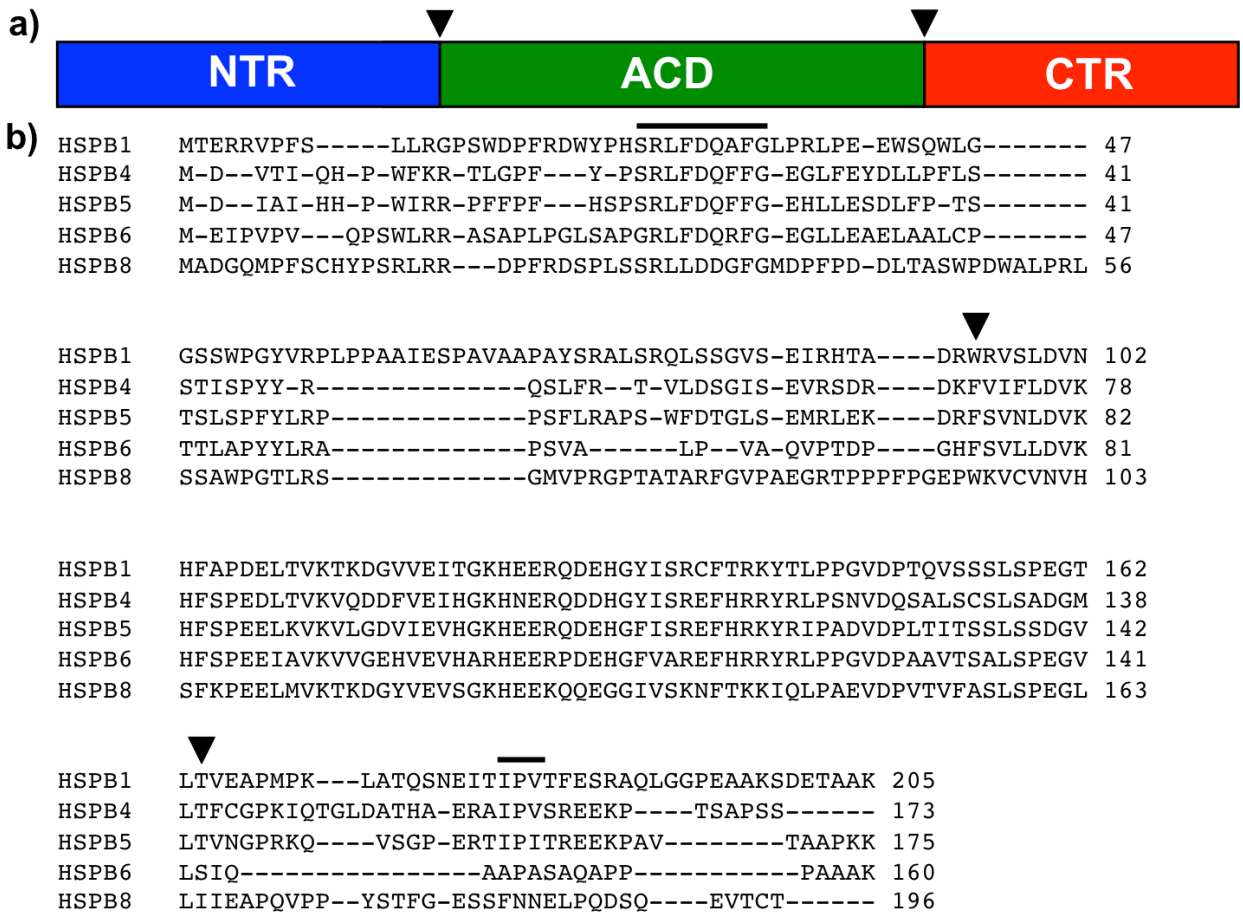


Figure 1.1- Domain architecture of sHSPs and sequence alignment of better-characterized human sHSPs. (a) sHSPs are characterized by three domains- the N-terminal region (NTR), α-crystallin domain (ACD), and C-terminal region (CTR); the color coding for these domains is maintained in the following figures. (b) Sequence alignment of the known human sHSPs shows high conservation of the ACD (sequence between the triangle markers) and low conservation of the NTR and CTR. The conserved region in the NTR and common IXI motif in the CTR are each indicated with a line above the sequence.

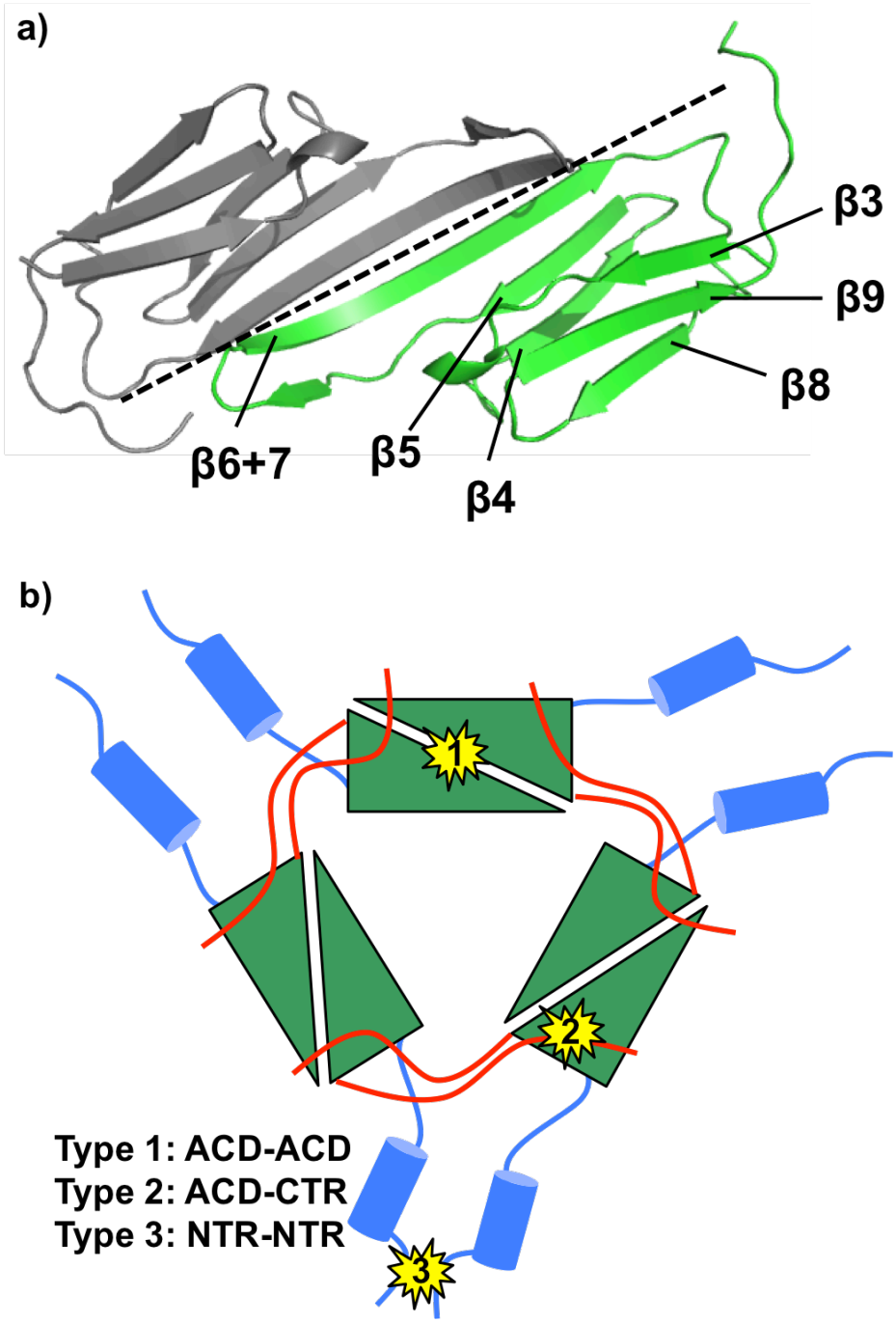


Figure 1.2- sHSP ACD structure and interacting regions among protomers. (a) NMR structure of truncated (ACD-only) human HSPB1 (2N3J), which forms the canonical dimer of β-sandwich protomers along the β6+7 strands. (b) sHSP oligomers are formed by three classes of self-interactions- (1) ACD-ACD as seen in (a), (2) ACD-CTR, in which a CTR IXI motif binds a neighboring ACD β4/β8 groove, (3) NTR-NTR, in which NTRs presumably interact but specific contacts are challenging to resolve.

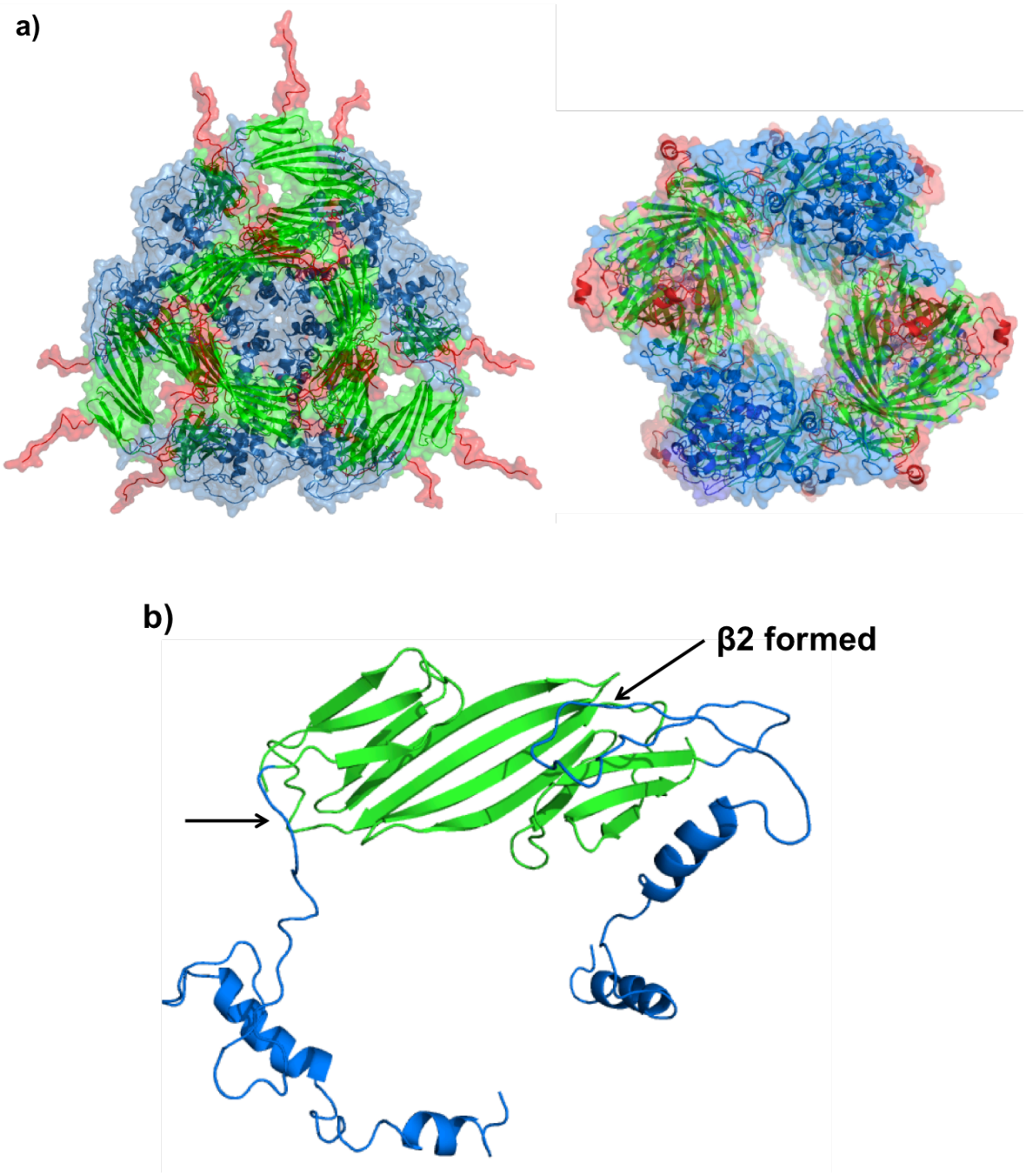


Figure 1.3- Multiple conformations of the NTR in human sHSP structures. (a) Left- model of an HSPB5 24-mer based on solid-state NMR, electron microscopy, and small-angle X-ray scattering (PDB ID: 3J07), in which two distinct N-terminal conformations are modeled. Right- model of an HSPB5 24-mer based on cryo-EM and fitting of an ACD-only structure (PDB ID: 2YGD). The two models differ primarily in arrangement of the NTRs. (b) One dimer subunit (NTR and ACD of each) from the left model in (a). Arrows indicate alternate positioning of the β2 region between protomers. Left arrow- NTR extends straight from the β3 strand. Right arrow- β2 strand forms antiparallel to the β3 strand and the NTR folds back again over the ACD.

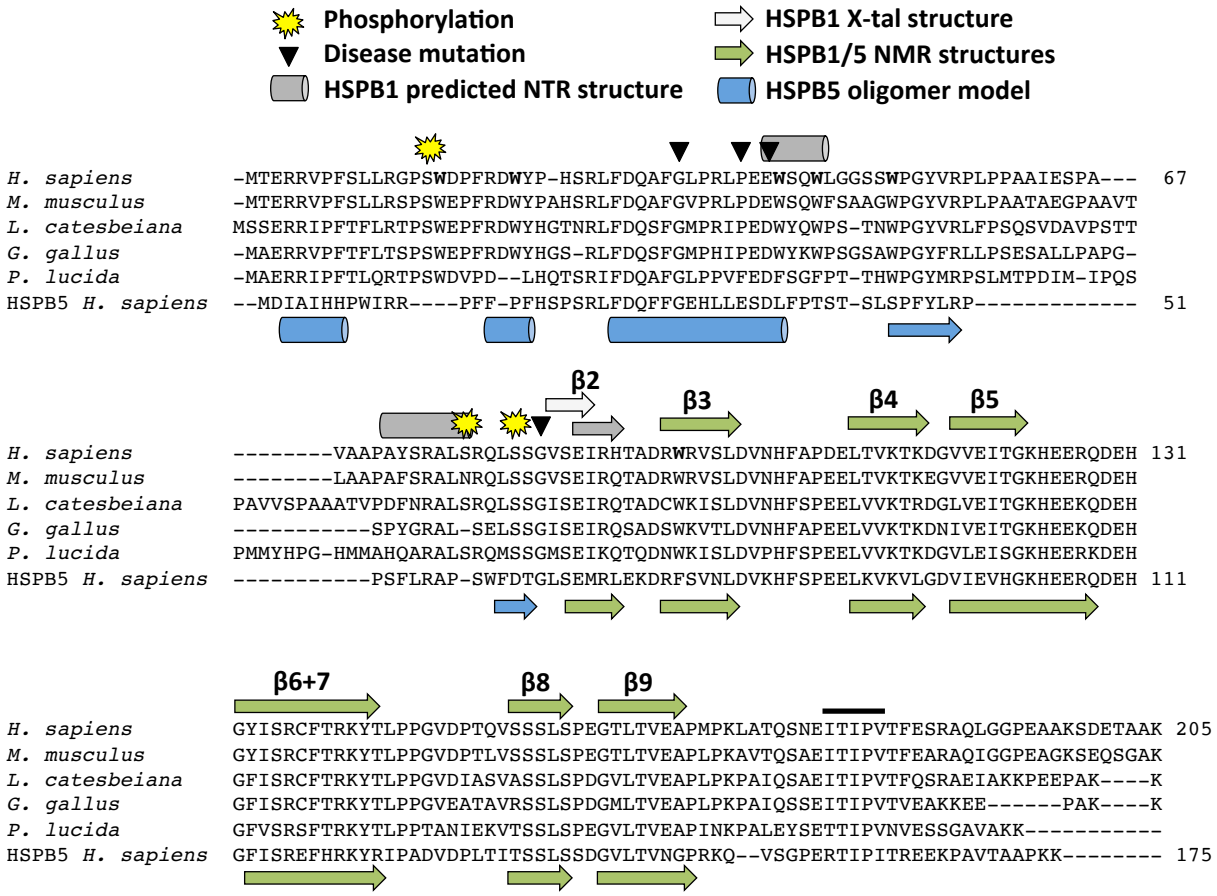


Figure 1.4- Sequence alignment of HSPB1 orthologs, residues of interest, and known and predicted secondary structure. HSPB1 sequences from human, mouse, frog, chicken, and fish show considerable conservation throughout most of the NTR, in contrast to other sHSP types as seen in the human HSPB5 sequence at the bottom. Yellow stars indicate conserved phosphorylation sites (S15, S78, and S82), and triangles indicate disease-associated point mutations in the NTR (G34R, P39L, E41K, and G84R). Gray helices and strands on top of the sequence are based on PSIPRED secondary structure prediction. The light gray β2 strand is based on the HSPB1 ACD-only crystal structure (PDB ID: 4MJH). The blue helices and strands below the sequence are based on secondary structure in the HSPB5 oligomeric model (PDB ID: 3J07). Green strands (below and above sequence) are based on HSPB1 and HSPB5 NMR ACD-only structures (PDB IDs: 2N3J and 2N0K). The black bar above the CTR indicates the IXI motif in HSPB1, which is actually a triple motif (ITIPV).

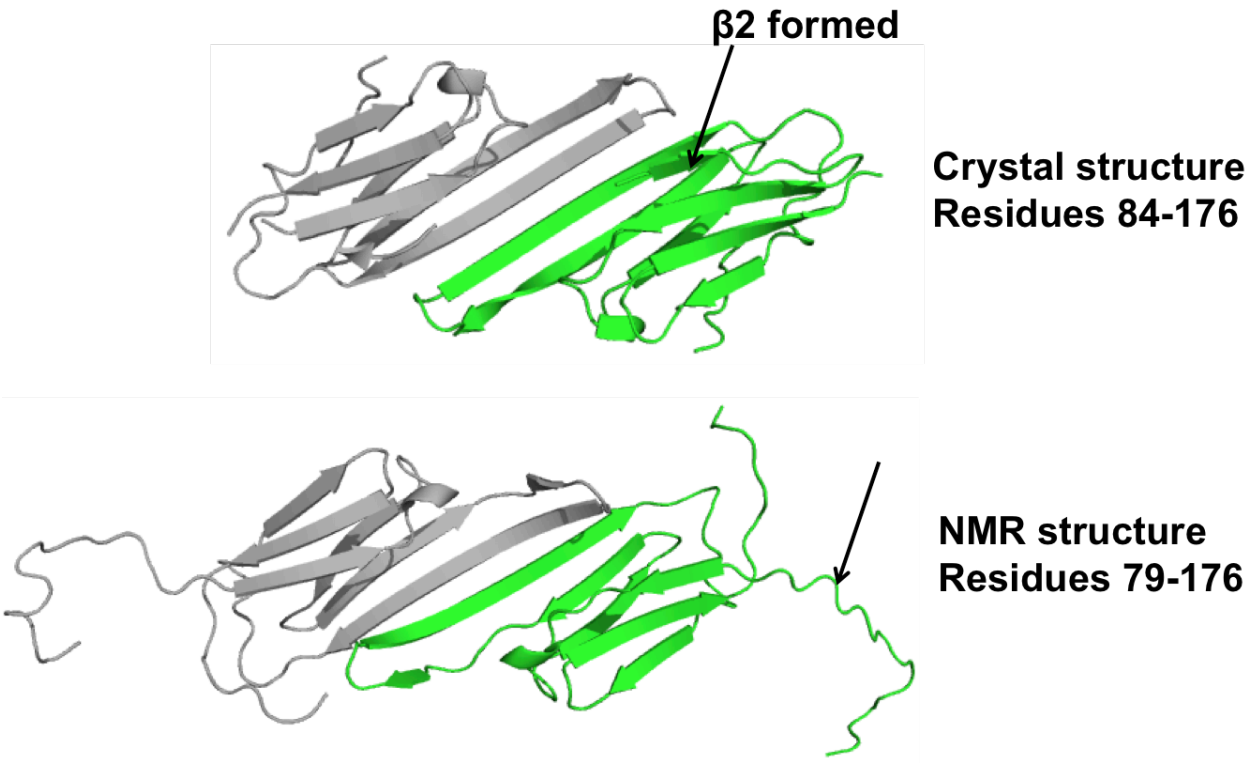


Figure 1.5- Comparison of the orientation of the β2 region in HSPB1 ACD-only structures. In the crystal structure of ACD-only (PDB ID: 4MJH, construct begins at residue 84), the β2 strand is formed antiparallel to the β3 strand of the same protomer. In the NMR structure (PDB ID: 2N3J), the β2 strand does not form and the end of the NTR (construct begins at residue 79) is modeled aiming away from the ACD.

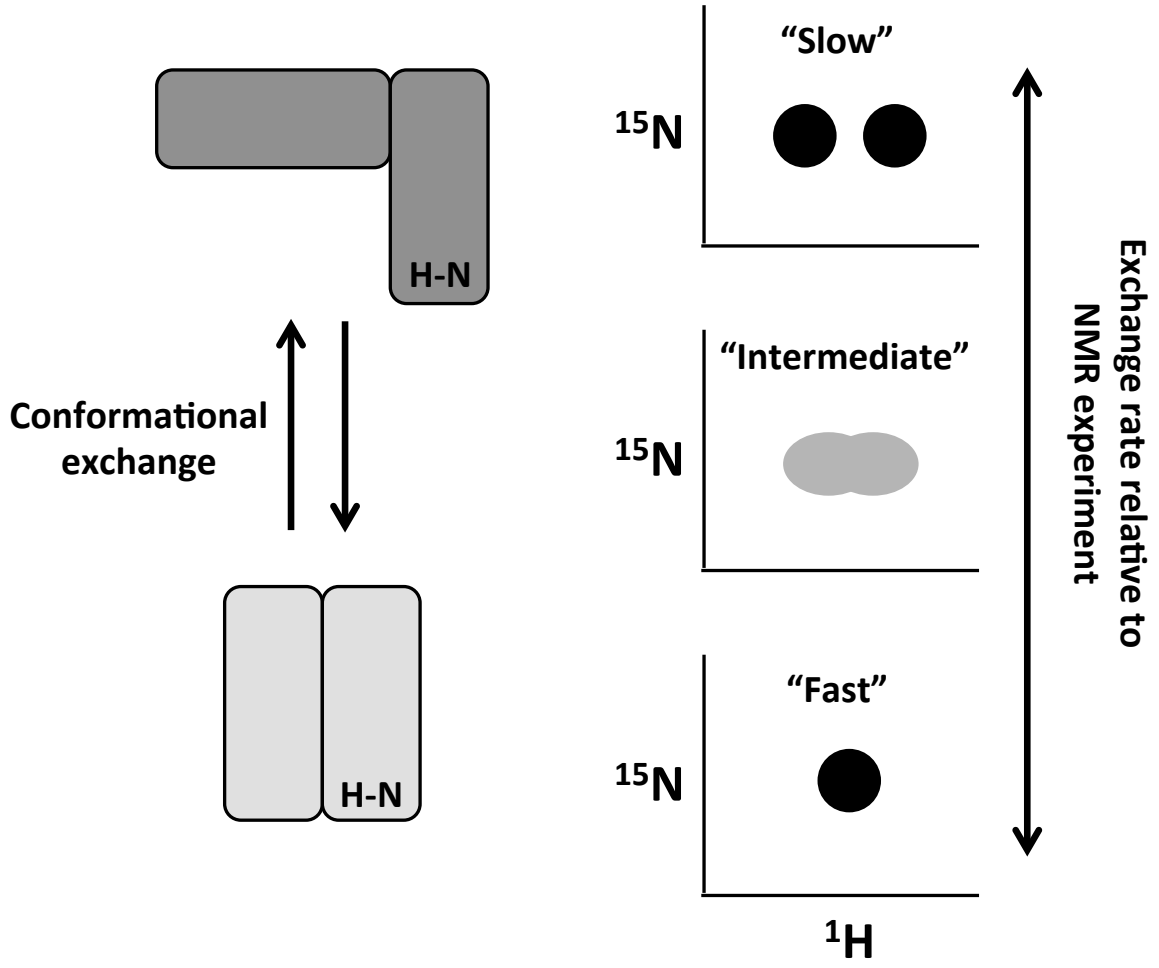


Figure 1.6- Observation of conformational heterogeneity by solution NMR. Most proteins adopt multiple structural conformations and interchange among these states at distinct rates that relate to function. If exchange between two states is slow relative to the NMR experiment (milliseconds and longer), two resonances will be observed. If exchange is similar to the NMR timescale, the resonances will broaden. If exchange is faster than the experiment, one resonance will be observed representing an average of these two states.

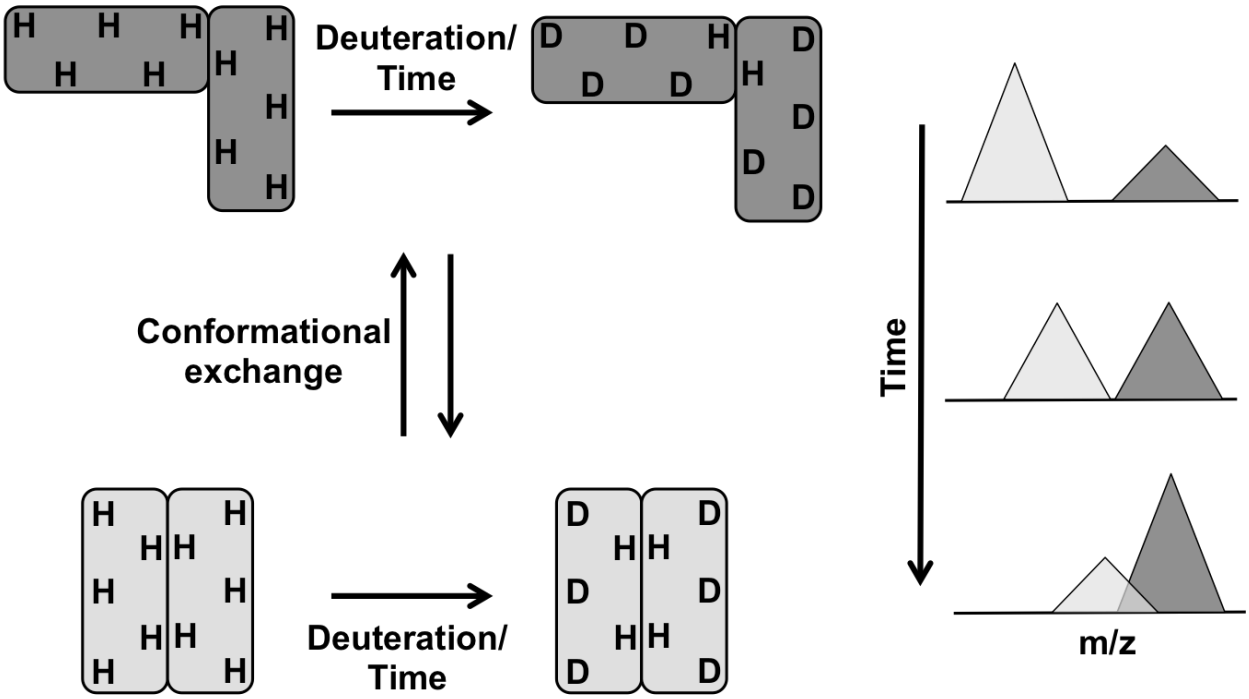


Figure 1.7- Observation of conformational heterogeneity by hydrogen-deuterium exchange mass spectrometry. If conformational exchange between two states of a protein is slower than the initial deuteration time points in an HDXMS experiment (usually seconds, but can be slowed at lower temperatures), two distinct isotopic profiles can be observed. The inherent dynamics of each state (deuteration levels) can be determined, and if the change in proportion of each state changes in the experimental timescale, one can also determine the exchange rate between states.

2. METHODS

2.1 Mutagenesis, expression, and purification of HSPB1

2.1.1 Vectors and mutagenesis

Human HSPB1 (accession # P04792) had previously been cloned into pET23a and pET151d vectors (ampicillin resistant). The ACD-only construct had previously been optimized to truncation of the full-length sequence from Gln80 to Ser176. Site-directed mutagenesis using the QuikChange protocol was used to introduce substitution mutations throughout the sequence. All mutants made of HSPB1 and their general purpose are listed in Table 2.1.

2.1.2 Protein expression

Several protocols for protein expression were used to obtain different isotopically labeled samples. In almost all cases (unless otherwise specified) for full-length HSPB1, BL21(DE3) *E. coli* cells were used and a final concentration of 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce protein expression. For the truncated ACD-only construct, 0.5 mM IPTG was used.

For natural abundance (or non-isotopically labeled) protein, colonies were scraped from a plate and added to 0.5 L of lysogeny broth (LB) with 100 μg/mL ampicillin. With pure protein yields of up to 100 mg from a 1 L culture, 0.5 L cultures were generally used. Cells were grown at 37°C in a shaking incubator until OD_{600} ~0.6. IPTG was then added to a final concentration of 1.0 mM and the temperature reduced to 22°C. Protein was expressed in a shaking incubator for ~22 hours. Cells were harvested by centrifugation and resuspension in lysis buffer (50 mM Tris, pH 8.0, 100 mM NaCl, 1 mM ethylenediaminetetraacetic acid [EDTA]). Cell pellets were stored at -80°C prior to protein purification.

For ^{15}N-labeled (no deuteration) protein, MOPS minimal media was used. Per 1 L culture, 1 g of $^{15}NH_4Cl$ was used for isotopic labeling. 4 g/L of glucose was used. Under these conditions, up to 100 mg of pure protein could still be obtained from a 1 L culture. Growth and expression steps were identical to those used for natural abundance protein. In this media, the doubling rate of the *E. coli* cells during log phase is ~1 hour.

For specific "unlabeling" of particular amino acids,^{15}N-labeled MOPS minimal media was used. However, the rate of protein production had to be limited for specific incorporation. Therefore, 3 g/L of glucose was used. Cells were grown at 37°C in a shaking incubator until OD_{600} ~0.6. Unlabeled amino acids were added ~15 minutes prior to induction to a final concentration of 1 g/L. Three classes of amino acids (among which scrambling occurs) were used- I/L/V, H, and F/Y. In each case, the total amino acid content added to the cultures was 1 g/L (e.g. 500 mg of F and 500 mg of Y per 1 L). Expression was induced with only 0.5 mM IPTG and cells were harvested after only 18 hours.

For $^{2}H^{15}N^{13}C$-labeled (partial deuteration) protein, cells were grown in stages to acclimate the cells to deuterated minimal media. Expression in D_2O based minimal media generally leads to ~75% deuteration. For all ^{13}C-labeling, 3 g/L of ^{13}C-glucose was used. One colony was grown in 3 mL of LB for ~5 hours and then centrifuged to pellet the cells. These cells were resuspended and grown in 50 mL of H_2O-based M9 minimal media to an OD_{600} ~0.6. Cells were again pelleted, resuspended, and grown in 100 mL of D_2O-based M9 minimal media to an OD_{600} ~0.6. Cells were then transferred directly to a larger 500mL (total) D_2O-based M9 minimal media and grown to an OD_{600} ~0.6. After induction and reduction of temperature, protein was expressed for ~48 hours. In the D_2O media, the doubling rate of the *E. coli* cells during log phase is ~2 hours.

For $^{2}H^{15}N^{13}C$-labeled (perdeuteration) or Ile-δ-$^{1}H^{13}C$-labeled (all other sites deuterated) protein, cells were grown in additional stages to acclimate the cells to deuterated minimal media (Figure 2.1). Expression in D_2O and ^{2}H-glucose minimal media generally leads to >95% deuteration. 3 g/L of $^{2}H^{13}C$-glucose was used. Stocks for deuterated M9 minimal media were also prepared in D_2O. One colony was grown in 3 mL of LB for ~5 hours and then centrifuged to pellet the cells. These cells were resuspended and grown in 50 mL of H_2O-based M9 minimal media to an OD_{600} ~0.6. Cells were again pelleted, resuspended, and grown in 100mL of D_2O-based M9 minimal (no deuteration) media for 1-2 hours. Cells were again pelleted, resuspended, and grown in 200mL of D_2O-based and ^{2}H-glucose-based M9 minimal media to an OD_{600} ~0.6. Cells were then transferred directly to a larger 1 L (total) D_2O-based and ^{2}H-glucose-based M9 minimal media. For uniformly perdeuterated protein, cells were grown to an OD_{600} ~0.6 and induced. For methyl-labeled protein, amino acid precursors (60 mg/L α-ketobutyrate for Ile labeling, 100 mg α-ketoisovalerate for Leu/Val labeling) were added at the final transfer step, and expression was induced at an OD_{600} ~0.25. After induction and reduction of temperature, protein was expressed for ~48 hours. In the D_2O-based and ^{2}H-glucose-based media, the doubling rate of the *E. coli* cells during log phase is ~3 hours.

2.1.3 Purification

Cells containing full-length HSPB1 were lysed by freeze-thaw and incubation with lysozyme and protease inhibitors in lysis buffer (specified in 2.1.2) on ice for 20 minutes. Generally 0.5 L cultures of cells were lysed in one tube to maximize yield and purity. Deoxycholate was added to the lysed cells and placed on a shaking incubator at 37°C for 15 minutes. DNase, RNase, and magnesium chloride were then added and shaking incubation continued for 15 minutes. Cell lysate was centrifuged at high speed at 4°C. Ammonium sulfate

was added to the supernatant to 40% saturation on a slow shaker at room temperature and allowed to equilibrate for 30 minutes. Ammonium sulfate precipitation is temperature dependent, therefore it is important to conduct these steps around room temperature. Ammonium sulfate precipitate was centrifuged at high speed at 20°C and the supernatant discarded. Pellets were used immediately or stored at -80°C for up to 1 week.

Ammonium sulfate pellets were resuspended in anion exchange buffer ("AEX"- 20 mM Tris, 10 mM $MgCl_2$, 30 mM NH_4Cl, pH 7.6) at room temperature. Remaining solids were pelleted briefly at 20°C. Protein was desalted using a G25 column in AEX buffer at room temperature. Light precipitated material was pelleted briefly at 20°C. At this desalted stage, protein will precipitate at cold temperatures. Desalted protein was separated by an anion exchange DEAE column with a step gradient of AEX buffer with increasing sodium chloride at room temperature. After anion exchange separation, protein was generally stored at 4°C. Protein fractions were analyzed for purity by SDS-PAGE, pooled, and concentrated to 2-4 mL for SEC. Preparations that yielded more protein were generally split into multiple SEC injections. Extra concentrated pre-SEC protein was often flash frozen in liquid nitrogen and stored at -80°C. Concentrated protein was separated on Superdex 200 or 75 columns in 50 mM NaPi, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer. Oligomeric proteins (WT, disease mutants, intermediate phosphorylation mimics) were separated on a Superdex 200 column, while smaller constructs (D3, D3GXG, D3/176del) were separated on a Superdex 75 column. SEC separated fractions were analyzed for purity by SDS-PAGE, pooled, and concentrated to the desired concentration. Concentration was determined by 280 nm absorbance (extinction coefficient of 40,450 $M^{-1}cm^{-1}$). Aliquots were flash frozen in liquid nitrogen and stored at -80°C.

For cysteine-free proteins or proteins containing only the native cysteine at position 137, no reducing agent was added during purification. With the native cysteine present, the resulting protein was generally >95% oxidized at the dimer interface as seen by dimer formation by non-reducing SDS-PAGE. For proteins containing non-native cysteines (for fluorophore or spin labeling), the reducing agent dithiothreitol (DTT) was included at each stage of purification to avoid disulfide formation.

2.2 Functional (chaperone) assays

Samples were prepared in a 96-well plate with final protein concentrations of 600 uM bovine α-lactalbumin and 60 μM HSPB1 in 50 mM sodium phosphate, 100 mM NaCl, and 0.5 mM EDTA at pH 7.5. Samples were incubated at 37°C for 30 minutes prior to triggering aggregation with final concentrations of 20 mM DTT and 5 mM EDTA. Absorbance at 360 nm was measured in a BioTek plate reader at 37°C every 2.5 minutes, with shaking prior to each measurement. Absorbance of each sHSP alone was measured to confirm that sHSPs were not aggregating under these conditions.

2.3 Global biophysical characterization

2.3.1 Analytical size-exclusion chromatography

Samples were prepared in 50 mM sodium phosphate, 100 mM NaCl, and 0.5 mM EDTA buffer at pH 7.5. Samples were incubated at the desired concentration at 37°C for at least one hour prior to injection. 150 μL samples were loaded onto a 100 μL loop and injected onto a 24 mL Superose 6 10/300 mm column (GE) at room temperature. For the indicated reduced samples, 2 mM DTT was present in both the sample and buffer used for column equilibration and elution.

2.3.2 SEC coupled with multi-angle light scattering

Chromatograms were collected on a GE AKTA Pure coupled to a Wyatt miniDAWN TREOS and Optilab T-rEX differential refractive index detector. Molar mass was calculated from the Raleigh ratio based on multi-angle (static) light scattering and protein concentration from the change in refractive index (dn/dc = 0.185). Analysis was performed using Wyatt ASTRA VI software, and curves were calibrated with the protein standard apoferritin. 100 μM samples of HSPB1 were prepared in 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, at pH 7.5. Samples were incubated at 37°C for one hour prior to placement in a 4°C autosampler, which injected 25 μL samples onto a 2.4 Superose 6 Increase 3.2/300 mm column (GE) equilibrated with the matched buffer at room temperature.

2.3.3 Circular dichroism spectroscopy

Samples were prepared in 25 mM sodium phosphate, 50 mM NaCl, and 0.25 mM EDTA buffer at pH 7.5. This is the standard buffer used in other experiments at half the concentration to limit absorbance due to buffer. Samples were incubated at the desired concentration at room temperature for several hours prior to measurement. Cuvettes ranging from 0.2 mm to 1 mm were used depending on the concentration of the protein. Specifically, more concentrated protein samples (e.g. 100 μM) would not allow enough light to pass through a 1 mm cuvette for reliable measurements, based on the recommended high-tension voltage cutoffs for the detector. All measurements were made on a Jasco J-1500 CD spectrometer with Peltier temperature control. Aside from thermal melts, all measurements were collected at 20°C, with 1 nm bandwidth, and averaged over three scans. For CD thermal melts, wavelength scans were collected every 10°C as indicated with a temperature ramping rate of 1°C/min. 1 nm bandwidth was maintained, but spectra were averaged over two scans.

All data were normalized to units of mean residue ellipticity (MRE), which accounts for protein concentration and number of peptide bonds so that any protein or peptide can be compared to another. Data were collected at 0.1 nm intervals and then smoothed linearly across 1 nm. All data presented has a high-tension voltage below the recommended cutoff for the detector (800 V).

2.3.4 Negative-stain electron microscopy

Protein samples were prepared at 100 μg/mL in 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer and placed on glow discharged carbon-coated copper grids. Phosphotungstic acid was used for the negative-stain, as it can be used at neutral pH, and HSPB1 behavior is pH-dependent. A Tecnai T12 microscope was used at 120 kV high-tension and ~10 μA emission current. Approximately 20 micrographs were collected for each protein at 26k magnification and -1 μm defocus. Images were processed using EMAN2 (Tang et al. 2007). Approximately 2000 particles were used for producing 50 class averages.

2.3.5 Fluorescence-based subunit exchange

WT and disease mutant constructs were generated with a cysteine introduced at position 174, analogous to similar fluorescence studies in HSPB5 (Peschel et al. 2013), and the native cysteine at position 137 mutated to serine. Protein was buffer exchanged into 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer with 2 mM tris-(2-carboxyethyl)-phosphin (TCEP), which does not interfere with maleimide reactions. Protein was incubated with 3X molar excess of Alexa Fluor 488 maleimide (stock in DMSO) and incubated at 37°C (to facilitate subunit exchange) for at least one hour (Figure 2.2). Protein was separated from free dye using gravity desalting columns. Fractions collected from desalting columns were analyzed for separation by comparing absorbances at 495 nm (fluorophore extinction coefficient of 73,000

$M^{-1}cm^{-1}$) and 280 nm, with an assumed A_{280}/A_{495} ratio of 0.11. The resulting pool of fractions was diluted to desired protein concentrations and fluorophore labeling percentages

Samples were incubated at 37°C at the desired protein concentration and labeling percentage for at least one hour prior to measurement. For the experiments presented here, samples with 30% fluorophore labeling were mixed 1:2 with unlabeled protein (Figure 2.2c). The resulting dequenching from homo-FRET among fluorophores in oligomers was measured as a function of time. Protein concentrations of 5 μM, 20 μM, and 40 μM were examined and different labeling ratios were used to confirm there was no fluorophore-concentration-dependent bias. Measurements were collected on a Horiba Fluorolog-3 with double excitation and emission monochromators and Peltier temperature control. Excitation and emission wavelengths of 518 nm and 498 nm were used, respectively.

The resulting kinetic data was fit to the following exponential equation using the R software package to obtain subunit exchange rates:

$$y = A + B * e^{-kt}$$

2.4 Hydrogen-deuterium exchange mass spectrometry

2.4.1 Sample preparation

200 μM protein samples were equilibrated at room temperature (22°C - 25°C) in 50 mM NaPi, 100 mM NaCl, 0.5 mM EDTA at pH 7.5 for several hours. The experimental scheme is decribed in Figure 2.3a. Samples were diluted 10X into deuterated buffer (prepared identically but with D_2O) for a final concentration of 20 μM, and incubated at room temperature for various periods of time to allow for hydrogen-deuterium exchange. At the desired time point, the deuteration reaction was quenched by adding an equal volume of quench buffer on ice. The quench buffer (mixture of formic and trifluoroacetic acids) was adjusted to bring down an equal

volume of D_2O buffer to a final pH of 2.5. At this low pH, proton-deuteron exchange at amides is limited and pepsin cleavage is optimized. Quenched samples were then immediately flash frozen in liquid nitrogen and stored at -80°C. Undeuterated samples were prepared in a similar fashion but replaced addition of D_2O buffer with normal buffer. Fully deuterated samples were made similarly but from denatured protein (3M guanidine HCl and high heat for at least 30 minutes) incubated in D_2O for several hours.

2.4.2 Mass spectrometry

Samples were stored in liquid nitrogen until 5 minutes prior to injection to maintain consistent levels of back-exchange of deuterium to hydrogen from the quench buffer. Samples were injected into a Waters SYNAPT G2 QTOF LC/MS system. Initially, the sample was loaded onto a column containing immobilized pepsin, in 0.1% TFA , at 1°C for digestion of the protein into peptides. Digested peptides were then loaded onto a C18 reverse-phase column with the same buffer and temperature conditions. Peptides were eluted using a gradient into a high acetonitrile concentration. Separated peptides were then analyzed by time-of-flight MS.

2.4.3 Data analysis

Initial samples of WT and D3GXG protein were analyzed by tandem MS to identify peptides of interest. ProteinLynx Global SERVER (Waters) software was primarily used for assigning peptides along with manual screening of spectra for additional peptides and confirming assignments. Most peptides were common among all mutants, and peptides containing a mutation (but identical otherwise) were generally found eluting at slightly different retention times. Certain mutations altered the pepsin digestion pattern, but suitably similar peptides were identified for qualitative comparisons. MassLynx (Waters) software was used to align spectra of various time points at the appropriate retention times, and HX-Express2 (Guttman et al. 2013)

was used to extract and quantify the isotopic distribution of each spectrum. The deuteration level at each time point was calculated from deuteration levels of undeuterated and fully deuterated spectra for each peptide. Briefly, the HX-Express2 program fits the spectrum at each time point to a binomial distribution expected for that peptide (different numbers of amides). In cases where the fit was very poor due to very broad isotope distributions, an alternative fitting was used to a bimodal distribution, representing two distinct deuteration states of the peptide at a given time point (Figure 2.3b). Specifically, deuteration levels are provided for both states as well as relative proportions of each state. In most cases the bimodal behavior was clearly visible in the spectrum. However, if the two states were similar in deuteration levels and therefore overlapping considerably or one state was lowly populated (< 10%), the exact values reported are less reliable and are not over-interpreted in the subsequent analyses.

2.5 Nuclear magnetic resonance spectroscopy

All NMR experiments were carried out on either 600 or 800 MHz Bruker spectrometers equipped with cryoprobes. All samples were prepared in 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer. Spectra were collected at 30°C, which was the optimal temperature for increased signal and limited breakdown.

2.5.1 Triple resonance experiments for assignments

Several TROSY-based triple resonance experiments were implemented to assign peaks in the D3GXG and D3/176del spectra to particular residues. HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HNCB, and HNCOCANNH ("NNH") experiments all provided useful spectra for the assignment process. Non-uniform sampling (NUS) at a sampling rate of 25% was used for longer experiments to reduce the time required for the same level of signal-to-noise. For NUS datasets, an iterative soft threshold algorithm was used to reconstruct full spectra

(Hyberts et al. 2012). A maximum protein concentration of 600 μM was used as inter-dimer interactions were evident at higher concentrations (many peaks broadened).

2.5.2 Intensity and CSP analysis of 2D spectra

To measure intensities and positions of peaks in 2D spectra, NMRViewJ was used (Johnson 2004). The following equation was used to calculate chemical shift perturbations (CSPs) between peaks in two spectra:

$$CSP = \sqrt{(\delta_H)^2 + (\delta_N/5)^2}$$

2.5.3 Paramagnetic relaxation enhancement

Spin-label constructs were made with a D3/176del/C137S background and cysteines introduced at positions throughout the NTR, one at a time. Samples were prepared in 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer and 10 mM DTT to fully reduce all cysteines. Reducing agent was then removed from protein samples using a desalting column. Immediately after removing reducing agent, 5X molar excess (1-Oxyl-2,2,5,5-tetramethylpyrroline-3- methyl)-methanethiosulfonate (MTSL) spin label (in DMSO) was added to protein samples and allowed to incubate overnight at 4°C (Figure 2.4). Excess MTSL was then removed from the labeled protein using a desalting column. The resulting NMR samples contained ~400 μM protein. 2D ^{1}H-^{15}N HSQC-TROSY spectra were collected for each spin-labeled mutant protein. The unpaired electron of the spin label was then quenched in each sample by addition of ascorbate (5 mM final concentration). Identical spectra were collected for each quenched sample for intensity comparison between quenched and unquenched spectra.

2.5.4 Methyl-TROSY

Ile-δ-methyl ^{1}H-^{13}C labeled protein was buffer exchanged into D_2O-based 50 mM sodium phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.5 buffer. The limited amount of each protein

was concentrated to at least 300 μM for maximum signal-to-noise. ^{1}H-^{13}C-HMQC spectra were collected in the "methyl" region, centered around 1 ppm (^{1}H) and 12 ppm (^{13}C), over the course of several days at 37°C.

Table 2.1- HSPB1 constructs used in this thesis. Mutations are listed for each region of the protein- NTR, ACD, and CTR.

	Mutant name	NTR mutations	ACD mutations	CTR mutations
Mutation of interacting regions	Single phosphorylation mimic	S15D, S78D, or S82D	-	-
	Double phosphorylation mimic	S15D/S78D, S15D/S82D, or S78D/S82D	-	-
	D3	S15D/S78D/S82D	-	-
	GXG	-	-	I179G/I181G/V183G
	Single phosphorylation mimic/GXG	S15D, S78D, or S82D	-	I179G/I181G/V183G
	D3GXG	S15D/S78D/S82D	-	I179G/I181G/V183G
HDXMS constructs	C137S	-	C137S	-
	D3GXG/C137S	S15D/S78D/S82D	C137S	I179G/I181G/V183G
NMR constructs	D3/176del	S15D/S78D/S82D		Truncate after S176
	D3/176del/ C137S	S15D/S78D/S82D	C137S	Truncate after S176
	Spin label on D3/176del/ C137S	T2C, S65C, S74C, OR T91C S15D/S78D/S82D	C137S	Truncate after S176
Disease mutant constructs	Disease mutant	G34R, P39L, E41K, OR G84R	-	-
	Disease mutant/ fluorophore labeled	G34R, P39L, E41K, OR G84R	C137S	T174C
	Disease mutant/ C137S	G34R, P39L, E41K, OR G84R	C137S	-
	Disease mutant/ D3GXG	G34R, P39L, E41K, OR G84R S15D/S78D/S82D		I179G/I181G/V183G
	Disease mutant/ D3GXG/C137S	G34R, P39L, E41K, OR G84R S15D/S78D/S82D	C137S	I179G/I181G/V183G

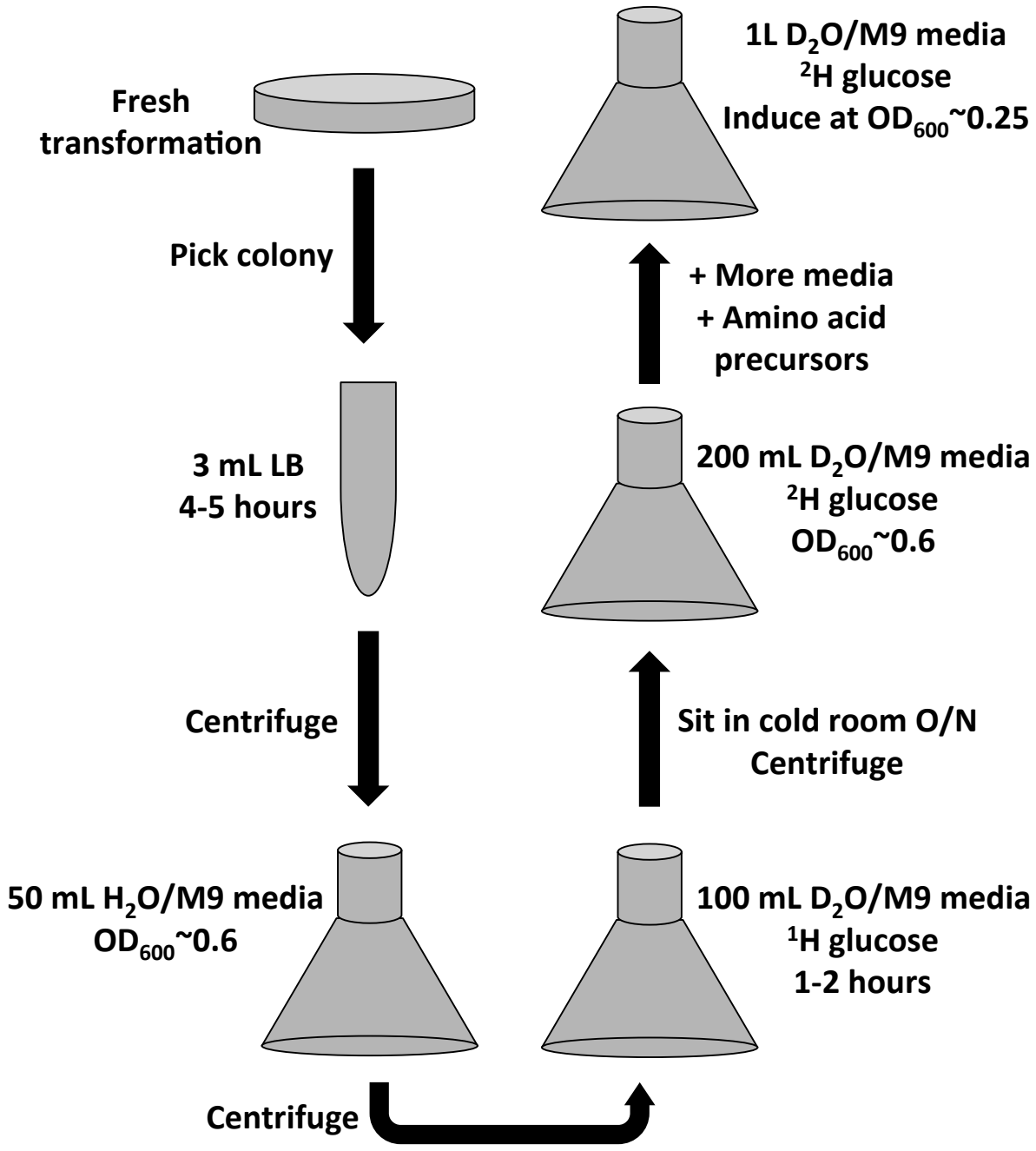


Figure 2.1- Schematic of procedure for making methyl-labeled perdeuterated protein. There are numerous steps to acclimate *E. coli* cells to fully deuterated media. Acclimation is accomplished in stages with increasing volumes of media. The exact schematic is for methyl-labeling perdeuterated protein, but the same process is used for uniformly labeled perdeuterated proteins except that amino acid precursors are not added and the final culture is grown to a higher optical density before induction.

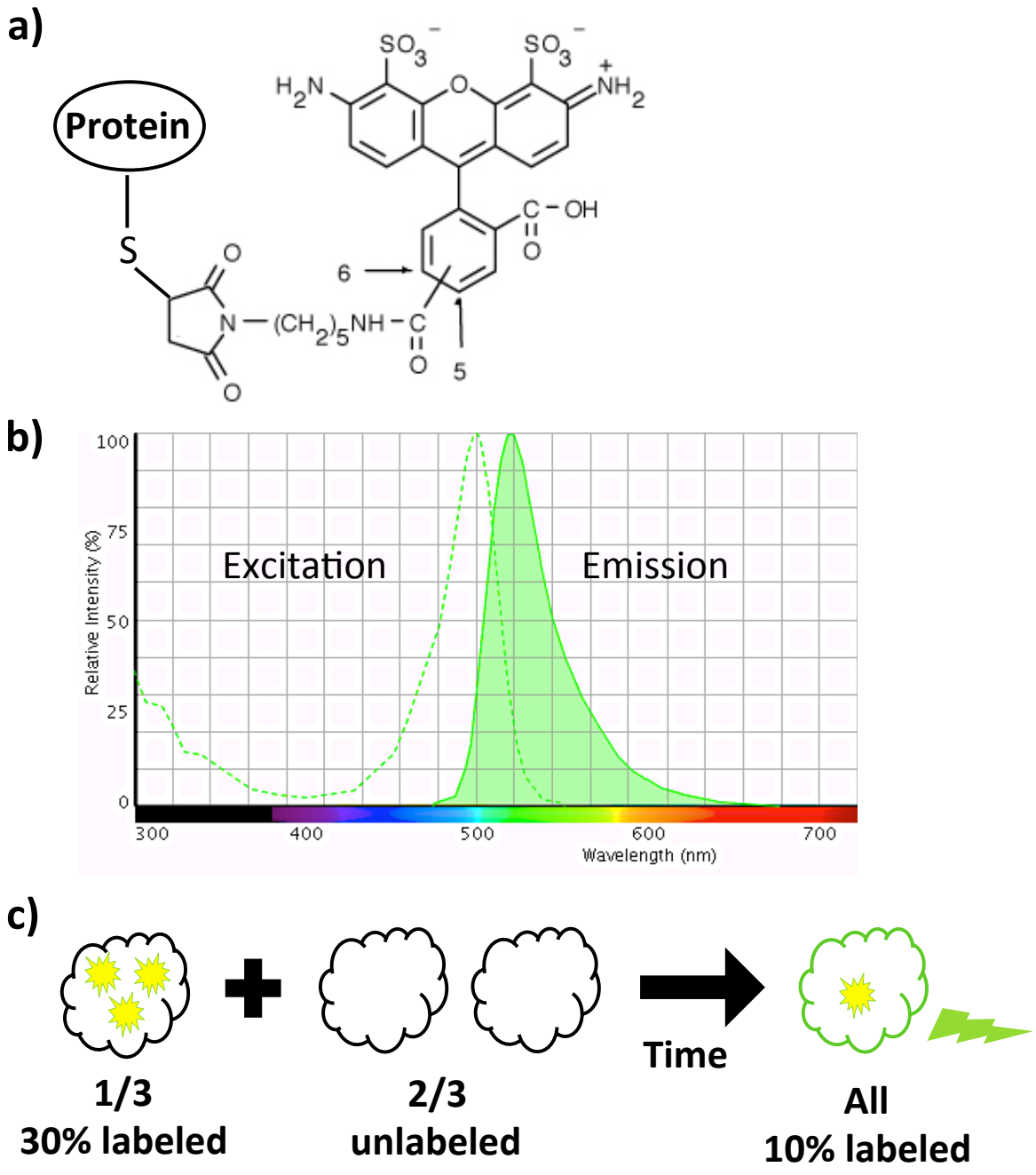


Figure 2.2- Properties of the fluorophore used for subunit exchange experiments. Alexa Fluor 488 maleimide (Thermo Fisher) was attached to cysteines on HSPB1. (a) Chemical structure of Alexa Fluor 488 maleimide and attachment point to a cysteine on a protein. The cysteine-maleimide bond is resistant to some reducing agents, such as TCEP. (b) Excitation and emission spectra for the fluorophore, showing substantial overlap (taken from Thermo Fisher Spectra Viewer). (c) Experimental design for measuring subunit exchange among sHSP oligomers. At high local concentrations (in an oligomer), fluorescence is quenched via homo-FRET. Fluorophore-labeled oligomers are mixed with unlabeled oligomers and the resulting dequenching is measured with time.

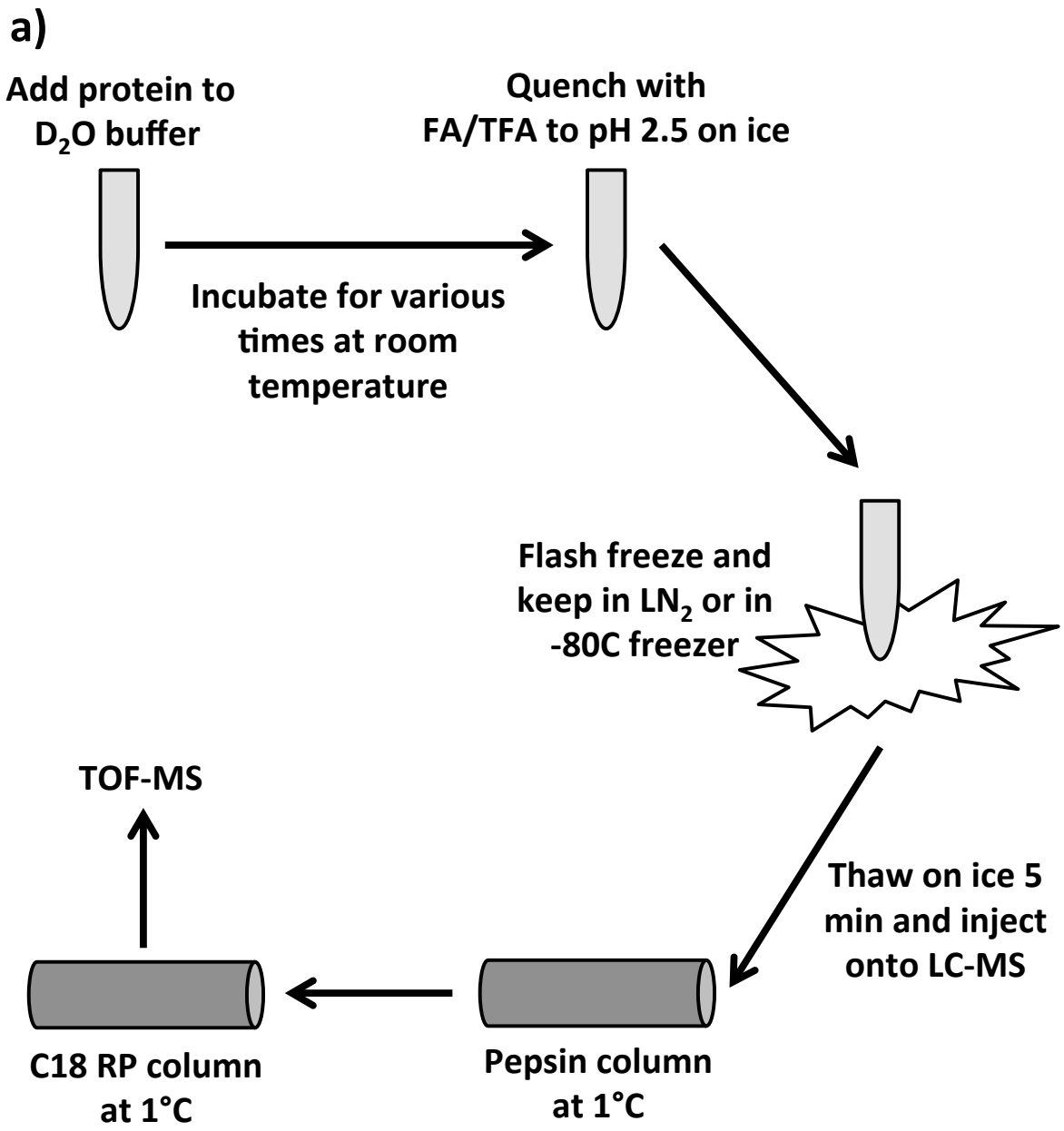
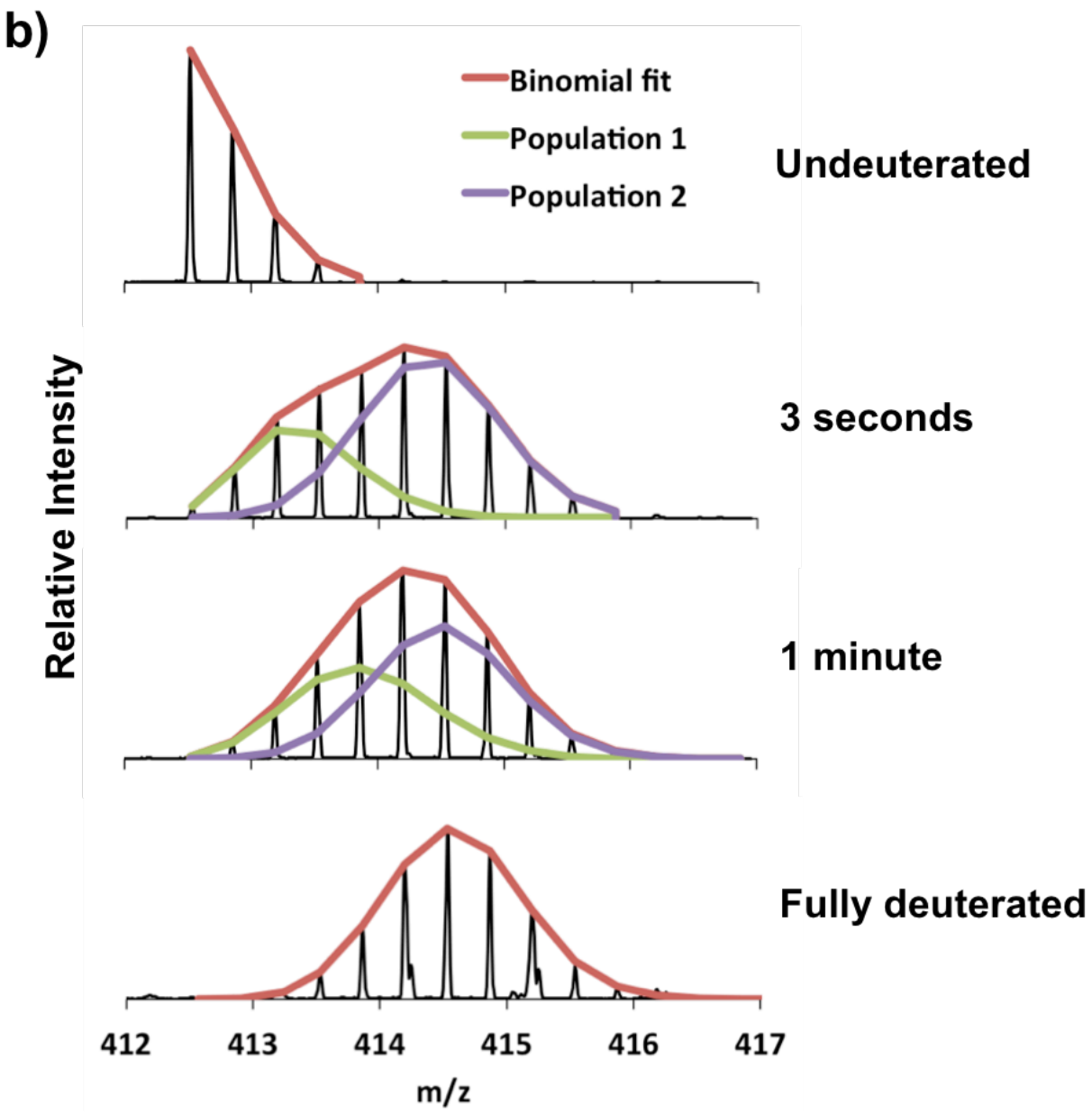


Figure 2.3- Schematic for HDXMS experiment and analysis of bimodal peptides. (a) Protein is incubated in D_2O-based buffer for different lengths of time. The reaction is quenched with acid on ice and flash frozen. Quenched protein samples are digested along a pepsin column and resulting peptides are separated along an acetonitrile gradient on a reverse-phase C18 column. Peptides are then analyzed by TOF-MS. (b) For many sHSP-derived peptides, bimodal distributions were observed and spectra were deconvoluted into two populations with different deuteration levels.


b)
Binomial fit
Population 1
Population 2
Undeuterated
3 seconds
1 minute
Fully deuterated
Relative Intensity
412
413
414
415
416
417
m/z

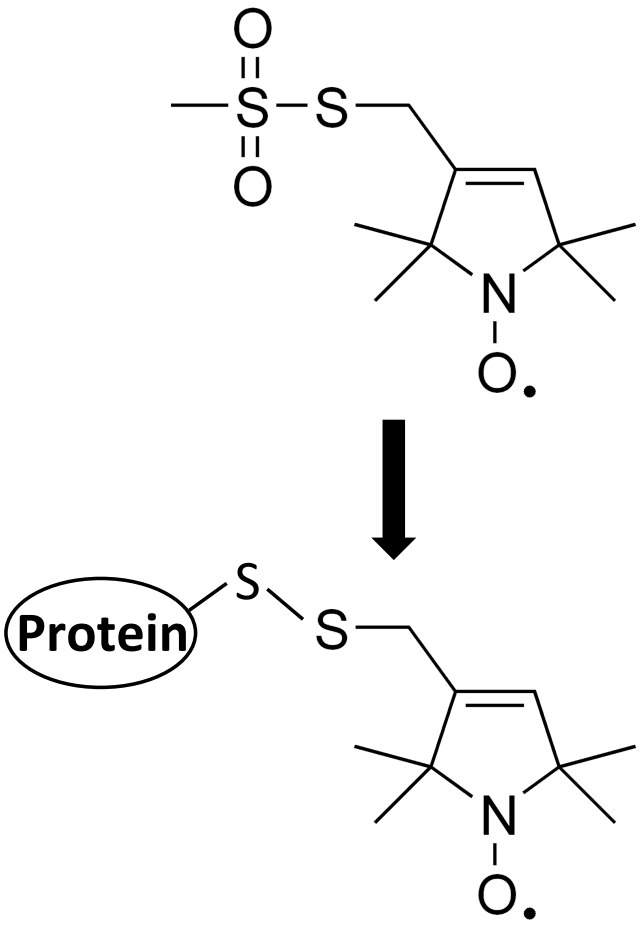


Figure 2.4- MTSL spin label used for NMR-PRE experiments. The unpaired electron of the nitroxide is sterically protected by flanking methyl groups. The thiosulfonate ester reacts with a cysteine on a protein, leaving behind sulfinic acid. The resulting disulfide bond is vulnerable to any reducing agent.

3. HSPB1 OLIGOMERIZATION IS MODULATED BY THREE TYPES OF INTERDEPENDENT INTERMOLECULAR INTERACTIONS

sHSPs that form large oligomers have three major types of intermolecular interactions- ACD-ACD, ACD-NTR, and NTR-NTR. All three types of interaction have been demonstrated or implied in HSPB1 oligomers. The following is an examination of how perturbations to each of these types of interaction structurally and functionally impact HSPB1. Combinatorial effects of perturbing mutations are also explored to understand the interplay of interaction types.

3.1 Phosphorylation-mimicking mutations in the NTR alter local structure and lower oligomeric propensity

3.1.1 WT and the triple phosphorylation mimic ("D3")

There are three conserved (among orthologs) phosphorylation sites in HSPB1, all located within the NTR (Figure 1.4- S15, S78, and S82). Numerous studies have demonstrated for HSPB1 and other human sHSPs that phosphorylation of these sites or phosphorylation-mimicking mutations (to Asp or Glu) shift the average oligomeric size toward smaller states. The triple phosphorylation mimic ("D3"- S15D, S78D, and S82D) is used to mimic a hyper-phosphorylated state of HSPB1.

The quaternary structure, or oligomeric size, of sHSPs is routinely characterized by size-exclusion chromatography (SEC). In this method, molecules of larger hydrodynamic radii elute earlier than those with smaller radii. Therefore SEC does not provide an actual measurement of the mass of a protein or complex but instead depends upon shape of the molecule. SEC can be used to compare sizes (implying number of subunits) of different sHSP oligomers with the assumption that the oligomer shape and density of subunits (loosely or tightly packed in an oligomer) have not changed dramatically. Figure 3.1a shows SEC elution profiles of WT and D3

HSPB1 oligomers that are injected onto a column at high (100 μM, monomer concentration) and low (20 μM) concentrations. It should be noted that protein samples become diluted while traveling through a column. As HSPB1 oligomers are dynamic, they rapidly re-equilibrate during SEC. Therefore the HSPB1 elution profiles do not precisely measure the states of these proteins at the injected concentrations but rather represent their behavior at high concentrations approaching 100μM and at lower concentrations approaching 20μM. WT oligomers elute early (~13 mL) at both high and low concentrations, and at lower concentrations there is only a slight shift right towards slightly smaller species. Therefore the average oligomeric size is relatively "stable" at these concentrations common for SEC. The D3 mutant oligomers elute much later (~16 mL) relative to WT. At the lower concentration, the D3 profile shifts moderately to the right towards smaller species. Therefore the D3 mutant undergoes changes in quaternary structure within this concentration range. For both proteins, the eluted peaks are very broad, which indicates the presence of a range of differently sized species. Both WT and D3 oligomers, though different in average oligomeric size, show a high level of polydispersity (a range of multimers coexisting) under these conditions.

The overall secondary structure of WT and D3 HSPB1 oligomers was probed by circular dichroism (CD) spectroscopy. Far-UV CD provides an average of secondary structure in a protein, with distinct spectral features attributable to α-helical, β-sheet, and random coil structure. While it is possible to quantify structural features for small proteins and peptides, analysis is far less accurate for proteins of longer sequences. Each secondary structure type differs in the maximum intensity of mean residue ellipticity (MRE, normalized for protein concentration and number of residues). A protein or peptide that is primarily helical will yield a CD profile with minima reaching -30,000 MRE. Conversely, β-sheet and disordered proteins

will generally not reach minima below -10,000 MRE. Additionally, α-helical and β-sheet profiles sharply increase around 200 nm (very positive), while the minimum for random coil (-10,000 MRE) is at 200 nm. Figure 3.1b shows far-UV CD profiles of WT and D3 oligomers and the truncated ACD-only construct (dimer) at a concentration of 20 μM (protomer concentration). ACD-only HSPB1 has a typical β-sheet profile (minimum around 218 nm), consistent with known ACD structures, including solved structures of HSPB1. Full-length WT HSPB1 appears to have a mixture of all three secondary structure types (α, β, and random coil). The ellipticity increases in absolute intensity and reaches local minima around 208 nm and 220 nm, typical of α-helical structure. This observation is consistent with predicted helical structure in the NTR and modeling of helices in HSPB5 structural models (Braun et al. 2011; Jehle et al. 2011). The WT profile only reaches a minimum of -3,000 MRE around 220 nm, indicating that there is not a large proportion of α-helical structure. The very small intensity at 200 nm (500 MRE) suggests there are disordered regions negating the expected high intensity at this wavelength.

The D3 mutant is different throughout the spectrum relative to WT protein. D3 shows a very distinct positive signal around 230 nm, a feature that is not attributable to any standard secondary structure formation. Instead, this unusual feature is more likely explained by formation of an exciton couplet between two tryptophan side-chains (Grishina and Woody 1994). Specifically, aromatic rings in close proximity to each other can interact and yield this type of CD signal. HSPB1 has six tryptophans (bold in Figure 1.4), five of which are in the NTR with the remaining tryptophan at the very beginning of the ACD (W95). Therefore, most likely there is a rearrangement of tryptophans in the D3 mutant such that two become close in proximity in the smaller oligomeric state. The local structural implications of a possible exciton couplet are explored and discussed further in the next chapter. The increase in ellipticity at 200

nm for D3 relative to WT suggests that there is a loss of disordered structure and gain of α or β structure. There were no notable changes in the CD profiles of either WT or D3 protein at higher concentrations with respect to profiles at lower concentrations (data not shown).

The triple phosphorylation-mimicking mutant of HSPB1 (D3) forms smaller oligomers than WT, but there is still polydispersity in the range of oligomeric species observed. The phosphorylation mimicking mutations or the induced change in quaternary structure results in altered secondary structure relative to WT. Furthermore, at least two tryptophans adopt a different orientation in D3 oligomers and provide a structural feature that is easy to monitor. The underlying cause (mutation of an involved residue or oligomeric state) of changes in local structure is determined in the following sections.

3.1.2 Single phosphorylation mimics

To determine if global and local structural changes are dependent upon specific phosphorylation-mimicking mutations, single phosphorylation mimics (S15D, S78D, or S82D) of HSPB1 were analyzed by SEC and CD. Figure 3.2a shows SEC elution profiles of each mutant at high and low concentrations (injected at 100 μM and 20 μM). At the higher concentration, all single phosphorylation mimics elute similarly to WT (~13 mL) but with a later-eluting shoulder. At lower concentration, single phosphorylation mimics predominantly elute later (~16 mL), closer to the D3 profile, with a substantial shoulder eluting earlier (~13 mL). The simplest explanation for these results is that phosphorylation-mimicking mutations induce dissociation of oligomers into a distinct class of smaller oligomers (D3-like). The exchange between WT-like and D3-like oligomeric states is very slow as it is observable by SEC, where an experiment takes ~30 minutes at room temperature. In general, if exchange among states is slower than the time required to take a measurement by a particular method,

distinct states can be observed. Additionally, all single phosphorylation mimics have very similar SEC profiles and presumably have similar oligomeric properties.

CD profiles of single mutants are shown in Figure 3.2b. Each single phosphorylation mimic at both high (not shown) and low concentrations shows little change in their CD profile relative to WT protein. Even though the SEC results show that a proportion of single phosphorylation mimics exist in a smaller D3-like state, there is no positive peak in their CD spectra at 230 nm as observed for D3. Together the data indicate that there are no detectable local structural changes with single phosphorylation mimicking mutations, despite partial oligomer dissociation toward a D3-like state.

3.1.3 Double phosphorylation mimics

To further probe specific effects of phosphorylation-mimicking mutations, the three possible combinations of double phosphorylation-mimicking mutations (S15D/S78D, S15D/S82D, or S78D/S82D) were analyzed by SEC and CD in an identical manner to the single mutants. Figure 3.3a shows SEC elution profiles for each double mutant injected at 100 μM and 20 μM (monomer concentration). At low concentration, the double phosphorylation-mimics elute late, similar to the D3 mutant (~16 mL). At high concentration, each double mutant elutes broadly with species at WT-like and D3-like volume. However, the relative proportions of large and small oligomers vary among the double mutants. S15D/S78D at 100 μM elutes almost entirely around the later 16 mL position. S78D/S82D at 100 μM shows a slight majority of protein eluting earlier. S15D/S82D elutes with large/small oligomer proportions intermediate relative to the proportions seen in the other two double mutants. Although the same trends are observed for each double mutant, the intensity of the effect is variable depending upon the pair

of mutations. The less pronounced effect on dissociation for the S78D/S82D mutant could be due to redundancy of the two mutations in close proximity.

CD profiles of the double mutants at high and low concentrations are shown in Figure 3.3b. Double phosphorylation mimics partially adopt the D3-like secondary structure profile with the positive 230 nm peak. At lower concentration, where SEC shows predominantly smaller oligomers, each double mutant is highly D3-like in its CD spectrum. At high concentration, the CD profiles of the double mutants are more WT-like. The magnitude of change from WT-like to D3-like secondary structure is different among double mutants and correlates with their ability to cause large oligomers to dissociate into smaller species. Specifically, the S15D/S78D mutant, which has a large proportion of small oligomers at low concentration relative to the other double mutants, has fully D3-like secondary structure at low concentration and partial D3-like structure at high concentration. The S78D/S82D mutant, which has the highest proportion of large oligomers of the three double mutants, has only partial D3-like secondary structure at low concentration and retains WT-like structure at high concentration.

Local structural changes correlate strongly with the average oligomeric size of "phosphorylated" HSPB1, suggesting that small HspB1 oligomers have distinct secondary structure elements relative to large oligomers. Although each phosphorylation site might be associated with different degrees of oligomer dissociation, there appears to be a common mechanism of effect (global and local structural changes) among the phosphorylation sites.

3.1.4 Mixing different phosphorylation mimics

To further demonstrate the independence of specific phosphorylation-mimicking mutations on size effects, different single mutants were mixed and analyzed by SEC. If a

particular phosphorylation site in one protomer were strongly influenced by the same or different phosphorylation site in a neighboring protomer, then mixing the two mutant proteins to form hetero-oligomers would presumably lead to an oligomeric distribution different from that of each individual mutant. For example, if S15D in one protomer experienced charge repulsion from S78D in a neighboring protomer, then mixing these two mutants would lead to further dissociation of the total population. Conversely, if S15D were repulsed by a neighboring S15D, then mixing in a phosphorylation mimic without S15D might cause a "rescue" effect and result in larger oligomers. Figure 3.4 shows the SEC profiles of all three possible single phosphorylation-mimicking mixtures (1:1) injected at 100 μM and 20 μM (total monomer concentration). No substantial shifts in the elution profile were observed for any combination of single phosphorylation mimics. This result suggests that phosphoryl groups or mimicking mutations in different protomers do not interact in a manner that alters the distribution of oligomeric states.

Examination of the local (secondary) and global (quaternary) structure of all combinations of HSPB1 phosphorylation-mimicking mutants reveals a common mechanism of oligomer dissociation and subsequent change in local structure. With an increasing number of mutations, mimicking higher levels of phosphorylation, smaller oligomers are favored. Dissociation of large oligomers is accompanied by local structural changes as observed by the appearance of a distinct CD signal. These effects are generally phosphorylation-site independent, although the extent of structural changes varies slightly among mutants. Hetero-oligomers of different phosphorylation mimics do not exhibit properties different from their homo-oligomeric counterparts, suggesting that sites of mutation (i.e. phosphorylation) do not directly interact among protomers in an oligomer.

3.2 Changes in redox state of the ACD alter oligomeric properties

Among the structurally characterized human sHSPs, it is assumed that the ACD-ACD interaction has the strongest affinity, among the other types of interaction. The affinity of this interaction has been measured for ACD-only HSPB5, which is 2μM at 25°C, and ACD-only HSPB1 is > 90% dimeric at 100 μM, implying a low μM affinity (Rajagopal et al. 2015a and b). HSPB1 has a cysteine in the β6+7 strand that can form a disulfide bond with its symmetry-related neighbor under non-reducing conditions along the β6+7/β6+7 dimer interface. As the redox state of a cell changes with stress, the formation or reduction of the HSPB1 disulfide bond could have physiological relevance. The disulfide bond also provides a useful tool for studying the behavior of HSPB1 with the ACD-ACD interaction interface covalently linked or free to exchange. The default conditions for all experiments presented in this thesis are non-reducing unless specified as having a reducing agent present (DTT or TCEP) or utilizing a C137S mutation to mimic the reduced state of HSPB1. Non-reducing conditions are maintained throughout the purification of HSPB1. The presence of oxidized protein (usually > 90%) is confirmed at the end of the purification by non-reducing SDS-PAGE analysis. Therefore "non-reducing" and "oxidized" terms are used interchangeably below in regards strictly to HSPB1.

To determine the impact of redox state at the dimer interface on oligomeric propensity, several phosphorylation-mimicking mutants of HSPB1 were analyzed by SEC under reducing (with DTT) and non-reducing conditions. Figure 3.5 shows SEC profiles of representative single and double phosphorylation mimics (S15D and S15D/S78D) under reducing and non-reducing ("oxidized") conditions. The injection concentrations that yielded the largest elution shifts are shown. The intermediate phosphorylation mimics under reducing conditions show a dramatic shift in their distribution equilibrium of oligomeric states toward larger, WT-like oligomers

(~13-14 mL). In contrast, WT HSPB1 does not appear to substantially change in its oligomeric propensity between reduced and oxidized states under these conditions (data not shown). Therefore the ability of HSPB1 subunits to exchange at the ACD-ACD dimer interface under phosphorylation mimicking conditions leads to a larger average oligomeric state. Under stress conditions that generate phosphorylated HSPB1, the distribution of oligomers may depend on the oxidation state of the cell, which also changes under stress. This is a possible case of interplay between stress-regulated structural transitions in HSPB1.

3.3 Disruption of the CTR IXI motif is associated with larger, more polydisperse oligomers

The ACD of HSPB1 has been shown to bind a peptide consisting of the HSPB1 CTR sequence, including the IXI motif, via its β4/β8 groove (Hochberg et al. 2014). IXI motifs have been shown to bind the same groove in several other sHSPs, implying a conserved interaction (Delbecq et al. 2012; Sluchanko et al. 2017). However, it has been shown in HSPB5 that only a small population of the CTRs in an oligomer is in a bound state (Baldwin et al. 2012). Mutation of isoleucines and valines in the IXI motif to glycines inhibits this ACD-CTR interaction ("GXG" mutation).

To probe structural changes in HSPB1 when the ACD-CTR interaction is perturbed, the GXG mutation was introduced into the IXI motif. HSPB1 contains an extended IXI motif, therefore the mutations I179G/I181G/V183G were used. Figure 3.6a shows the SEC profile of GXG HSPB1 relative to WT, injected at 50 μM (monomer concentration). The GXG mutant elutes early (~11-12 mL) relative to WT, indicating formation of oligomers with larger hydrodynamic radii on average. The increase could be due to a change in shape and/or number of subunits in the average oligomer. The GXG peak is even broader than the WT profile, suggesting that an even greater range of oligomers are formed. Figure 3.6b shows the far-UV

CD profile of the GXG mutant relative to WT HSPB1. The GXG mutant shows a general decrease in ellipticity in its CD spectrum relative to WT. Its CD profile is more similar to that of the ACD-only construct, suggesting that there is a loss of helical structure in the NTR and/or CTR. In order to visualize the shapes of GXG oligomers, negative-stain electron microscopy (EM) was used to compare WT and GXG oligomers. This study was limited to an EM-class project and WT micrographs were collected and processed by a classmate. Representative micrographs and class averages are shown in Figure 3.6c. WT HSPB1 forms relatively spherical oligomers of ~15nm diameter, consistent with previous EM reports (Haley et al. 2000). The GXG mutant forms elongated oligomers of highly variable length but consistent width. The width of the GXG oligomers is roughly the same diameter as that of WT oligomers.

These data support a model that loss of the ACD-CTR interaction causes the system to "lose control" of oligomerization along at least one dimension. The remaining ACD-ACD and NTR-NTR interactions hold oligomers together and define certain dimensions and curvature of the oligomers, as seen by the consistent width between WT and GXG oligomers. Binding of the IXI motif is likely responsible for constraining HSPB1 oligomeric size. The reduced helical structure in the GXG mutant suggests that NTRs in these oligomers adopt altered local structure.

3.4 Combination of phosphorylation-mimicking and CTR mutations

A major goal of this thesis work was to identify a full-length mutant of HSPB1 that is amenable to conventional solution-NMR studies (described in the next chapter), i.e. small and monodisperse. The D3 mutant forms a range of small oligomers presumably held together by ACD-ACD and ACD-CTR interactions. To try to obtain a more monodisperse protein, phosphorylation-mimicking mutations and the GXG mutation were combined in a single construct. Figure 3.7a shows the SEC profile of the D3 mutations coupled with GXG mutations

("D3GXG"). D3GXG elutes as a monodisperse peak (sharp and symmetric) slightly later than small D3 oligomers. Even at high concentrations (injected 200 μM), this peak remains monodisperse. Multi-angle light scattering (MALS), which provides a measurement of the protein mass, coupled with SEC was used to confirm that the peak corresponds to a dimeric species (Figure 3.7b). Figure 3.7c shows the far-UV CD profile of D3GXG relative to WT, D3, and GXG forms of HSPB1. D3GXG has a nearly identical CD spectrum to D3 oligomers, including the unusual positive 230 nm peak. Thus addition of the GXG mutation does not appear to affect the average secondary structure. Therefore, D3GXG represents an extreme, completely dispersed dimeric form of phosphorylated HSPB1.

To see if the GXG mutation could promote predominantly dimeric species of intermediate phosphorylation mimics, single phosphorylation-mimicking mutations combined with the GXG mutation were also analyzed by SEC and CD. Figure 3.8a shows SEC profiles of S15D/GXG, S78D/GXG, and S82D/GXG mutants injected at 100 μM and 20 μM (monomer concentration). Each mutant eluted at distinct volumes, suggesting that two very different oligomeric states coexist for the combined mutants. Specifically, both large GXG-like and small D3-like oligomers or even more monodisperse D3GXG-like dimers are observed. As expected from study of the intermediate phosphorylation mimics, at lower concentrations the equilibrium of phosphorylation/GXG mutants shifts toward smaller species. However, there is a more pronounced ranking of dissociation effects depending on the site of phosphorylation. The S15D/GXG mutant has the lowest propensity to form large, GXG-like oligomers, while S82D/GXG has the highest propensity. This is consistent with the ranking of effects observed for double phosphorylation mimics here and in other reports (Jovcevsky et al. 2015). Figure 3.8b shows far-UV CD profiles of the same combinatorial mutants and high and low concentration.

The single phosphorylation mimicking/GXG mutants adopt D3-like secondary structure. The combinatorial mutants have a more D3-like CD profile if they form smaller oligomers by SEC. For example, S82D/GXG has a high proportion of large oligomers at high concentration, and the 230 nm peak in the CD spectrum is less pronounced at high concentration. These results are consistent with the correlation of size and secondary structure of intermediate phosphorylation mimics in the previous section.

The chaperone activity of mutants where specific oligomeric interactions are perturbed was analyzed using an *in vitro* "chaperone assay". Aggregation of the model client protein α-lactalbumin was triggered by reduction of disulfide bonds and subsequent destabilization of the soluble, folded state. Absorbance at 360 nm increases due to light scattering as proteins aggregate. Absorbance was measured as a function of time in the presence or absence of WT, D3, GXG, D3GXG, S82D, or S82D/GXG forms of HSPB1 (Figure 3.9). WT, D3, GXG, and S82D HSPB1 slightly delay formation of aggregates. The D3 result is consistent with certain previous reports that the extreme phosphorylated-mimicking state of HSPB1 is not as effective a chaperone as intermediate phosphorylation mimics (Rogalla et al. 1999). Combination of D3 and GXG mutations in the D3GXG mutant results in an extremely effective chaperone. Although the representative single phosphorylation mimic S82D does not have increased chaperone activity relative to WT, the S82D/GXG mutant does have increased activity. Importantly, there is no direct correlation between chaperone activity and either oligomeric size or local structure. A more likely explanation is different availability of possible client binding sites in each mutant. In this model, the increased availability of NTR binding sites in a phosphorylated state might only be beneficial for function if the β4/β8 groove is also available (GXG mutant) for client recruitment, and vice versa.

The combination of mutations perturbing NTR-NTR and ACD-CTR interactions results in proteins that favor distinctly different oligomeric states- very large GXG-like oligomers and small D3-like oligomers. Changes in secondary structure correlating with quaternary structure are consistent with results for phosphorylation mimics alone. There is evidently a synergistic effect on chaperone activity when multiple potential client binding sites become exposed. Combination of the triple phosphorylation mimic with GXG mutations yields a monodisperse, highly active dimer that is structurally examined in detail in the next chapter.

3.5 Discussion of the interplay of inter-protomer interactions in HSPB1 oligomerization

3.5.1 Perturbation of NTR-NTR interactions

Phosphorylation is a common post-translational modification (PTM) of many proteins. Kinases attach phosphoryl groups to typically serine, threonine, and tyrosine side chains. The process is reversible by the action of phosphatases. Phosphorylation of a protein can alter its structure, which can have a wide variety of effects depending upon the protein and phosphorylation site. An enzyme might be activated or deactivated upon phosphorylation, or the rate of a reaction might be modulated. For many proteins, and for the case of sHSPs, phosphorylation alters protein-protein interactions. For the cases of sHSP phosphorylation that have been characterized, phosphorylation (or a phosphorylation mimicking mutation) generally increases chaperone activity (or cell viability in *in vivo* assays) and decreases homo-oligomeric propensity. The kinase/phosphatase system provides a rapid means of regulation of protein function. HSPB1 is phosphorylated rapidly (within a few minutes) in response to stress *in vivo* by two stress-upregulated kinases- MAPKAP kinases 2 and 3 (Larsen et al. 1997; Rogalla et al. 1999). Phosphorylation has been observed at all three sites in one subunit *in vivo* under stress conditions, based on the levels of phosphorylation relative to the amount of protein present. For

the case of HSPB1, this form of regulation allows the available pool of HSPB1 proteins in a cell to be activated for chaperone activity when a cell is under stress. Expression of HSPB1 is also upregulated in response to stress, but this is a slower response (several minutes to hours).

Of all biochemical studies of HSPB1, phosphorylation is a widely covered topic. Numerous studies have demonstrated changes in activity and quaternary structure for both phosphorylated and phosphorylation-mimicking (Ser to Asp) forms of HSPB1. The studies presented in this thesis indicate that with increasing phosphorylation-mimicking mutations in a protomer, HSPB1 oligomers shift toward smaller species (Figure 3.10). Rather than a continuum of species, two predominant oligomeric states, WT-like and D3-like, are observed for intermediate phosphorylation mimics. Resolution of the states by SEC indicates that there is very slow exchange between two predominant states. A recent study using native mass-spectrometry (MS) quantified the number of subunits in HSPB1 oligomers and every combination of phosphorylation mimic as a function of concentration (Jovcevski et al. 2015). It was found that the triple phosphorylation mimic (D3) was predominantly dimeric at the concentrations examined (<50 μM). The single phosphorylation mimics most commonly formed 14-mers but still maintained a broad distribution of species. The double phosphorylation mimics adopted an even broader distribution that is shifted toward smaller species. Although the mass spectrometry analysis did not yield two classes of oligomers of different numbers of subunits, their SEC analysis did show similar partitioning of eluting species for intermediate phosphorylation mimics as is demonstrated in this thesis. This discrepancy could indicate very different shapes of oligomers of the same mass, resulting in two peaks by SEC elution. Although a 14-mer sHSP oligomer could be constructed via many different arrangements of subunits, analysis of phosphorylation-mimicking HSPB5 provides insight into possible architectures

(Peschek et al. 2013). The triple phosphorylation mimic of HSPB5 was analyzed by cryo-EM. The protein predominantly adopted 6-mer and 12-mer structures. The hexamer fit well with an arrangement of three ACD-ACD dimers linked via CTR-ACD interactions, with NTRs interacting between only two neighboring protomers. The 12-mer fit well with two of these hexameric rings interacting, presumably via NTRs. HSPB1 when phosphorylated might adopt similar arrangements of protomers. EPR studies of the NTR between WT and D3 oligomers show the NTR to be more flexible in the phosphorylation-mimicking state (McDonald et al. 2012). Therefore, it is likely that HSPB1 when phosphorylated dissociates into smaller oligomers with their NTRs more exposed.

Phosphorylation of the NTR in oligomeric sHSPs generally correlates with both increased NTR exposure and increased chaperone activity. In HSPB5, phosphorylation-mimicking mutants are more active chaperones relative to WT (lengthening the lag phase before aggregation) with some client proteins (Peschek et al. 2013). Intermediate phosphorylation mimics of HSPB1 increase chaperone activity relative to WT for some client proteins (Jovcevski et al. 2015). For these client proteins, the triple mutant (D3) is the most effective chaperone. However, in some studies the D3 mutant is a poorer chaperone relative to WT, while the intermediates remain strong chaperones (Rogalla et al. 1999). Although the chaperone activity of intermediate phosphorylation mimics with our model client, α-lactalbumin, was not explored in-depth in this thesis, the D3 and S82D mutants had only similar activity to WT. It is likely that increased exposure of the NTR presents more hydrophobic binding sites for client proteins. Therefore, if chaperone function of sHSPs requires interaction of the NTR with client proteins, then phosphorylated states would generally be more active. There are many possible reasons for differences in function for certain model client proteins and the extreme phosphorylation-

mimicking state of HSPB1. Different client proteins likely have different preferences for interacting regions of a sHSP (NTR vs. β4/β8 groove). Also, too many or too strong hydrophobic interactions between NTRs and client proteins might promote aggregation and negate chaperone effects. Therefore multiple phosphorylation sites in an oligomeric sHSP provide a means to broadly regulate their function, avoiding sHSP-client interactions when they are not needed.

sHSPs are implicated in numerous other cellular pathways outside of directly aiding in preventing aggregation (cytoskeleton, apoptosis, muscle proteins, etc.). Therefore, phosphorylation of sHSPs might regulate additional functions. Some sHSPs that do not form large oligomers (HSPB6 and HSPB8) also have phosphorylation sites in their NTR. In these cases, phosphorylation must serve a purpose other than control of oligomeric size. Phosphorylated HSPB6 was found to interact with a 14-3-3 protein, a key type of protein for modulating protein-protein interactions (Sluchanko et al. 2017). A reported crystal structure of the HSPB6/14-3-3 protein complex has residues of the NTR of HSPB6, including a phosphorylation site, bound in a pocket of a 14-3-3 protein. Similarly, HSPB8 has been shown to interact with BAG3, a stress-inducible protein important for coordinating many protein-protein interactions (Rauch et al. 2017). In particular, BAG3 brings together and modulates the activity of HSPB8 and HSP70, two different types of chaperone. However, phosphorylated and certain phosphorylation-mimicking mutants of HSPB8 have decreased activity relative to WT protein for some client proteins (Shemetov et al. 2011). Therefore phosphorylation can regulate interactions between sHSPs and aggregating client proteins as well as other cellular targets. For HSPB1, changes in secondary structure correlate with oligomeric size independent of the site of phosphorylation-mimicking mutation. It is unknown if phosphorylation directly promotes

dissociation (e.g. sterics or charge repulsion) and if secondary structure changes are a secondary effect. Conversely, phosphorylation might promote altered secondary structure that can only be accommodated in a smaller oligomeric state (e.g. NTRs are less restrained). Therefore, the mechanism of regulation by phosphorylation could be both indirect, by dissociation of oligomers and a resulting increase in binding surfaces, and direct, by altering the local structure of the NTR to conformations that favor certain interactions.

3.5.2 Perturbation of the inter-subunit ACD-CTR interaction

An IXI motif is found in the CTR of many human sHSPs (HSPB1, 2, 4, 5, and 7) and has been shown to bind the β4/β8 groove of the ACD for some of these sHSPs. A crystal structure of truncated ACD-only HSPB1 shows a peptide mimicking the CTR (including the IXI motif) of HSPB1 binding the β4/β8 groove (Hochberg et al. 2014). Similarly, truncated ACD-only HSPB5 has been shown to bind its CTR-mimicking peptide via the analogous groove and it fails to bind a similar peptide with the IXI motif mutated to GXG (Delbecq et al. 2012). In WT HSPB5 oligomers, the IXI region of the CTR is observed bound to β4/β8 grooves of neighboring protomers by NMR and included in oligomeric models (Braun et al. 2011; Jehle et al. 2011). However, solution- and solid-state NMR studies focusing on the CTR of HSPB5 in the oligomeric context demonstrate that the bound population of IXI motifs varies widely with temperature (Baldwin et al. 2011 and 2012). At sub-freezing temperatures, both bound and unbound states of the CTR are observable by solid-state NMR. At above freezing temperatures, the unbound state is predominantly observed by both solution- and solid-state NMR. Therefore, it is believed that not all types of interaction occur in full for one subunit in an oligomer. Oligomers are instead held together by only a fraction of the possible inter-subunit interactions that have been characterized.

Mutation of the conserved IXI motif in the CTR to GXG in HSPB1 presumably abrogates interaction of the CTR and ACD. Paradoxically, perturbation of this ACD-CTR interaction results in formation of very large oligomers (Figure 3.10) as seen by SEC. This result is consistent with a study examining a truncated form of HSPB1, where residues 182-205 of the CTR were removed (includes the end of the IXI motif), resulting in formation of larger oligomers (Lelj-Garolla and Mauk 2012). There are also disease-associated truncations of the CTR reported for HSPB1 (T174-stop) and HSPB5 (G154-short tail due to frameshift), but there has been no biochemical characterization of these mutants (Datskevich et al. 2012). Negative-stain EM of the GXG mutant shows formation of elongated oligomers of highly variable length. However, the narrow diameter of the elongated oligomers remains the same, ~15 nm, which is consistent with the average diameter of WT HSPB1 oligomers. Therefore, abrogating the ACD-CTR interaction (with the ACD-ACD interaction disulfide linked) might promote NTR-NTR interactions. There is a delicate balance of these types of interactions that maintains a controlled range of oligomeric species. The resulting NTR-NTR interactions in the variably elongated oligomers are likely less restrained and might adopt additional orientations. The GXG oligomers have an average secondary structure different from WT as seen by CD. Specifically, the GXG mutant has more ACD-only-like secondary structure with a loss of helical content. Bound and unbound forms of the unstructured CTR likely would not produce different CD spectra, therefore the change in secondary structure is likely due to altered NTR structure (loss of helical structure) in the elongated oligomers. Although the GXG mutation presumably results in increased availability of β4/β8 grooves, this mutant does not have increased chaperone activity with the model client protein α-lactalbumin. The chaperone activity of the 182-205 truncation mutant

(Lelj-Garolla and Mauk 2012) is temperature dependent for a model client protein, with activity only observed at higher temperatures.

There are several single-point disease-associated mutations reported in the CTR of human sHSPs that might result in similar structural changes to those of the GXG mutant, although few have been characterized biochemically or structurally. Mutations in or near the IXI motif are reported in HSPB1 and HSPB5. In HSPB5, R157H has been reported in patients with cardiomyopathies. The IXI motif of HSPB5 begins at position 159, therefore this mutation could likely impact the ACD-CTR interaction. In HSPB1, the mutations T180I, P182L, P182S, and S187L have been reported in patients with motor neuropathies (dHMN) (Echaniz-Laguna et al. 2017; Datskevich et al. 2012). Two of these mutation sites are within the elongated IXIXV motif at the "X" sites and could affect binding. SEC analysis of two of these mutant oligomers reveals little change in quaternary structure for the T180I mutant and a large increase in average size for the P182L mutant (Chalova et al. 2014). Prolines are often found in the "X" position in sHSP IXI motifs. Prolines are unique relative to other residues in that there is a substantial propensity (up to 15%) for the preceding peptide bond to be in a *cis* conformation. The *cis* conformation is rarely observed in peptide bonds between other residues. *Cis/trans* proline isomerization essentially alters the relative orientation of flanking regions, possibly resulting in large structural changes. *Cis/trans* proline isomerization was reported for P194 in the CTR of HSPB1 oligomers, well beyond the IXI motif (Alderson et al. 2017). Proline isomerization was also observed for P182 in the context of a peptide mimicking the CTR. Therefore, mutation of prolines or nearby residues in the CTR of sHSPs could alter local structural propensities that affect the ACD-CTR interaction.

3.5.3 Combinatorial effects of perturbing NTR-NTR and ACD-ACD interactions

Changes in oligomerization propensity were explored as a function of redox state at the ACD-ACD dimer interface. A cysteine naturally occurs at position C137 in the HSPB1 dimer interface. Under non-reducing conditions, a disulfide bond readily forms and can be observed via non-reducing SDS-PAGE. Although the cytoplasmic environment is generally reducing, certain stressors can change the environment toward more oxidizing. Therefore, the disulfide bond in HSPB1 could provide an additional level of stress regulation. WT oligomers do not substantially change average size between reduced and oxidized states. The most intense effects were seen for intermediate phosphorylation mimics, where multiple distinct oligomeric states are observable by SEC. For single and double phosphorylation mimics, reducing the disulfide bond at the dimer interface shifts the proportions of oligomeric states, favoring the larger state (Figure 3.10). Therefore, exchange at the dimer interface appears to promote NTR-NTR interactions.

Numerous studies have probed the effects of disease-associated mutations and changes in pH at the dimer interface of human sHSPs. In HSPB1, S135F/M, R136W/L, R140G, and K141Q mutations at the dimer interface have been reported in dHMN and CMT-2 patients (Datskevich et al. 2012). In HSPB4, HSPB5, and HSPB8, mutations analogous to R140G in HSPB1 have been reported (R116C/H, R120G, and K141E/T/N, respectively). Many of these mutants have been characterized biochemically and generally show changes in quaternary structure and chaperone activity. Detailed studies on the effects of pH changes on HSPB1 and HSPB5 structure and function showed dramatic global effects resulting from perturbation of the ACD-ACD dimer interface (Clouser and Klevit 2017; Rajagopal et al. 2015b). Specifically, a conserved histidine in the ACD was shown to be a pH-sensitive switch. Slightly lowered pH (in the range of 6.5 to 7.5, consistent with cellular acidosis) destabilizes the ACD-ACD interaction.

Mutating this histidine to lysine mimics a permanently protonated, low-pH state. The resulting chaperone activity of the low-pH-mimicking mutants in the full-length context is increased. For both the mimicking mutants and WT HSPB1 and HSPB5 at low pH, larger oligomers are formed. Therefore multiple types of perturbations to the ACD-ACD interface alter sHSP oligomeric properties and chaperone function.

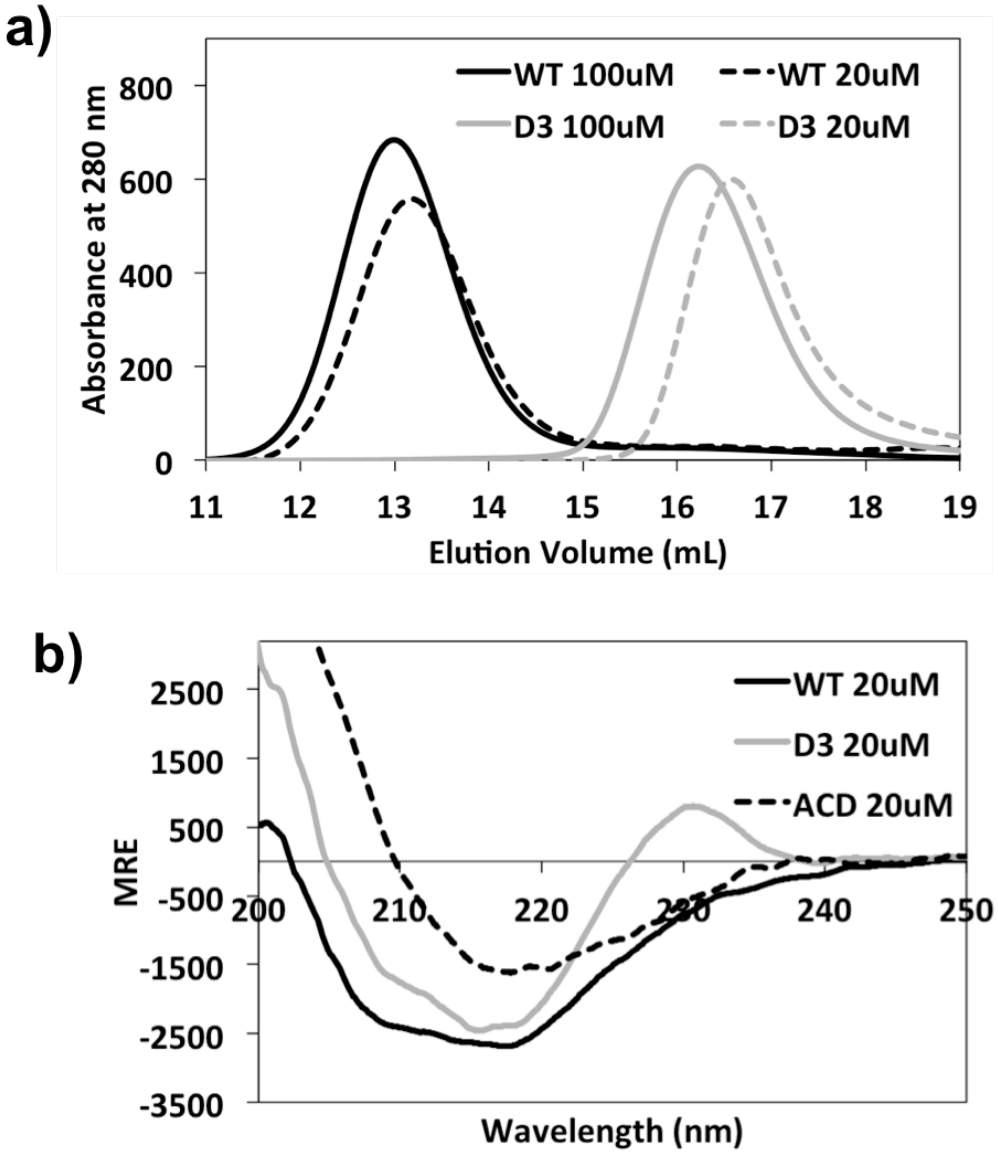


Figure 3.1- SEC and CD profiles of WT and D3 HSPB1 oligomers and ACD-only. (a) 100μM or 20μM samples of WT or D3 protein (monomer concentration), in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer were incubated at 37°C for one hour. Samples were then injected onto an SEC column (24 mL Superose6) at room temperature. WT oligomers elute much earlier than D3 oligomers, indicating the presence of much larger oligomers. (b) 20μM samples of WT, D3, and ACD-only protein in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. The ACD-only has a typical β-sheet profile. WT and D3 show an increase in helical content (low points around 208 and 222 nm) relative to the ACD-only. The D3 mutant shows distinct changes in secondary structure relative to WT, particularly in the appearance of a positive peak around 230 nm.

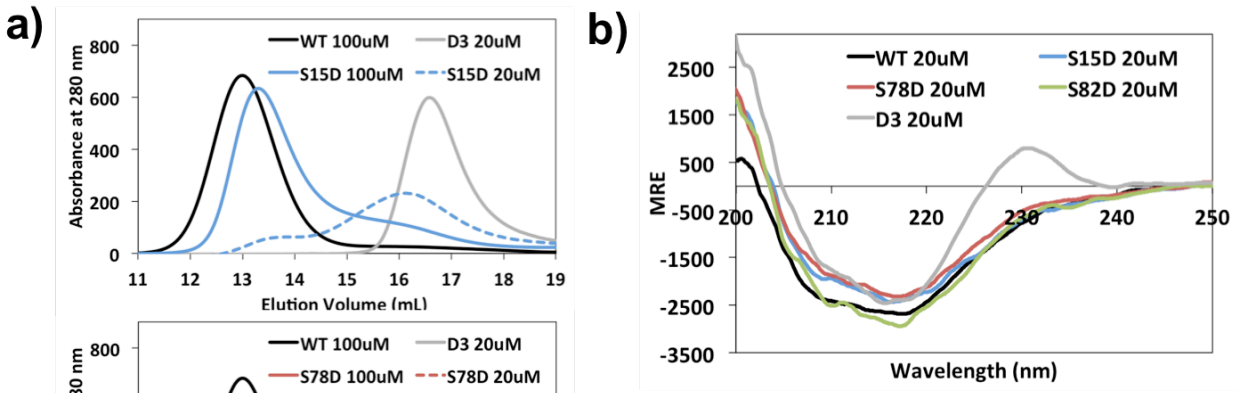


Figure 3.2- Concentration dependence of SEC and CD profiles of single phosphorylation mimics of HSPB1. (a) S15D, S78D, and S82D mutants were incubated at 100μM and 20μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. WT and D3 traces (identical for Figure 3.1a) are shown for comparison. 100μM single phosphorylation mimics elute only slightly later than WT with a small shoulder after. 20μM single phosphorylation mimics elute over a broad volume in two humps- one near WT elution and one near D3 elution. (b) 20μM samples of single phosphorylation mimics in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. Even at 20μM, where smaller oligomers are present, single phosphorylation mimics adopt WT-like secondary structure.

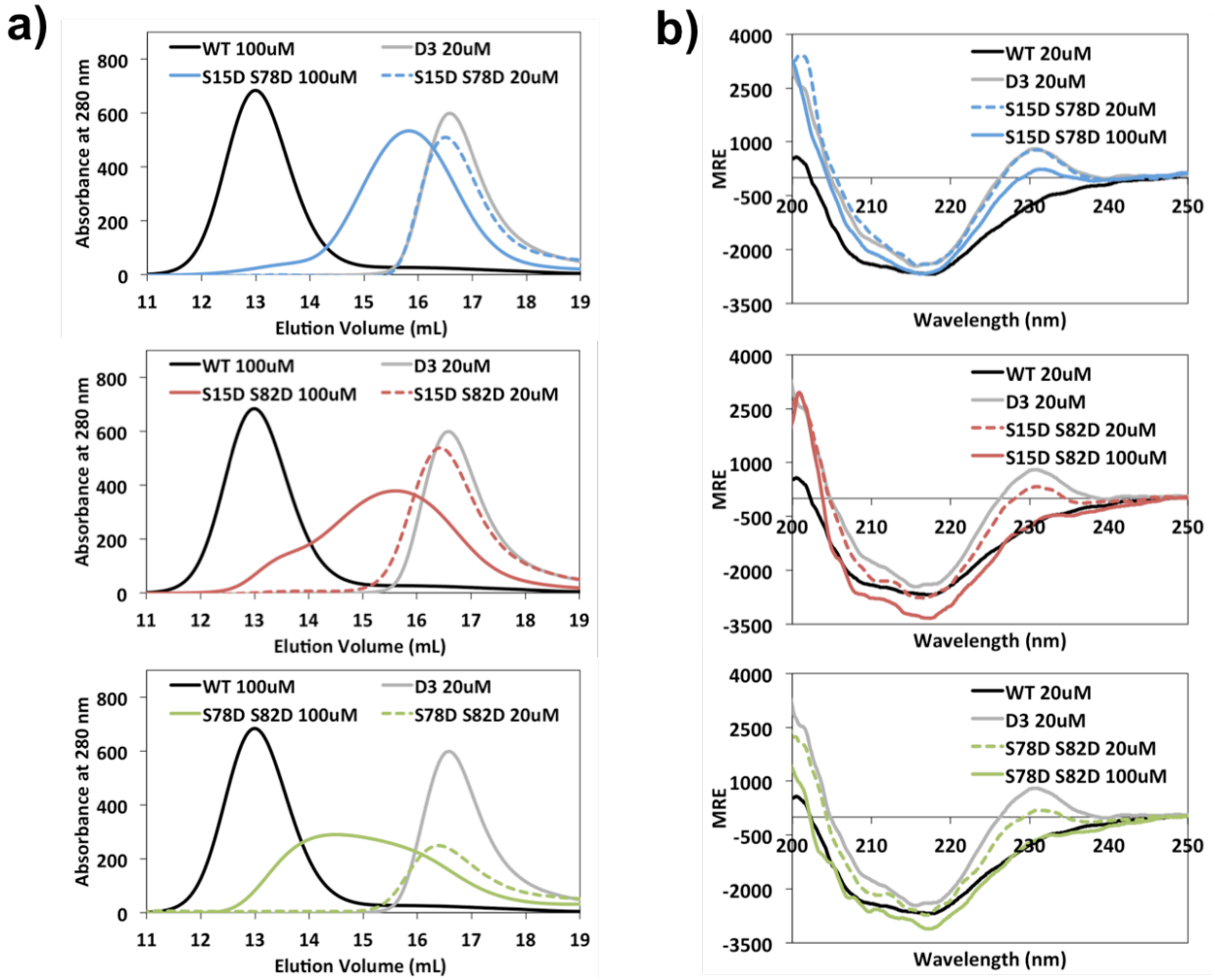


Figure 3.3- Concentration dependence of SEC and CD profiles of double phosphorylation mimics of HSPB1. (a) S15D/S78D, S15D/S82D, and S78D/S82D mutants were incubated at 100μM and 20μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. WT and D3 traces (identical for Figure 3.1a) are shown for comparison. 100μM double phosphorylation mimics elute much later than WT and are closer to the D3 elution but are very broad. 20μM double phosphorylation mimics elute less broad and nearly overlay with the D3 elution profile. (b) 100μM and 20μM samples of double phosphorylation mimics in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. At 20μM, where D3-sized oligomers are present, double phosphorylation mimics adopt D3-like secondary structure.

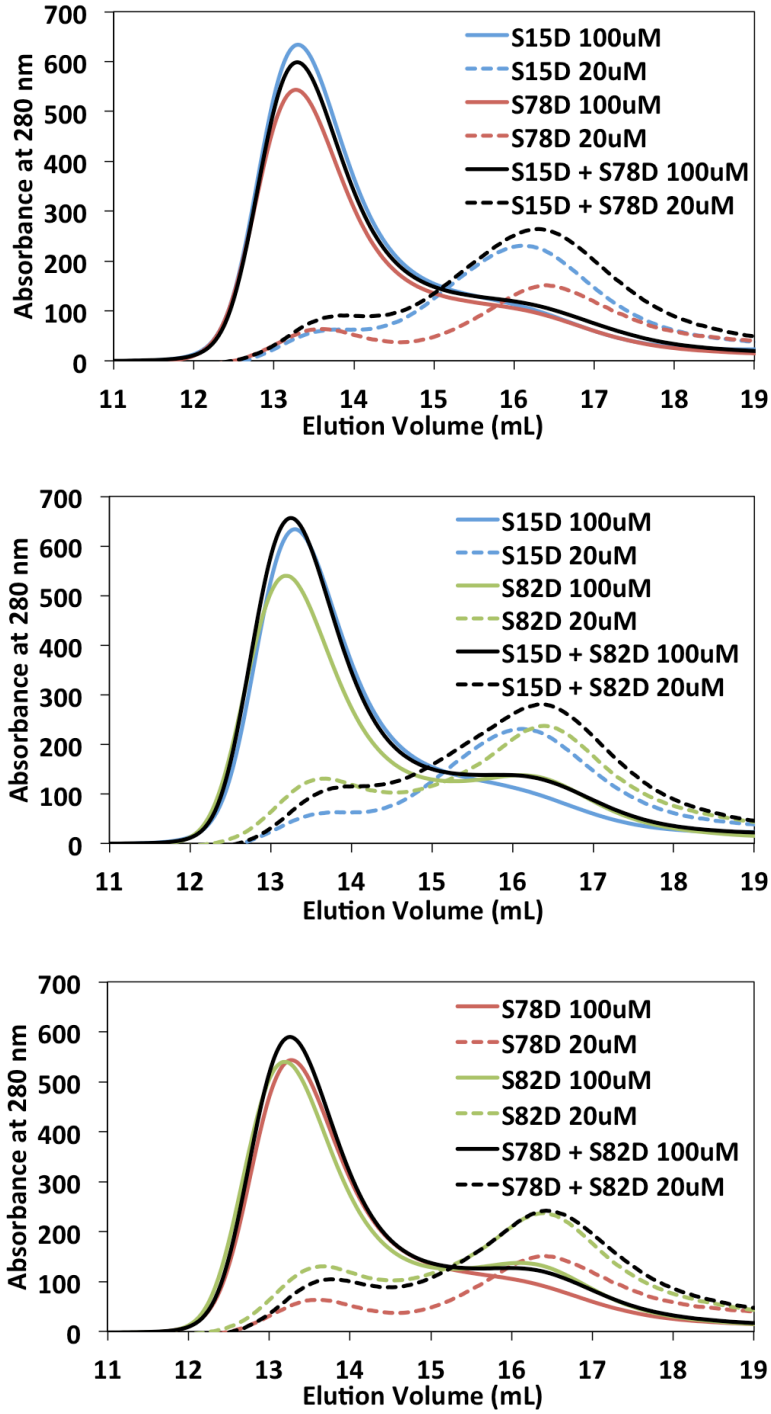


Figure 3.4- SEC profiles of mixtures of single phosphorylation mimics of HSPB1. S15D, S78D, and S82D mutants were mixed in pairs at a 1:1 ratio with final protein concentrations of 100μM and 20μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer. Samples were incubated at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. Single phosphorylation mimic traces (identical for Figure 3.2a) are shown for comparison. For all three mixtures at both high and low concentrations, the elution profiles do not change substantially relative to the unmixed samples.

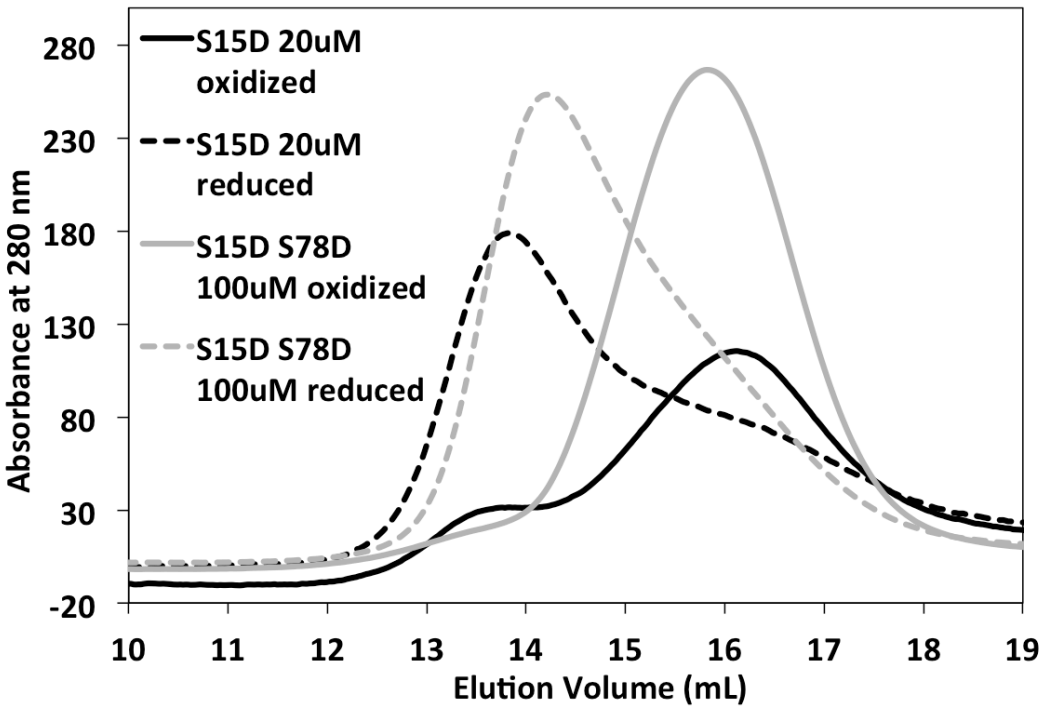


Figure 3.5- Redox effects on SEC profiles of intermediate phosphorylation mimics of HSPB1. 20μM samples of S15D and 100μM samples of S15D/S78D were prepared in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer with (reducing) and without (oxidized) 2mM DTT. Samples were incubated at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature and run with the identical reducing or non-reducing buffer. For both mutants, reducing the protein resulted in a dramatic shift in the elution profile towards larger oligomers.

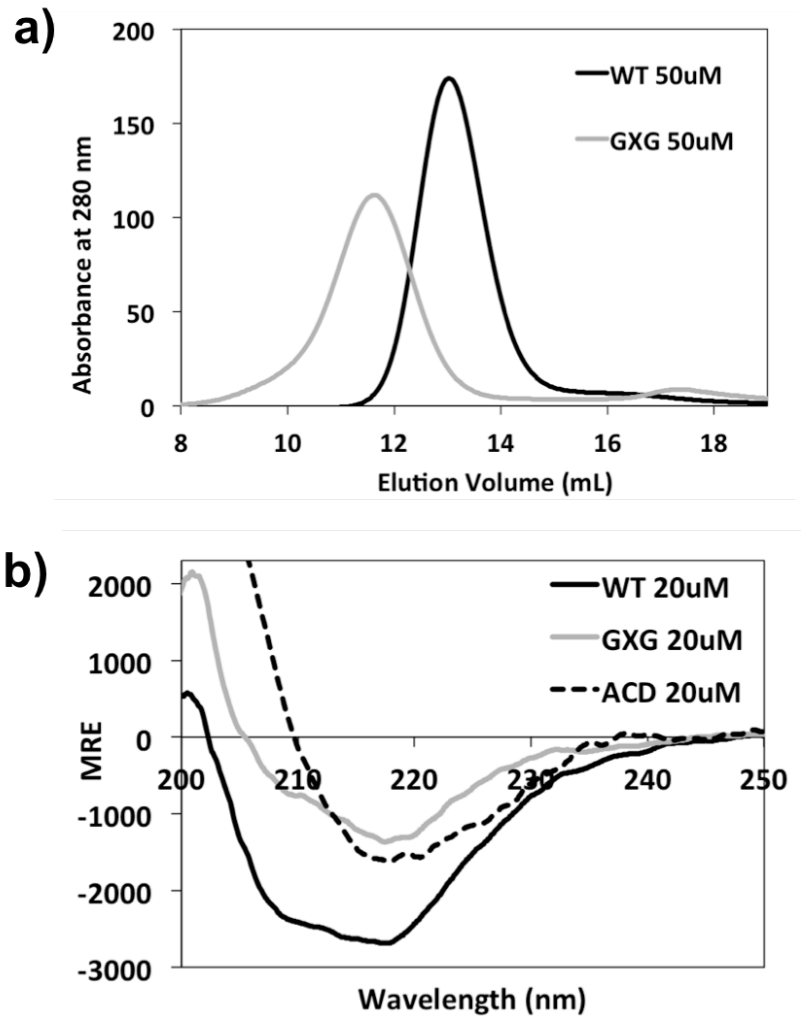
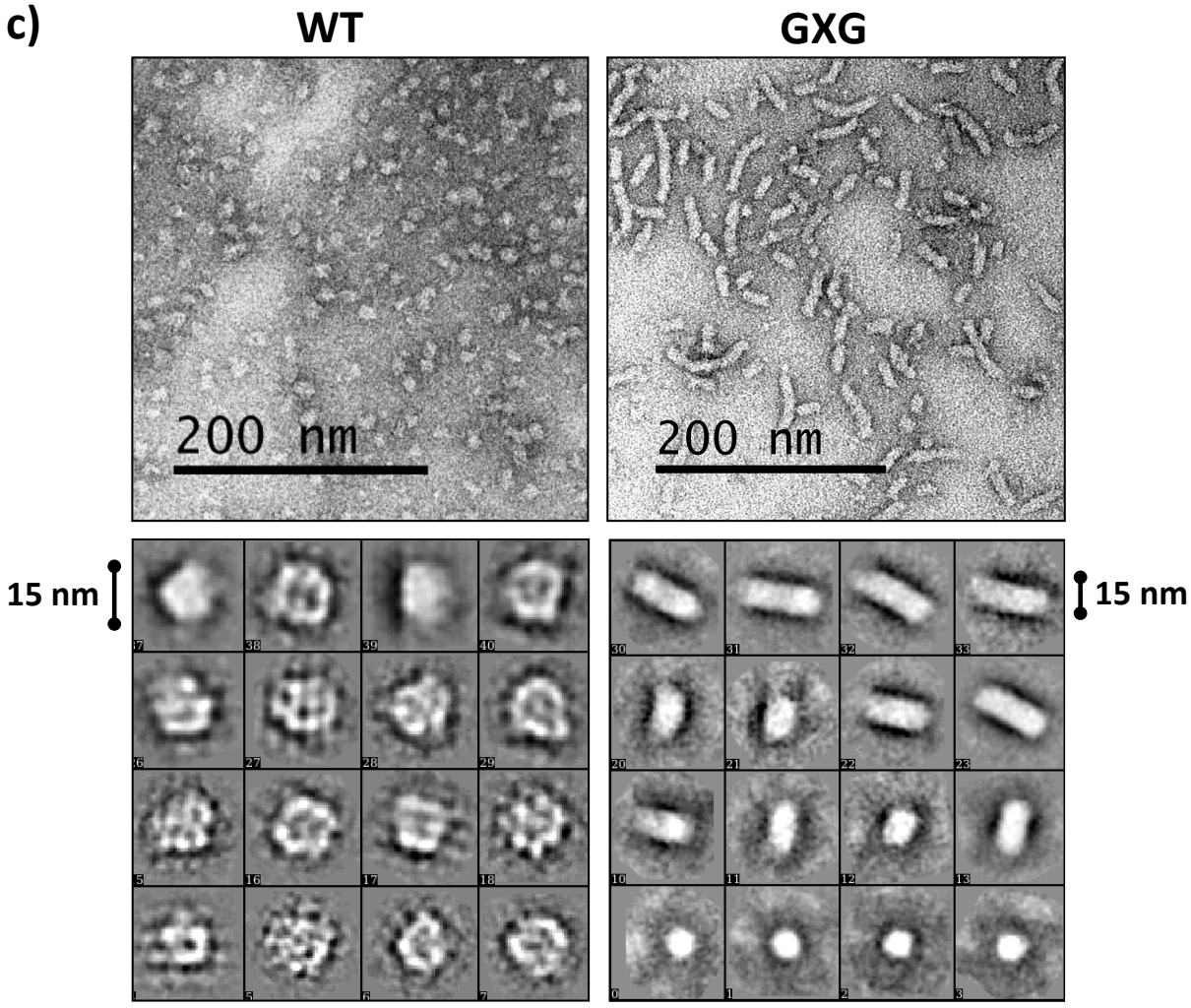


Figure 3.6- Characterization of the GXG mutant of HSPB1 by SEC, CD, and negative-stain EM. (a) The GXG HSPB1 mutant was incubated at 50μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. The sample was then injected onto a SEC column (24 mL Superose6) at room temperature. A WT trace at the same concentration is shown for comparison. The GXG mutant elutes much earlier than WT, indicating formation of larger oligomers. (b) A 20μM sample of the GXG mutant in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer was analyzed by far-UV CD spectroscopy at 20°C. The sample was incubated at room temperature for 3 hours prior to measurement. The GXG mutant has generally decreased ellipticity relative to the WT. Specifically, GXG has a more similar profile to ACD-only, suggesting a loss of helical structure in GXG relative to WT, which is not present in the ACD-only construct. (c) Representative micrographs and class averages of WT and GXG forms of HSPB1 analyzed by negative-stain EM are shown. Samples were prepared and equilibrated at 4μM (monomer concentration) and in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer. WT protein forms fairly spherical but polydisperse oligomers of ~15 nm diameter as previously reported. The GXG mutant forms a range of oligomers ranging from spherical to highly elongated, but all GXG oligomers have an average diameter in one dimension similar to that of the WT.


c)
WT
GXG
200 nm
200 nm
15 nm
15 nm

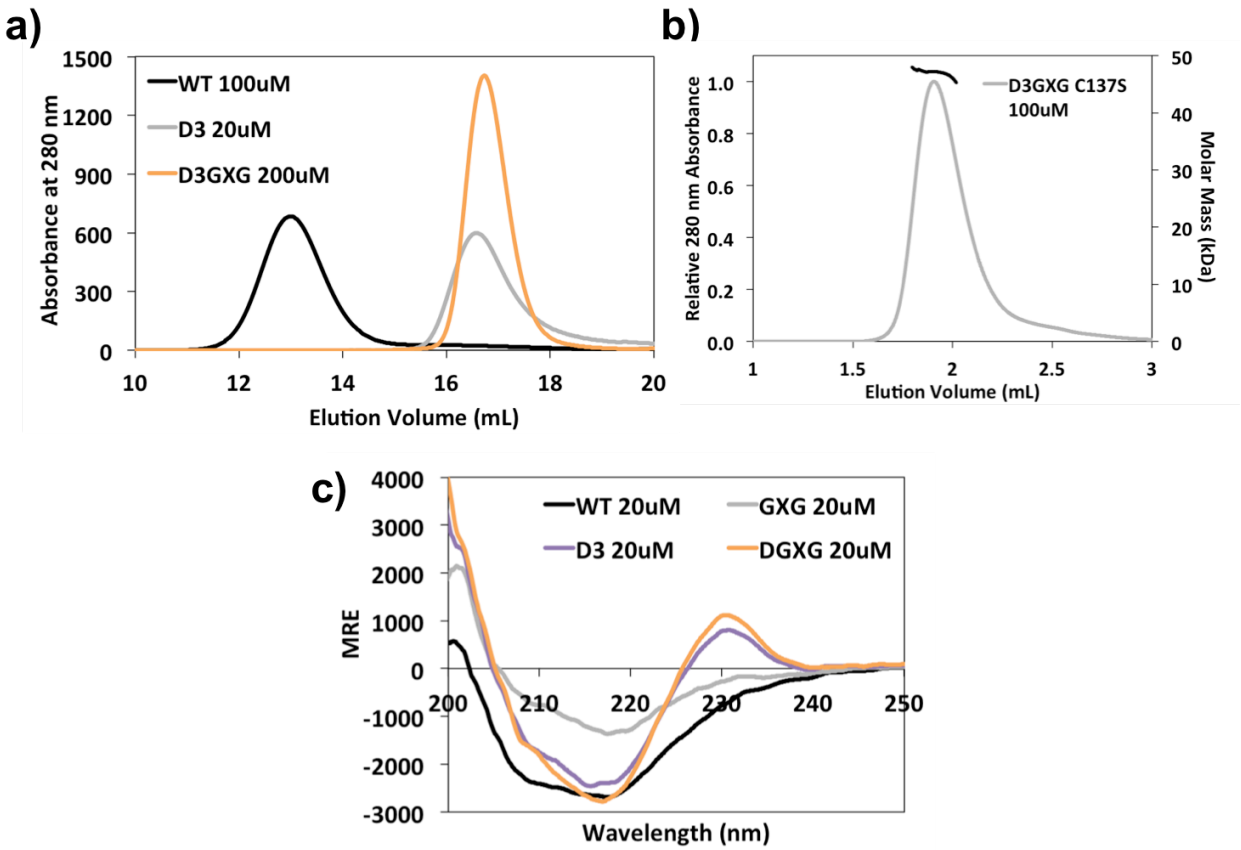


Figure 3.7- SEC, MALS, and CD profiles of the D3GXG mutant of HSPB1. (a) The D3GXG HSPB1 mutant was incubated at 200μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. The sample was then injected onto a SEC column (24 mL Superose6) at room temperature. WT and D3 traces are shown for comparison. The D3GXG mutant elutes much later than WT and slightly later than D3. The peak for D3GXG is also very sharp, indicating monodispersity. (b) The D3GXG HSPB1 mutant was incubated at 100μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. The sample was then injected onto a SEC column (2.4 mL Superose6) at room temperature. Eluted sample was analyzed by absorbance, multi-angle light scattering (MALS), and change in refractive index. Molar mass calculation based on the relationship between light scattering and concentration (from change in refractive index) is plotted in black. (c) A 20μM sample of the D3GXG mutant in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer was analyzed by far-UV CD spectroscopy at 20°C. The sample was incubated at room temperature for 3 hours prior to measurement. The D3GXG mutant has a highly similar profile to D3, indicating that these two mutants have the same average secondary structure, regardless of the GXG mutation.

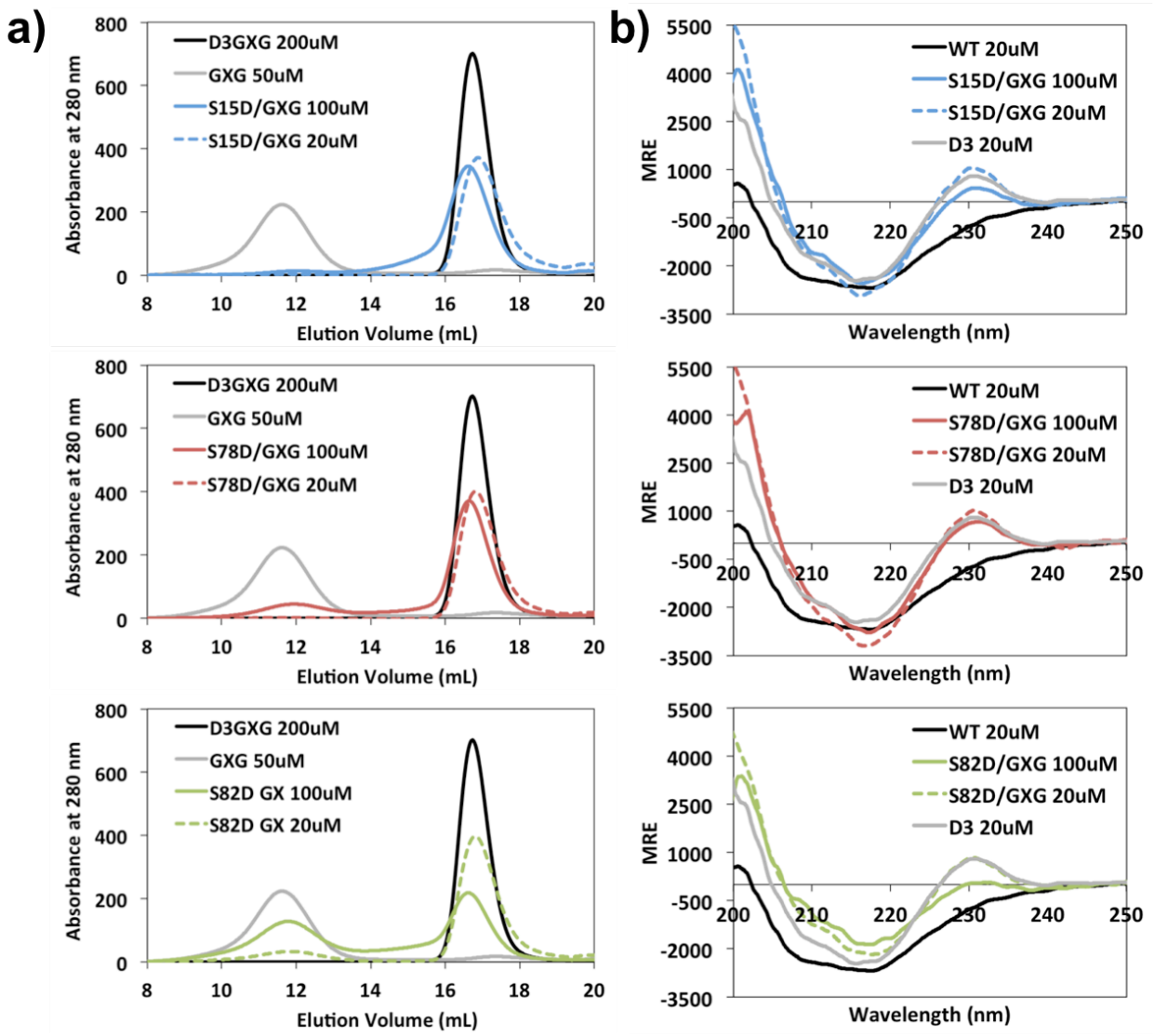


Figure 3.8- SEC and CD profiles of single phosphorylation mimics with the GXG mutation in HSPB1. (a) S15D/GXG, S78D/GXG, and S82D/GXG mutants were incubated at 100μM and 20μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. GXG and D3GXG traces (identical to Figures 3.6a and 3.7a) are shown for comparison. Each combinatorial mutant elutes in two regions- early, similar to GXG oligomers and later, similar to D3GXG oligomers. The proportions between very large and small oligomers shifts toward smaller species at low concentration and for particular mutants (S15D has the strongest effect). (b) 100μM and 20μM samples of single phosphorylation mimics/GXG in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. At 20μM, where D3-sized oligomers are present, combinatorial mutants adopt D3-like secondary structure. Depending on the proportion of small oligomers seen by SEC, each mutant at 100μM also shifts toward a D3-like secondary structure profile.

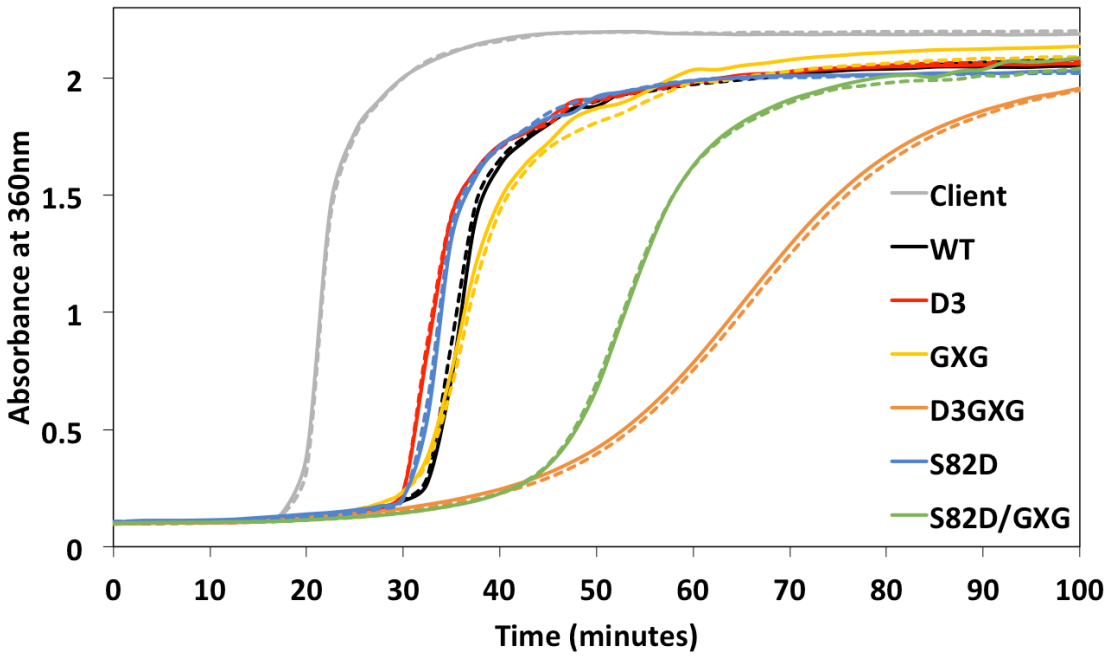


Figure 3.9- Chaperone assay of WT, D3, GXG, and D3/GXG forms of HSPB1. The chaperone assay was performed in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C. Aggregation of the model client protein α-lactalbumin (600μM) was triggered by reduction of disulfide bonds (addition of 20mM DTT) and subsequent destabilization. Absorbance was then measured as a function of time in the presence or absence of WT, D3, GXG, S82D, S82DGXG, or D3GXG forms of HSPB1 at 30μM (monomer concentration). Solid and dotted lines represent duplicates within one assay. Absorbance increases as the client protein aggregates due to light scattering. WT, D3, GXG, and S82D HSPB1 mildly delay formation of aggregates. D3GXG delays aggregation substantially, while S82D/GXG has an intermediate effect.

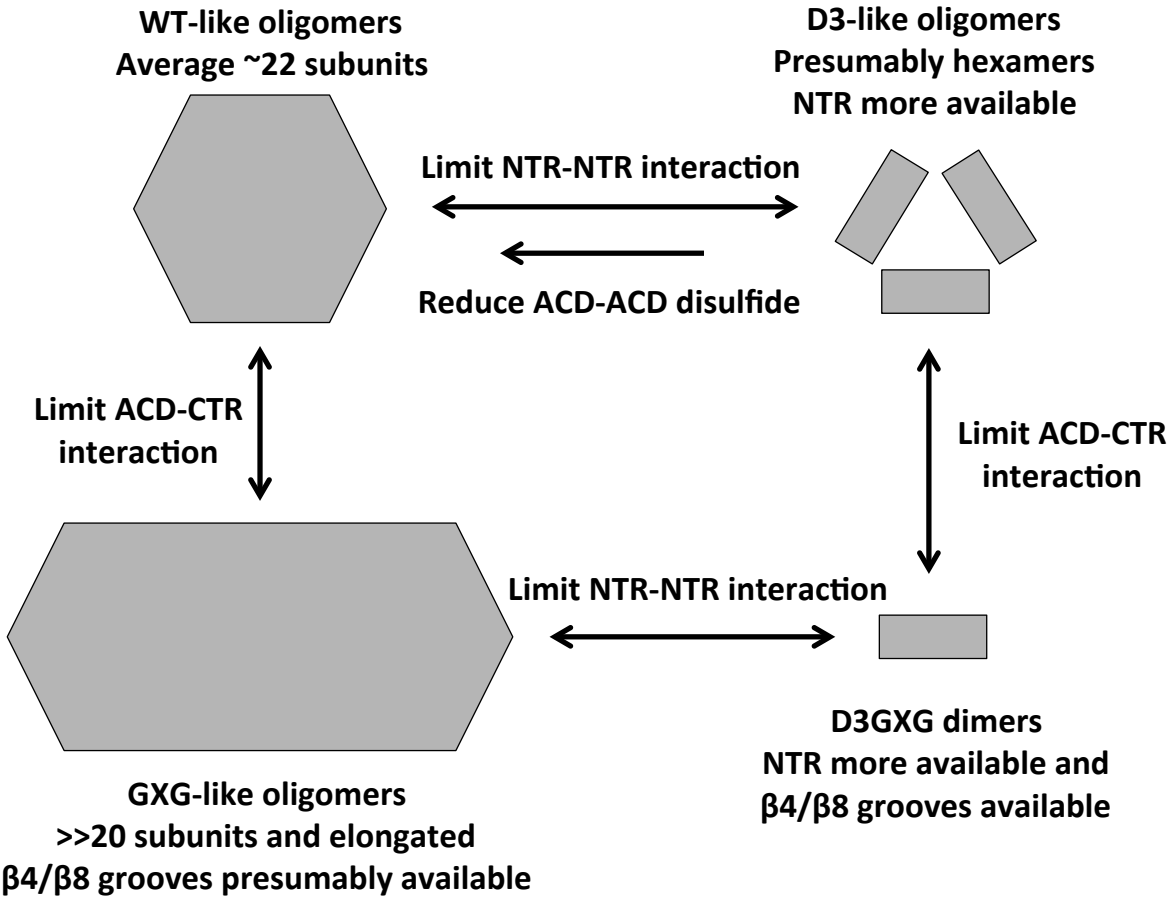


Figure 3.10- Summary of effects of perturbing inter-protomer interactions in HSPB1. Limiting NTR-NTR and ACD-CTR interactions have disparate effects on oligomeric propensity, leading to smaller or larger oligomers, respectively. In each case, different potential client binding sites are presumably more exposed. Combining mutations that perturb both types of interaction leads to a dimeric protein (D3GXG) that is highly chaperone active.

4. STRUCTURAL CHARACTERIZATION AND COMPARISON OF WT OLIGOMERS AND D3GXG DIMER

Although there are now a few atomic-level structures of the HSPB1 truncation ACD-only, little is known about the structures adopted in the NTR. One might be inclined to make inferences from structural models of HSPB5, but there is little homology in this region. Furthermore, for no polydisperse sHSP has there been an analysis of changes in local structure amongst the many oligomeric forms. The D3GXG mutant described in the previous chapter provides a tractable state, a discrete dimer of HSPB1, to characterize. As dimers are thought to be a basic building block for oligomerization, I chose to compare the dimeric (albeit "phosphorylated") state of HSPB1 with larger WT oligomers.

Far UV CD spectra of D3GXG and other phosphorylation mimics (Figures 3.2b, 3.3b, and 3.7c) demonstrate that there are local structural changes in smaller oligomeric states of HSPB1. Higher resolution techniques are required to characterize and locate structural features to particular regions or residues. As described in Chapter 1.3, polydisperse sHSPs are not particularly amenable to most traditional structural approaches. Therefore, hydrogen deuterium exchange mass spectrometry (HDXMS) was used to identify solvent-protected regions of dimeric and WT-oligomeric states at peptide-level resolution. Nuclear magnetic resonance (NMR) spectroscopy was used to characterize the D3GXG dimer at the residue-level. Many challenges arose in the NMR analysis due to the multiple conformations present in even the HSPB1 dimer. NMR of the large oligomers of HSPB1 yielded limited information and was not pursued in great detail. Nevertheless, comparison of peak heterogeneity and positions between WT oligomers and the D3GXG dimer allowed for numerous inferences to be made.

4.1 Comparison of the solvent accessibility of regions in WT oligomers and D3GXG dimers

4.1.1 Technical considerations

HDXMS was used to determine the solvent accessibility of the entire sequence of HSPB1 in the context of WT oligomers and D3GXG dimers. C137S forms of each protein were used to avoid possible complications from reducing agent required to avoid potential disulfide formation across the dimer interface. The deuteration exchange reaction was performed under standard buffer conditions at pH 7.5. After the exchange step, the reaction is quenched, allowing for digestion into peptides for MS analysis. In general it is backbone amide protons that are probed, specifically the rate of hydrogen bond formation among backbone amides in secondary structure. Side-chain amides and hydroxyl groups exchange too quickly for observation, as they are not often involved in more protected hydrogen bonding. Aliphatic protons will not exchange, as the carbon-hydrogen bond is covalent. One major caveat to this method that particularly affects experiments on sHSPs is the fact that the protein sample is diluted into deuterated buffer (ten-fold in this case). For monomeric proteins, the protein concentration is not expected to impact local dynamics. However, sHSPs may dissociate into smaller species upon dilution, depending on the concentration range involved and the overall stability of the oligomer. At time points after dilution into deuterated buffer, there is the chance that deuteration profiles could be affected by not just local dynamics in an equilibrated state, but also by slow oligomer dissociation with time. In this experimental design, WT and D3GXG proteins were equilibrated at 200 μM at room temperature (~25°C) and diluted ten-fold in deuterated buffer to a final concentration of 20 μM for either species. Based on SEC results dissociation is not a concern in this concentration range. The GXG mutations alter pepsin cleavage in the CTR, therefore identical peptides could not be compared between WT and D3GXG for the proximal region of the CTR. Similar peptides

between the two proteins can still be compared qualitatively. Representative peptides (denoted by residue number, e.g. "1-10") are shown in Figure 4.1; for a complete listing of observed peptides and their deuteration profiles see Appendix A.

4.1.2 ACD and CTR

In WT oligomeric HSPB1, over the time course of the experiment (3 seconds to 1 hour), the ACD is the most highly protected region (Figure 4.1, slow deuterium uptake). Most peptides derived from the ACD show only 50% deuterium exchange after one hour at room temperature (~25°C). This result is consistent with the stable (hydrogen bonded) β-sandwich characteristic of this domain. The CTR is not protected from solvent exchange, as expected for a highly exposed region. The peptide that includes the IXI motif (179-185) is predominantly unprotected, with a small, protected population observed at the earliest time point. These results are consistent with findings that only a small population of CTRs in oligomers are bound to β4/β8 grooves.

The deuteration profile of the dimeric D3GXG mutant is similar for the ACD and CTR relative to WT oligomers. ACD peptides are slightly less protected than in WT. Increased accessibility of the ACD in a dimeric context is suggestive that the same regions in the oligomeric context are likely more protected due to the overall avidity from other inter-protomer interactions. Specifically, there are NTR-ACD interactions (discussed in section 4.4) that could vary between dimeric and oligomeric states. Although identical peptides for the start of the CTR cannot be compared, the deuteration profile of the 164-185 peptide in D3GXG is consistent with the total deuteration observed for peptides 164-178 and 179-185 in WT oligomers. Aside from detectable IXI binding in WT oligomers, the CTRs for both proteins are likely similarly disordered.

4.1.3 NTR

Even in the context of large HSPB1 oligomers, the NTR is generally unprotected. This behavior is expected for a disordered and/or highly exposed region of a protein. However, multiple peptides show both protected and unprotected conformations, confirmed by several overlapping peptides (Appendix A).

The most distal peptide (1-10) exists in two distinct states, one moderately protected and one unprotected. As the isotopic distributions of the two states overlap substantially in the mass spectra, it is challenging to quantitatively deconvolute multiple states. Specifically, the exact proportion and protection of each state should not be heavily relied upon. The observation that the protected population of the 1-10 peptide in WT protein does not substantially decrease in proportion in the early three time points suggests that the exchange between the two states of this region is relatively slow. Similar heterogeneity is observed in D3GXG with similar deuteration profiles of the two states. However, the proportions between the two states changes more rapidly for D3GXG, possibly suggesting a faster exchange rate between states. As this heterogeneity is observed in both the oligomeric and dimeric proteins, the deuteration behavior observed for the protected state of this peptide is more consistent with a local structural feature, such as formation of a helix or intramolecular β-strands or binding in a groove, as opposed to protection in a buried core in an oligomer. This structural feature could be strictly local or involve interaction with another part of the protein. The observed change in conformational exchange rate between the WT and D3GXG proteins could be dependent upon restraint in the oligomeric context. Specifically, in the WT context a subunit might need to leave an oligomer in order for exchange to occur in this region of the NTR.

Peptide 11-28 shows moderate protection at early time points for WT but less protection for D3GXG. As this peptide contains the S15 phosphorylation site (mutated to Asp in D3GXG), it is likely that this region exists in different conformations or has different accessibility between oligomeric and dimeric states.

Peptide 29-41 shows two distinct states of protection for WT, similar to peptide 1-10. This peptide contains the single highly conserved region of the NTR (among sHSPs), has moderate helical propensity based on secondary structure prediction, and includes several sites of disease-associated mutation. In D3GXG no heterogeneity is observed and the peptide is much less protected. However, the 33-41 peptide, which covers the latter part of the same region, is more protected in D3GXG than in WT. This suggests different structural features are formed in this region between the oligomeric and dimeric states. The protected state in D3GXG is towards the latter part of this region.

The 55-62 peptide is unprotected for both WT and D3GXG. The 63-76 peptide in WT and D3GXG is unprotected. Therefore, much of the NTR (roughly position 54 through 76) is highly solvent exposed, not involved in hydrogen bonding, and not involved in tightly packed hydrophobic cores in either oligomeric or dimeric proteins. However, transient hydrophobic interactions of side-chains are possible and might not be reported on by this method.

Finally, peptide 77-99 shows two distinct states in WT oligomers. Several peptides in this region also show heterogeneity, but deconvolution of the two states is most reliable for the 77-99 peptide. This region includes the sequence that is predicted to form the β2 strand as well as the conserved β3 sequence. The two populations might be attributed to formation of the β2 strand (more protected state) in one population and a different orientation of this sequence from the ACD in the other (less protected state). The deuteration profile of this peptide in D3GXG

more closely mimics the less protected state in WT. The simplest explanation would be that the D3GXG state and the less protected state in WT are structurally similar, as opposed to a new conformation being adopted.

It is also important to note that the loss of heterogeneity in D3GXG peptides 29-41 and 77-99 might be coupled. Also, peptide 11-28 loses protection in D3GXG, which is the location of the S15D phosphorylation mimicking mutation. The 77-99 peptide contains the S78D and S82D mutation sites. It is tempting to speculate that these regions could be spatially close in the oligomeric context (intra- or inter-subunit) but are not associated in the D3GXG dimer.

The HDXMS analysis of HSPB1 WT oligomers shows that the NTR is highly solvent-exposed relative to the ACD. This suggests that the NTRs in these oligomers are not tightly packed in a hydrophobic core. Instead, oligomers must be held together by hydrophobic interactions and/or very limited inter-protomer hydrogen bonding, comparable to a molten globule state. Remarkably, comparison of WT oligomers and D3GXG dimers shows that the NTR in the oligomeric context is not substantially more protected than in the dimer context, where the NTRs are presumably completely exposed. Peptides 11-28 and 29-41 are moderately more protected in WT oligomers. These differences could be attributed to weak hydrophobic interactions or formation of secondary structure (e.g. a short-lived α-helix) in the oligomeric context. Overall, the HDXMS approach identified regions of structural heterogeneity in different quaternary states of HSPB1.

4.2 Stability of structural features in WT oligomers and D3GXG dimers

4.2.1 Thermal melt analysis of WT and D3GXG

To compare the stabilities of D3GXG dimers and WT oligomers, each protein was thermally melted and their far UV CD spectra measured. Figure 4.2 shows CD spectra for WT

under reducing conditions and D3GXG/C137S every 10°C between 15°C and 95°C. Reducing conditions are used (or mimicked by C137S) so that a covalent bond at the dimer interface does not impact unfolding. For WT protein, there is a large transition (increase in negative ellipticity) at shorter wavelengths (~205 nm) around 65°C, with the 65°C scan representing an intermediate along this transition. For D3GXG, this similarly large transition occurs between 55°C and 65°C with no intermediate scan, therefore a presumable melting temperature is ~60°C. It appears there is a slight decrease in stability for D3GXG, but a higher resolution melt (and on identical constructs) would be required to better define the difference. Overall, D3GXG mutations do not substantially destabilize HSPB1 nor do the transient intermolecular interactions in WT oligomers afford greater stability.

Loss of the structural feature responsible for the positive 230 nm peak observed in D3GXG spectra begins to occur at lower temperatures than for the global unfolding around 65°C. The 230 nm peak decreases in intensity across the 15°C to 55°C range. This earlier unfolding suggests that the responsible structural feature is less stable relative to the global stability of the protein. It is interesting to note that the D3GXG spectrum at 55°C is highly similar to that of WT (20°C spectrum is overlaid in dotted black line). It is likely that once transient local structures have been destabilized in WT and D3GXG proteins, their overall secondary structures are very similar.

4.2.2 Characterization of N-terminal peptide proposed to form β-hairpin

The positive 230 nm peak observed in CD spectra of D3GXG dimers and other small, phosphorylation mimicking species is most likely attributable to an exciton couplet between two aromatic tryptophan side chains in close proximity. There are six tryptophans in HSPB1, five of which are in the NTR. One tryptophan (W95) is located at the NTR/ACD boundary. In a

sequence-based search of the PDB, the NTR of HSPB1 shows a small stretch of homology with yeast isomaltase (Figure 4.3a and b, PDB ID: 3AXH). The region of HSPB1 that aligns with isomaltase extends from P39 to W51. The aligned region in isomaltase forms a β-hairpin in which tryptophan side chain indole rings are interacting. Tryptophan exciton couplets are common in β-hairpin structures stabilized by tryptophan interactions (Eidenschink et al. 2009). Therefore, the exciton couplet observed in D3GXG could be due to formation of a β-hairpin in the aligned region, with W42 and W51 interacting.

To test this hypothesis, a synthetic peptide containing this sequence (40-53) of HSPB1 was purchased for CD analysis. N-terminal acetylation and C-terminal amidation were included in the peptide synthesis to avoid non-native charges at the termini that might bias hairpin formation. Figure 4.3c shows the far UV CD spectrum of the peptide under identical buffer and temperature conditions (20°C) as the overlaid WT and D3GXG spectra. The peptide yields a positive peak around 225 nm, shifted from the peak observed in D3GXG. At shorter wavelengths (~205 nm), the peptide yields highly negative ellipticity indicative of some random coil structure. While the positive 225 nm peak may be attributed to an exciton couplet, consideration of the full spectrum suggests that this positive peak may also be attributed to β-turn or extended helical structure. A thermal melt of the peptide was performed to determine if unfolding of any structure occurs. Figure 4.3d shows CD spectra of the peptide every 10°C ranging from 5°C to 95°C. Although there is no dramatic, cooperative transition, there are changes in spectral features with increasing temperature. Specifically, ellipticity around 200 nm decreases (approaches zero), and the 225 nm peak decreases steadily with temperature. These transitions with increasing temperature indicate that some sort of secondary structure is present

in the peptide at low temperatures. The nature of the structure of this HSPB1 NTR sequence in a peptide context remains to be more fully investigated.

4.3 Structural analysis of D3GXG dimer by NMR

4.3.1 Technical considerations

The D3GXG dimer is approximately 45 kDa, which is at the upper size limit for routine NMR structural analysis. There are many factors to optimize when studying a new protein by NMR. Thorough analysis of the ACD-only HSPB1 dimer (~22 kDa) by others in the Klevit lab (Rajagopal et al. 2015a) provided a starting point for examination of the full-length dimer, and some of the parameters that were optimized are discussed here.

(1) Variation of temperature and buffer- As temperature can modulate the intensity and number of resonances visible in the NMR spectrum, a range of temperatures was tested for D3GXG. In general, lower temperatures (below 25°C) did not offer increased signal or any other advantages. Higher temperatures (room temperature to 37°C) yielded spectra with stronger signals, but at 37°C there was considerably more breakdown of the protein, which is detrimental to long 3D experiments (several days) necessary for assignment. Therefore 30°C was selected as the optimal temperature for a compromise of signal and minimizing sample breakdown. The buffer used for studying the ACD-only construct was 50 mM phosphate buffer with 100 mM NaCl at pH 7.5, therefore it was preferable to use similar conditions to be able to compare the full-length protein. Lowering the ionic strength, which can affect signal when using a cryoprobe, did not alter the spectrum. As exchange of amide protons is slowed as the pH is lowered from neutral to ~2.5, yielding sharper resonances, a pH buffer lower than neutral is preferable for NMR studies of many proteins. However, HSPB1 (and HSPB5) is highly sensitive to even slight lowering of pH, therefore this path was not pursued. As with the structural analysis of the ACD-

only protein, non-reducing conditions were used for studies of D3GXG as locking the dimer appears to limit exchange and conformational heterogeneity, yielding sharper peaks.

(2) Deuteration of the protein- As the size of a protein increases, the NMR signal broadens substantially due to relaxation effects during the NMR experiments. The effects are from two types of relaxation, T_1 (also called longitudinal or spin-lattice) and T_2 (also called transverse or spin-spin). T_1 relaxation is related to the time necessary for spins to relax back to the z-axis. Larger proteins generally have shorter T_1 relaxation times. T_2 relaxation can be thought of as a loss of coherence of the spins while they are in the x-y plane, due to interactions with neighboring nuclei. As interactions with aliphatic protons are a substantial source for loss of coherence, deuterating a protein (replacing these protons with deuterons) removes some relaxation pathways, leading to sharper signals. For D3GXG, partially deuterating the protein (~75% by growing in D_2O) leads to dramatically sharper signals and was necessary for 3D experiments. Perdeuteration (~99% by growing in D_2O and deuterated glucose) sharpened signals somewhat further, but as it is cost prohibitive perdeuteration was used only for the less sensitive HNCB experiment.

(3) Truncation of the protein- As decreasing protein size aids in both relaxation effects and spectral overlap, a CTR-truncated form was used for most NMR experiments ("D3/176del"). By truncating the D3 protein after position 176, identical to the CTR truncation in the ACD-only construct, the dimer size was reduced to ~40 kDa. This truncation yielded sharper NMR signals and reduced spectral overlap of resonances corresponding to unstructured residues in the NTR and CTR (Figure 4.4, center of the spectrum). No other effects were observed in the rest of the protein with this truncation. The CD spectrum of this mutant (data not shown) contains the same positive 230 nm peak as for D3GXG, indicating that the distinct structural feature observed in

HSPB1 small oligomers and dimers is present in this construct. The most notable difference with D3/176del compared to D3GXG was temperature-dependent precipitation. While D3GXG is soluble at high concentration at low temperature, D3/176del precipitates at 4°C, forming a cloudy gel. The precipitation is reversible and disappears after warming the protein solution to room temperature. The charge-enriched CTR likely functions as a solubility tag, counteracting hydrophobic intermolecular interactions of the rest of the protein.

4.3.2 Assignment of the NMR spectrum of the full-length dimer

The NMR spectrum of the D3GXG mutant of HSPB1 was assigned through use of standard triple resonance experiments. Assigned chemical shifts are listed in Appendix B. A combination of TROSY-based HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HNCB, and NNH experiments were performed on the D3/176del mutant, which yields sharper peaks and less spectral overlap than the full-length protein. HNCO, HNCA, and HNCACB experiments were performed on the full-length D3GXG in order to assign the distal CTR residues. A total of 217 peaks from ^{1}H-^{15}N TROSY HSQC spectra were seen reproducibly (not noise) for D3/176del (Table 4.1). These arise from either backbone amides or tryptophan side chain indole H-N groups (Figure 4.4). Other side chain H-N groups (belonging to Gln and Asn) that appear in the upper right region of this type of spectrum were not included in this count. Out of the 176 residues in this construct, 19 are prolines and will not yield resonances in a ^{1}H-^{15}N TROSY spectrum or the HN-based 3D experiments listed above. Therefore, there are 60 resonances more than expected, indicating alternate conformations of the protein. To observe two resonances for one residue, the conformational exchange must be slow relative to the NMR timescale (milliseconds or longer).

General characteristics of the observed peaks in the ^{1}H-^{15}N TROSY HSQC spectrum provide insights into characteristics of the protein. A protein with ordered structure (little disorder) that predominantly adopts one conformation will yield a spectrum of well-dispersed peaks (typically ranging from 6 to 10 ppm in the ^{1}H dimension) of equal intensity with one peak per residue. A disordered protein will yield a spectrum of poorly dispersed peaks that are very intense. Flexibility of disordered regions reduces local relaxation effects and thus enhances signal. However, the lack of diverse chemical environments in a disordered region leads to overlapping peaks in the center of the spectrum (~7.5-8.5 ppm in the ^{1}H dimension). This region corresponds to random coil-like chemical shifts. The D3/176del protein has a mixture of strong, dispersed peaks, weaker, dispersed peaks, and very strong peaks in the central region. The strong, dispersed peaks mostly arise from (as discussed below) residues in the structured ACD. The strong peaks localized in the center of the spectrum belong to residues in unstructured regions of the protein, mostly of the NTR and CTR. The weak, dispersed peaks likely correspond to alternate and lowly populated states of structured regions, although very few of these peaks could be rigorously assigned to specific residues. Assignments of peaks to certain residue types are listed in Appendix B in addition to the specific residue assignments. Side chain indole H-N groups of tryptophans typically appear in the lower left corner of ^{1}H-^{15}N TROSY HSQC spectra. There are six tryptophans in HSPB1, five of which are in the NTR. However, only three strong peaks are observed in this region that likely do not correspond to backbone amides due to their absence in 3D spectra. There are additional weak peaks in this area that might correspond to tryptophan side chains.

4.3.3 Comparison with ACD-only construct

Overall, many spectral properties of the ACD are conserved between the ACD-only and D3/176del constructs. Their ^{1}H-^{15}N TROSY HSQC spectra overlay, but only to the trained eye as almost all peaks have shifted slightly (Figure 4.5a). Most assigned peaks of the ACD-only construct are observable under the conditions used to study the full-length dimer (30°C vs. 25°C used for ACD-only assignments). The majority of these ACD peaks were also assigned in D3/176del, although there are unassigned gaps in the sequence at distinct regions (Figure 4.5b). For these unassigned residues, there are no nearby candidate peaks in the spectra. Therefore, residues in these parts of the full-length protein are in a very different chemical environment (shifted far) relative to ACD-only or in intermediate exchange such that the corresponding peaks would be broadened substantially. Residues assigned in ACD-only but not D3/176del are indicated in yellow in Figure 4.5d. Resonances corresponding to most of β3, the loop between β3 and β4, and β8 are either unobserved or unassigned in D3/176del. There are also parts of the β5/β6+7 loop and β7 strand that do not have observable/assigned peaks in the D3/176del spectrum. The absence of (or inability to assign) certain peaks for D3/176del indicates that these regions are in distinctly different chemical environments or exchange regimes between the ACD-only and D3/176del forms.

Of residues assigned for both proteins, there are several that shift substantially between the two forms or that exhibit heterogeneity in the D3/176del context (Figure 4.5c). Residues that show more dramatic chemical shift perturbations (CSPs) are highlighted in orange in Figure 4.5d. The most affected regions are the β4 strand and parts of β9. Three residues in the ACD definitely show multiple peaks in the D3/176del spectrum and not in the ACD-only spectrum. These residues are highlighted in red in Figures 4.5b and d. T110 is located at the start of β4,

whose resonances are substantially shifted between the ACD-only and D3/176del spectra. K123 is at the start of the loop connecting β5 and β6+7 and is preceded by a small, unassigned region. V148 is located in the loop between β6+7 and β8, which is otherwise unperturbed. Altogether these comparisons with the ACD-only construct indicate that the presence of the NTR dramatically perturbs the β4 and β8 strands and has a lesser but noticeable effect on other regions of the ACD. Specifically, it is highly likely that the NTR is interacting with the ACD to cause these perturbations.

4.3.4 Insight into the NTR

A small fraction of the NTR in D3/176del was assigned from the triple resonance spectra (20 residues of ~85 belonging to the NTR). It is clear that parts of the NTR exist in multiple conformations based on the HDXMS results described earlier. Structural heterogeneity can affect the assignment process in two ways. (1) The signal for one residue is divided among many peaks representing the different chemical environments. (2) If the exchange among conformations is at a rate similar to the NMR timescale ("intermediate exchange"), peaks representing all conformations will be broadened out and difficult to detect. One region of the NTR was amenable to assignment- residues 62-79 (Figures 4.5b and 4.6a). Peaks corresponding to these residues were very intense and had random coil-like chemical shifts. The intensities of these peaks are comparable to those of the flexible CTR (Figure 4.6b). Therefore this region appears to be in a disordered and highly flexible conformation, a conclusion that is supported by a lack of solvent protection observed by HDXMS. However, for several of the assigned NTR residues there appear to be two distinct resonances (Figure 4.6a). The alternate conformation of residues 64-69 also appears to be disordered and represents a large fraction of the population (> 20% based on intensities). Therefore there are two observable, slowly interchanging states of this

disordered region of the NTR, and more structured flanking regions must modulate the slow exchange. A few other residues in the NTR were assigned in the 35-49 region, overlapping with the synthetic peptide examined in section 4.2.2. Two different peaks were assigned to S49, suggesting two environments for this region as well.

4.3.5 Assignment of the CTR

The most distal end of the CTR in the D3GXG construct was assigned (Figure 4.5b). The D3/176del truncation as well as a D3/198del truncation, which eliminated the second "AAK" repeat in the sequence, provided confirmation. Peaks corresponding to these residues are strong and located in the random coil region of the spectrum. The peak assignments are also confirmed by a recent report on the CTR of HSPB1 (Alderson et al. 2017). The inferred flexibility is consistent with what is known structurally of other sHSPs and with the HDXMS results. However, residues 177-191 were not assigned. Most peaks that are present in the D3GXG spectrum and not in the D3/176del spectrum are assigned to the more distal part of the CTR. Therefore, the 177-191 region is likely in intermediate exchange and relatively invisible by the ^{1}H-^{15}N- based experiments used here. As this region is highly unprotected in the HDXMS analysis, there must be intermediate exchange among several solvent-unprotected states.

4.3.6 Probing long-range interactions by PRE

As much of the NTR remains unassigned in NMR spectra, it was necessary to proceed to another NMR method capable of providing distance information in the absence of assignments. The paramagnetic relaxation enhancement (PRE) effect from a spin label broadens resonances corresponding to residues spatially near the label. The nitroxide spin label 2,2,5,5-tetramethyl-1-oxyl-3-methyl methanethiosulfonate (MTSL) was used here and attached to cysteines introduced in the NTR. The effects of MTSL can be observed up to 2.5 nm away.

To introduce a single spin label, the following mutants were made in the D3/176del/C137S background (cys-free)- T2C, D15C, S50C, S65C, S74C, T91C, or D107C. All mutants were evaluated for size (correct elution during SEC step of purification) and secondary structure (CD). Samples used in PRE experiments were not deuterated, thus many fewer peaks were analyzed than are assigned. Also, the C137S mutation, which mimics a reduced state of the protein, causes shifting and loss of certain peaks in the D3/176del spectrum consistent with changes observed in ACD-only spectra (oxidized vs. reduced). As both protomers of the dimer are ^{15}N-labeled and spin-labeled, broadening observed can be due to intra- and inter-protomer interactions. Overall the NMR intensities for the labeled samples were weak, so the ratios of intensities were noisy. Therefore, only regions where multiple residues showed peak intensity loss are discussed below. Also, as no single region was affected by a spin label at each of the four positions, it is reasonable to conclude that the spin label does not bind non-specifically to the protein.

Cysteine mutation sites were selected based on tolerable sites found in previous studies (McDonald et al. 2012) and not being in highly conserved regions. Purification and initial analysis of these mutants provided some insights into certain regions of the NTR. T2C, S65C, S74C, and T91C behaved just like their cys-free counterpart and are explored below. The D15C mutant (site of phosphorylation) behaved poorly during purification and eluted as a larger than dimer species by SEC. The D107C mutant was highly aggregation prone during the purification and gave a CD signal very different from that of D3/176del. The S50C mutant was well behaved in every respect and yielded an ideal control NMR spectrum. However, attachment of the spin label must have caused the protein to self-associate (no visible aggregates) as it became almost

invisible by NMR. The failures of these residues to tolerate mutation or spin label attachment suggest that they might be in interacting regions of the protein.

The spin label attached at position 2 gave strong PRE effects on two regions of the ACD: the loop between β4 and β5 and the loop between β6+7 and β8 (Figure 4.7). This indicates that in HSPB1 dispersed dimers the very distal end of the NTR is in close proximity to the ACD, specifically at the ends away from the ACD-ACD interface. Position 2 is far from the ACD structure in sequence space, therefore the interaction could be intra- or inter-protomer. It is also important to note that a spin label at this position did not detectably broaden unassigned peaks that might have corresponded to the region proximal to residue 2. This observation suggests that the corresponding peaks are not observed and are therefore in intermediate exchange, consistent with the heterogeneity observed for this peptide by HDXMS.

Spin labels attached at either position 65 or 74 caused similar PRE effects to each other. Peaks assigned to residues sequentially near 65 and 74 are substantially affected by the attached spin label, providing support that the assignments are correct. Also, as this region of the NTR is assigned, it is possible to judge the strength of interactions relative to the broadened signals in this region. The data obtained from these two spin label experiments are noisy due to the wide-ranging effects of spin labels at positions 65 and 74. Therefore, only regions where multiple peaks are affected are discussed here. Both spin label positions affect the loop between β5 and β6+7 and the earlier part of β9. Position 74 also appears close to the loop/small helix between β3 and β4. These two primary affected regions are spatially distant within a protomer or across protomers in a dimer, suggesting that the region of the NTR containing positions 65 and 74 could be in two conformations, with each closer to one of these regions of the ACD. This is consistent with the multiple assignments for residues in this region of the NTR. The intensity losses

associated with each position were substantial, but not as large as in the region of the NTR containing the spin label. The lessened effects on the ACD could be due to either a moderate distance between the spin label and the affected region or more likely that the region containing the spin label adopts two conformations that interact with different regions of the ACD.

Position 91 is at the boundary between NTR and ACD and would likely be part of the β2 strand if it is forming. The spin label at position 91 caused substantial broadening in β3, β9, the β6+7/β8 loop, the β5/β6+7, and the CTR. The β5/β6+7 loop interaction is likely inter-protomer as position 91 could not reach this loop in its protomer without release of the β3 strand. For the same reason, the remaining interactions are likely intra-protomer. The orientation of this proximal part of the NTR (containing the possible β2 strand) relative to the ACD is not clear from these data alone. However, the observation that assignments in this region do not correlate between ACD-only and D3/176del constructs and the PRE effects on the β5/β6+7 loop suggest that the 91 region might fold in on the ACD, different from the ACD-only NMR structure.

4.4 NMR of HSPB1 oligomers

WT HSPB1 oligomers are large (average ~450 kDa) and are not amenable to conventional NMR experiments, as the amide resonances typically probed are not visible. However, highly mobile methyl groups of certain side chains can provide detectable signal in an otherwise perdeuterated, large protein. As there are only six isoleucines in HSPB1, we chose to start with Ile-δ-$^{1}H^{13}C$ labeling (only the delta methyl group in the side chain), where the rest of the protein is $^{2}H^{12}C$. Two isoleucines are in the NTR or at the NTR-ACD boundary (I63 and I88), two are in the ACD (I120 and I134), and two are in the IXI motif of the CTR (I179 and I181). ^{1}H-^{13}C HMQC spectra in principle can reveal information about each isoleucine. There are more peaks than expected in the spectrum of WT HSPB1 oligomers, immediately implying

conformational heterogeneity (Figure 4.8). There are strong overlapping peaks in the random-coil center of the spectrum (~12 ppm in ^{13}C dimension), which likely correspond to isoleucine methyl groups in disordered regions. Four peaks of lower intensity are downfield (in ^{13}C, less shielded) of the random-coil peaks, and likely correspond to methyl groups in the ACD, based on HSPB1 ACD-only assignments and observation of ACD peaks in HSPB5 methyl-TROSY spectra (Baldwin et al. 2012). An additional low intensity peak appears upfield of the random-coil peaks and could correspond to a bound state of the IXI motif. In HSPB5 methyl-TROSY spectra, peaks upfield (^{13}C) were assigned to the bound state of the IXI motif.

Rigorous assignment of the oligomer Ile spectrum can only be accomplished by mutation of each Ile individually. Given the cost of preparing one sample, this approach is cost prohibitive. However, inferences can be drawn from comparison with other forms of HSPB1-D3GXG, D3 and ACD-only. For D3GXG spectra, the amino acid precursor to label the methyl groups was used so that the remainder of the side chain would be deuterated, but the protein was not otherwise perdeuterated. An HMQC spectrum of D3GXG, which has I179G and I181G mutations, shows only four distinct peaks, corresponding to the four indicated isoleucines and confirmed by mutational analysis. Mutation of each Ile to Val resulted in disappearance of one peak from the spectrum. In the D3GXG context I63 is in the random coil region as predicted, and the remaining isoleucines correlate well with their ACD-only counterparts. I120 and I134 are downfield (in ^{13}C) of the random coil region, consistent with being in the structured ACD. There are multiple peaks in the WT oligomer spectrum in this region, suggesting one or both of these isoleucines might be in multiple conformations. In the WT oligomer spectrum, there are multiple strong, broad, overlapping peaks in the central random coil region, which could be attributed to I63, I88, I179, and I181. I63 and I88 are likely flexible based on their chemical

shifts in the D3GXG spectrum and HDXMS analysis. I179 and I181 are part of the IXI motif that is expected to be bound in small populations in an oligomeric state, thus a large flexible population could exist. There is one more weak peak upfield (in ^{13}C) of the random coil region that is likely attributable to the bound state of the IXI motif. An HMQC spectrum of the D3 mutant, labeled identically to WT, could be construed as a combination of WT and D3GXG spectra. There are the expected intense, overlapping peaks in the random coil region. The I120 peak of D3GXG overlaps well with a peak in D3, although for D3 the peak is partially spread out toward the nearby WT peak that is presumably I120. Similarly, a peak for D3 roughly overlays with I134 in D3GXG, but low intensity peaks are also observed at WT peak positions that might correspond to alternate states of I134 in the oligomer. The most striking difference however is the pair of strong peaks just upfield of the random coil cluster that do not overlay with the weak WT peak in this region. If these peaks correspond to the bound IXI motif, this result suggests that there is more binding of the IXI motif in the D3 context compared to WT and that the exact orientations are different, due to changes in chemical shifts. As D3GXG represents an extreme dimer form of phosphorylated HSPB1 and D3 at high concentrations is a broad mixture of different oligomeric species (SEC), it is consistent that resonances corresponding to WT-like and D3GXG-like states would be observed for D3 oligomers.

Preliminary attempts were made to probe leucines and valines in HSPB1 oligomers. The amino acid precursor labels both types of amino acid at one of their δ (leucine) or γ (valine) methyl groups. There are numerous Leu/Val residues in HSPB1, particularly in the primarily unstructured NTR. The resulting HMQC spectrum therefore yields overwhelmingly intense, broad peaks in the random coil regions of the spectrum that drown out most distinct resonances if they are even observable (4.9). There are very weak peaks observed upfield (^{1}H) that could be

attributed to more structured states of these side chains, but the cost/benefit ratio did not warrant further experiments.

4.5 Discussion of structural analysis of WT oligomers and D3GXG dimers

4.5.1 Structural changes in the NTR between WT oligomers and D3GXG dimers

The combination of phosphorylation-mimicking mutations and the GXG mutations yield a monodisperse dimer (D3GXG) representative of an extreme phosphorylated state of HSPB1. Circular dichroism (CD) spectra confirm that D3 and D3GXG adopt very similar secondary structure. CD spectra of intermediate phosphorylation mimics adopt D3-like secondary structure at lower concentrations where smaller oligomers are favored. Therefore, local structural changes observed by CD are dependent only on oligomeric size of phosphorylation mimicking mutants and not on particular mutations. Only in a sufficiently small oligomeric state, presumably where the NTR maximally unrestrained, is there an observable change in secondary structure.

Using HDXMS, it was possible to compare regions of solvent protection between the WT oligomeric state and the D3GXG dimeric state and make inferences about differences in local structural features. In general, peptides derived from the ACD and CTR had similar protection profiles between the two states of HSPB1. Certain regions the NTR, however, differed in their solvent protection between WT and D3GXG. Peptide 11-28, which contains the S15 phosphorylation site, is moderately protected in WT and unprotected in D3GXG. This suggests that some type of structure in this region is favored in the oligomeric state or this region is involved in oligomeric contacts. Peptide 29-41 shows conformational heterogeneity in this region for WT oligomers, and contains the only moderately predicted helical region. The ordered conformation in WT oligomers is lacking in D3GXG (unprotected), suggesting a potential structural feature or interaction is absent. However, peptide 33-41 is more protected in

D3GXG, suggesting an alternate structural feature is formed in the dimer compared to WT oligomers. CD analysis of a peptide containing the 40-53 sequence of the NTR demonstrates that there is some non-standard secondary structure formation in this region, but it does not have helical propensity. There is also an observed loss of conformational heterogeneity in the 77-99 peptide for D3GXG. This coupled change suggests that the 11-41 region of the NTR might interact with the 77-99 region at the NTR/ACD interface in the WT oligomeric state. This particular interaction could be either intra- or inter-protomer in an oligomer. Alternatively, these regions might not be spatially close but experience an indirectly coupled structural effect.

4.5.2 Model of the D3GXG dimer

Through the use of NMR assignments, PRE measurements, and HDXMS, it is feasible to propose models for the general structure of a full-length HSPB1 dimer, D3GXG. Figure 4.10 visually and sequentially represents (parts a and b) all of these results for D3GXG (or D3/176del for NMR experiments).

It was possible to assign the 62-79 region of the NTR by NMR. The chemical shifts and strong intensities of peaks corresponding to residues in this region are random-coil-like, indicating a region of disorder. This is consistent with the lack of protection of this region seen by HDXMS. Interestingly, two distinct resonances (both in the random-coil spectral region) were observed for several residues in this region. Therefore even in a highly disordered part of the NTR, multiple conformations are observed. The slow exchange (as both states are observable by NMR) could be due to different structural constraints upstream or downstream of this region. The abundance of prolines in this region (and throughout the NTR) would likely promote heterogeneity via *cis/trans* proline isomerization. In unconstrained regions of a protein, the peptide bond preceding a proline is typically in a *cis* state 10% of the time. The *cis*

propensity is increased substantially (up to 40%) when a proline is preceded by a glycine or aromatic, both of which are common in the NTR. Via PRE experiments, positions 65 and 74 were found to interact with the β3/β4 loop, β5/β6+7 loop, and the β9 strand. As these regions are spatially distant both within a protomer and between protomers, it is likely that this region of the NTR (at least residues 65-74) has multiple orientations with respect to the ACD. It is unknown whether these contacts are intra- or inter-protomer. The assigned region encompasses the "HSPB1 insertion" that is unique to HSPB1 and shares no sequence homology with the other human sHSPs. This sequence was inserted between the NTR and ACD of a monodisperse sHSP, the archeal Hsp16.5. Normally this sHSP forms oligomers of 24 protomers and can be crystallized, although the NTR is unresolved. Insertion of the HSPB1 sequence resulted in formation of monodisperse 48-mers that also crystallized, but the NTR, including the HSPB1 insertion, remained unresolved (Mchouarab et al. 2012). It is apparent that this region of the HSPB1 NTR, although disordered, is important for modulation of oligomer size in certain contexts.

The fact that the remainder of the NTR could not be assigned with these datasets under these conditions suggests that other parts of the NTR are either: (1) exchanging among states at an intermediate time scale (relative to the NMR experiment) such that peaks are broadened or (2) signals are split among too many exchanging conformations. A comparison of residue assignments in the ACD between ACD-only and D3/176del constructs indicates regions of interest where the NTR might be interacting. Residues in the β4 strand have substantial CSPs between these two constructs, indicating a difference in chemical environment. Residues in the β8 strand were assigned in the ACD-only context under these conditions but no similar peaks could be assigned for the D3/176del construct. Therefore, the β8 strand is also undergoing

different behavior between these two forms of the protein. By HDXMS, peptide 1-10 shows two conformations at the distal end of the NTR- one unprotected and one moderately protected. The moderately protected state decreases in proportion within a few minutes, suggesting that it interchanges with the unprotected state rapidly. The PRE result that shows position 2 in close proximity to the edge of the ACD coupled with this protection suggests that the distal end of the NTR is interacting with the ACD. Either the 1-10 region or the 33-41 region of the NTR (which shows protection by HDXMS) could be interacting with the β4/β8 region, causing the observed CSPs relative to the ACD-only construct. As with discussion of the 62-79 region of the NTR, NTR contacts with the β4/β8 region of the ACD could be inter- or intra-protomer.

Although there is no direct evidence for the formation of the β2 strand antiparallel to β3 in the D3GXG dimeric context, the results suggest that the β2 region is not extended away from the ACD. This region (peptide 77-99) is not protected from solvent as seen by HDXMS. However, the inability to assign some of the β3 strand and following loop indicate there are differences in this groove between β3 strands (across the dimer) between ACD-only and D3GXG contexts. A recent crystal structure of full-length dimeric HSPB6 in complex with another protein does not show formation of this β2 strand (Sluchanko et al. 2017). Although much of the NTR is not resolved in this structure, the conserved region of the NTR is found to be inserted in the groove between β3 strands in the ACD dimer and on top of the dimer interface, with this region not adopting any ordered secondary structure (Figure 4.11). This HSPB6 N-terminal structure is in stark contrast to that of HSPB5 oligomeric models, but given the lack of sequence homology of the NTR among families, it is possible that each sHSP has very different structural features in the NTR. Given that there are no substantial CSPs deep in this β3-β3 groove (β6+7 strands) between ACD-only and D3/176del, it is unlikely that any part of the NTR

is truly bound *in* this groove. However, some part of the NTR or a weak β2 strand is likely in this region interacting weakly. It is also not clear whether NTR regions from both protomers in a dimer could occupy this β3-β3 groove. Two NTR regions inserted into the β3/β3 groove were not observed in the HSPB6 structure, suggesting the binding of one NTR region precludes binding of the other. Similarly for HSPB5 oligomeric models, β2 strands formed along β3 were not observed on both protomers of a dimer. Therefore, the two NTRs of one D3GXG dimer likely adopt different conformations, some of which are proposed in Figure 4.10.

4.5.3 Implications of NTR-ACD contacts for disease-associated mutants in the ACD

There are numerous disease-associated mutations reported in the ACDs of several sHSPs (Datskevich et al. 2012). These mutants are discussed in the previous chapter with respect to their ability to perturb the ACD-ACD dimer interface. Given the results presented here and HSPB5 oligomeric models, mutations in particular regions of the ACD would likely also affect interactions between the NTR and ACD. Specifically, mutations of residues whose side-chains make up the β3-β3 groove and residues along the β4, β8, and β9 strands, where NTR-ACD contacts are observed or predicted, should be considered as potentially perturbing multiple types of interactions in sHSPs.

Several reported disease-associated mutations in HSPB1 fit the criteria for potentially affecting NTR-ACD interactions. Mutation L99M is located in the β3 strand, which could directly affect formation of the β2 strand antiparallel to this strand. This mutant is one of few reported as autosomal recessive, which could be due to the relatively mild change from Leu to Met. Mutations R136W/L and R140G are along the dimer interface with side-chains facing the groove. These mutants alter both the local charge state by removing positively charged arginines and also the hydrophobicity/packing of the groove by introducing large hydrophobic side-chains

or removing a side-chain entirely. Mutations R127W (β5/β6+7 loop) and T164A (β9 strand) are in regions that were shown via spin probes to be close to positions 65 and 74 of the NTR. This general region of the NTR is hydrophobic, except for a Glu at position 64, and occupies two observed disordered states. Therefore, loss of a positive charge and introduction of a very hydrophobic side-chain (R127W) could alter presumed transient interactions with the 65-74 region of the NTR. Mutation T151I is located in the β6+7/β8 loop, where contacts were observed via spin probes at positions T2 and T91. A mutation in this region could affect both β2 strand formation/orientation and also the putative interaction with the most distal part of the NTR.

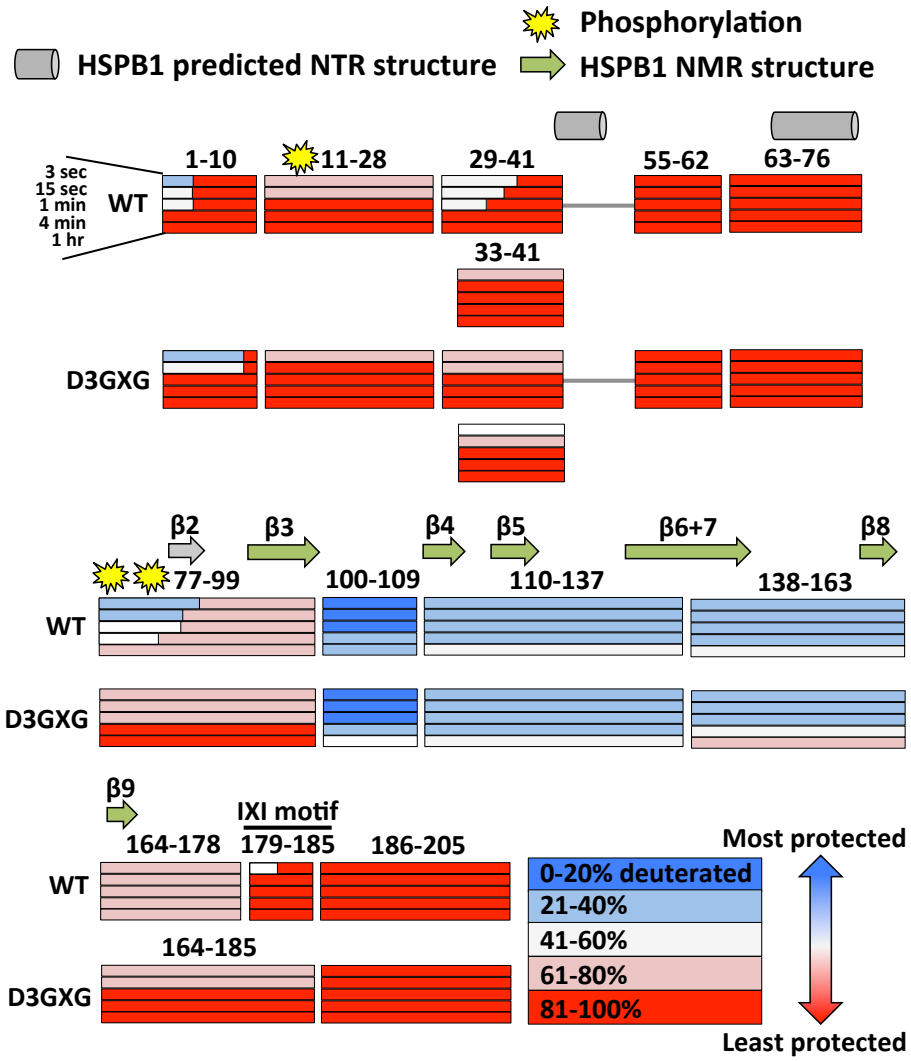


Figure 4.1- Comparison of solvent protected regions in HSPB1 WT oligomers and D3GXG dimers using HDXMS. WT/C137S and D3GXG/C137S samples were prepared at 200μM in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and equilibrated at room temperature for 3 hours. Exchange experiments were performed at room temperature (~25°C) by diluting the proteins 10X into an identical but D_2O based buffer. Peptides analyzed are shown as a series of horizontal bars (e.g. peptides consisting of 1-10, 11-28, etc.). Time points are represented as bars stacked vertically for one peptide. The color-coding represents deuteration level of each peptide (blue- most protected, red- least protected). For certain peptides, the horizontal bar is split into two sub-bars, which represent two distinct populations (conformations) observed in the mass spectrum for that peptide. The relative lengths of the sub-bars within one peptide bar correspond to relative proportions of the two conformations. Overall the ACD is highly protected for both WT oligomers and D3GXG dimers. The NTR and CTR are largely unprotected for both proteins, but distinct regions of some solvent protection are observed in the NTR. Within the NTR, there are moderate changes in solvent protection between WT and D3GXG for certain peptides.

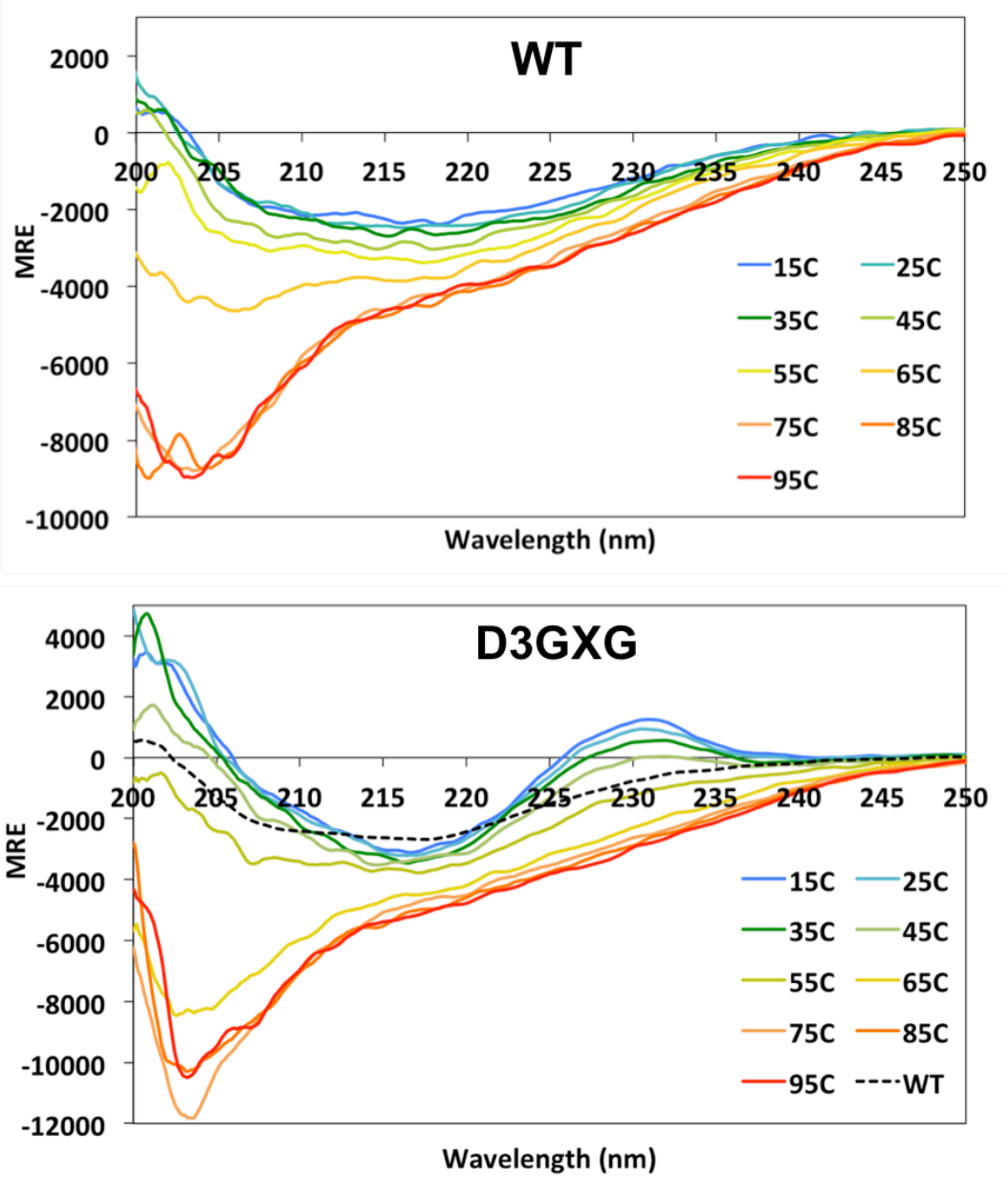


Figure 4.2- Far UV CD thermal melts of HSPB1 WT and D3GXG proteins. 20μM samples of WT (+ reducing agent) and D3GXG/C137S in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at temperatures ranging from 15°C to 95°C. Temperature was increased at a rate of 1°C/min and halted every 10°C for wavelength scans. WT protein shows a dramatic transition at shorter wavelengths around 65°C. D3GXG protein shows a similar transition at shorter wavelengths between 55°C and 65°C. The 230 nm positive peak is gone by 55°C, at which temperature the D3GXG CD spectrum looks highly WT-like (20°C WT spectrum shown as dotted black line).

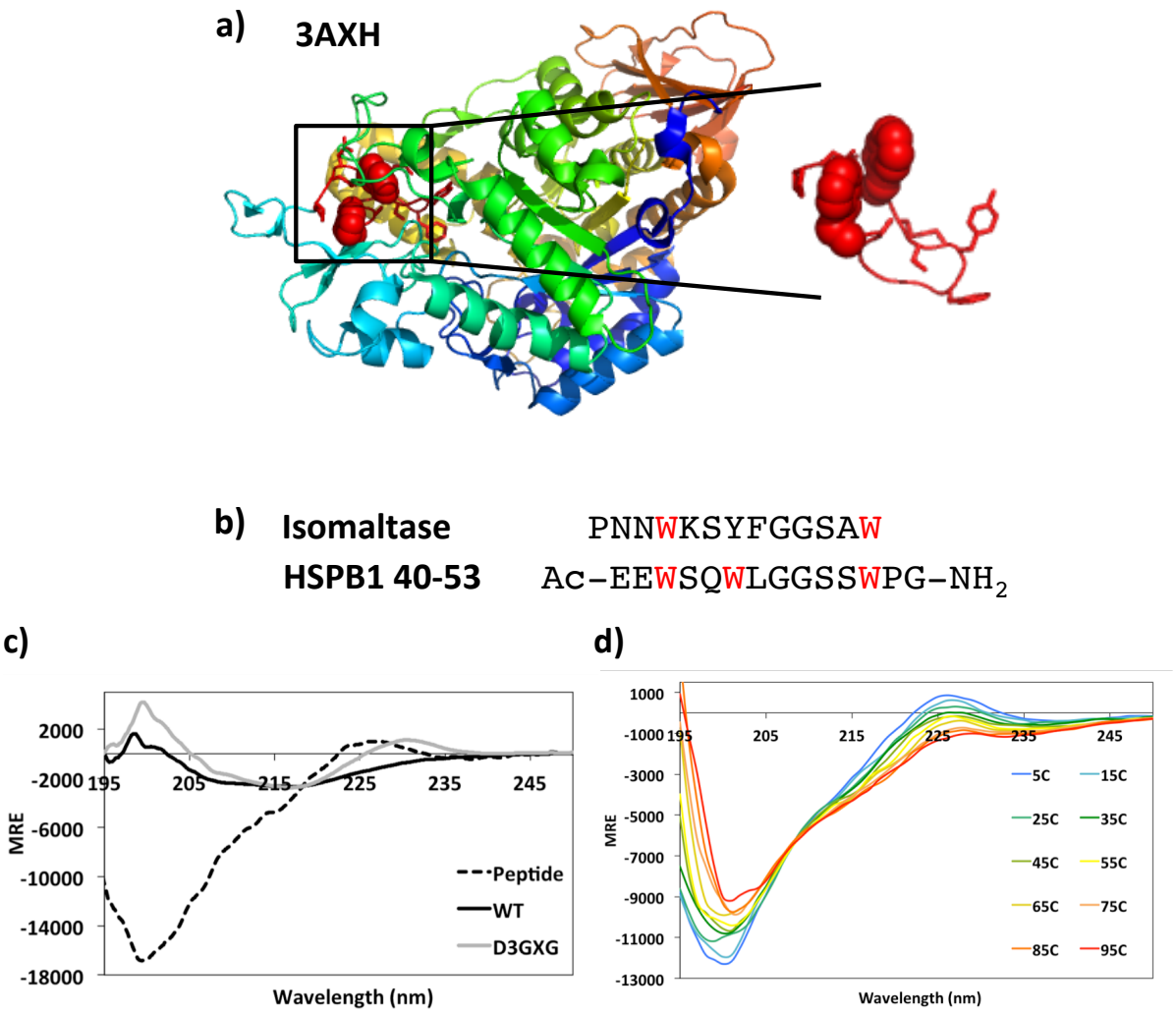


Figure 4.3- Analysis of HSPB1 NTR peptide consisting of residues 40-53. (a) In a search of the HSPB1 NTR sequence against the PDB, yeast isomaltase (PDBID: 3AXH) shows one region of sequence homology that forms a β-hairpin (red structure) containing stacked tryptophans. (b) Alignment of the isomaltase sequence depicted in red in (a) with residues 40-53 of HSPB1. Tryptophans are highlighted in red. This sequence of HSPB1 was ordered as a peptide with N-terminal acetylation and C-terminal amidation. (c) Far UV CD analysis of the NTR peptide in (b). 135μM peptide is compared with 20μM WT and D3GXG full-length proteins. All spectra were collected in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer at 20°C. The peptide yields a slightly positive peak around 225 nm but otherwise appears disordered (strong negative MRE at shorter wavelengths). (d) Far UV CD thermal melt of the peptide (same buffer and concentration as (c)) shows decrease of the 225 nm peak but a bump remains. There is also a decrease in ellipticity at shorter wavelengths, suggesting a non-cooperative denaturing of some type of secondary structure.

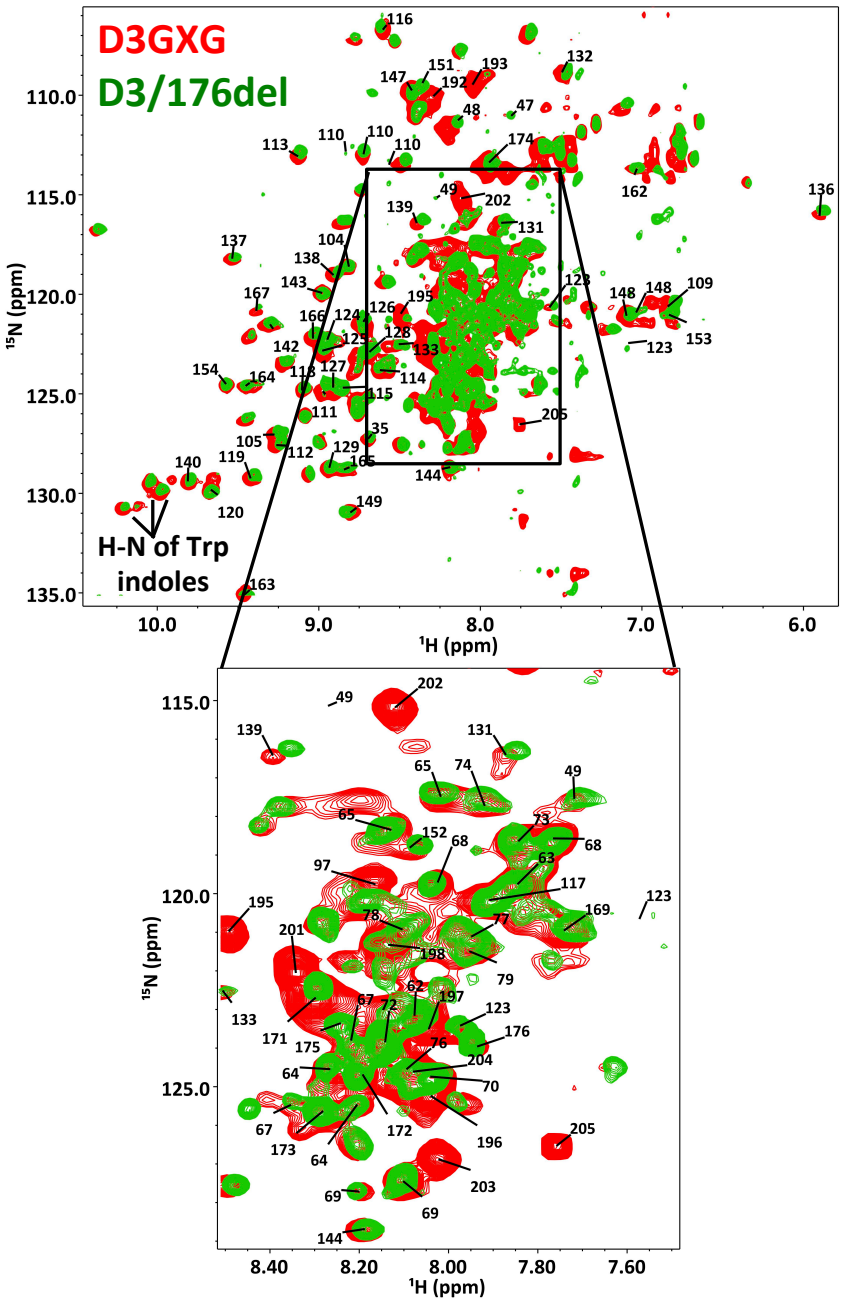


Figure 4.4- Comparison of D3/176del and D3GXG ^{1}H-^{15}N HSQC-TROSY spectra. The full-length dimer of HSPB1 yields a ^{1}H-^{15}N HSQC-TROSY spectrum with both dispersed peaks of moderate to weak intensity and many poorly dispersed peaks in the center of the spectrum that are very strong in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 30°C. The dispersed peaks generally correspond to more ordered regions of the protein (ACD), while the poorly dispersed peaks mostly correspond to disordered regions (NTR and CTR). The D3/176del mutant (green) has sharper peaks and less spectral overlap in the central region compared to D3GXG (red) due to the smaller size and truncation of the disordered CTR. Assignments to particular residues in full-length HSPB1 are indicated.

Table 4.1- Expected and observed resonances in D3/176del ^{1}H-^{15}N HSQC-TROSY spectra.

Total residues	176
Prolines	19
Expected backbone H-N resonances	159
Observed backbone H-N resonances	217
Expected Trp H-N resonances	6
Observed Trp H-N resonances	3

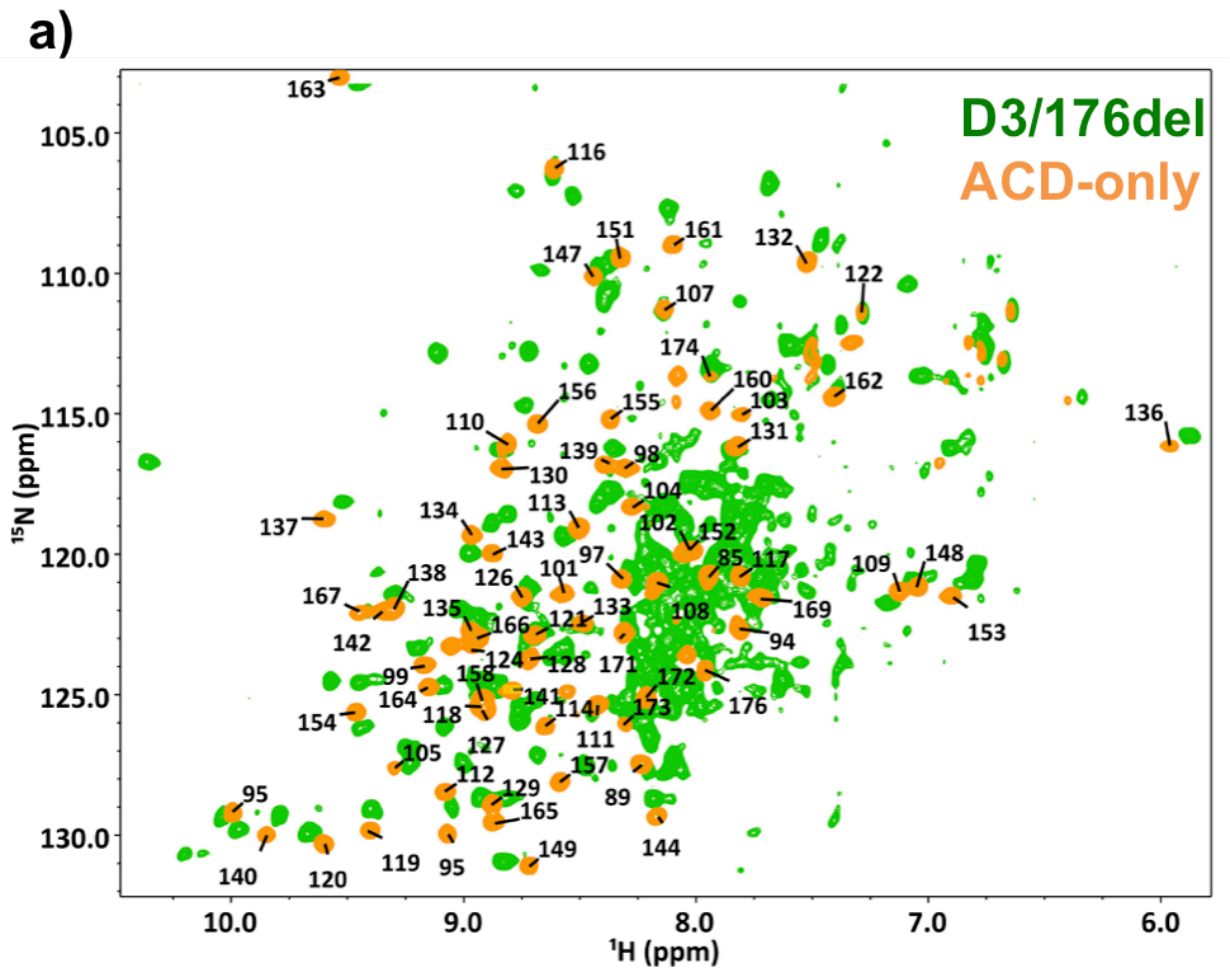
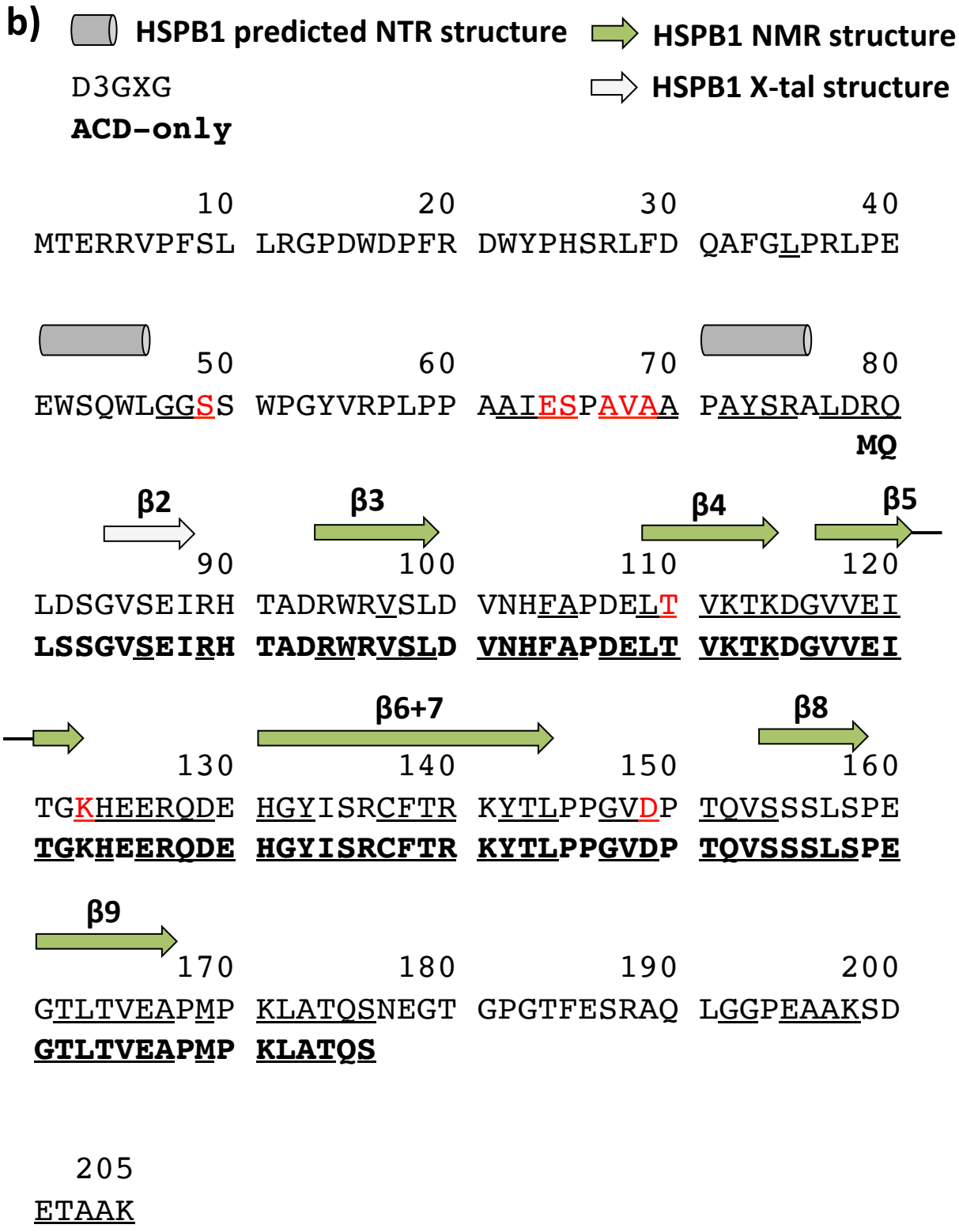
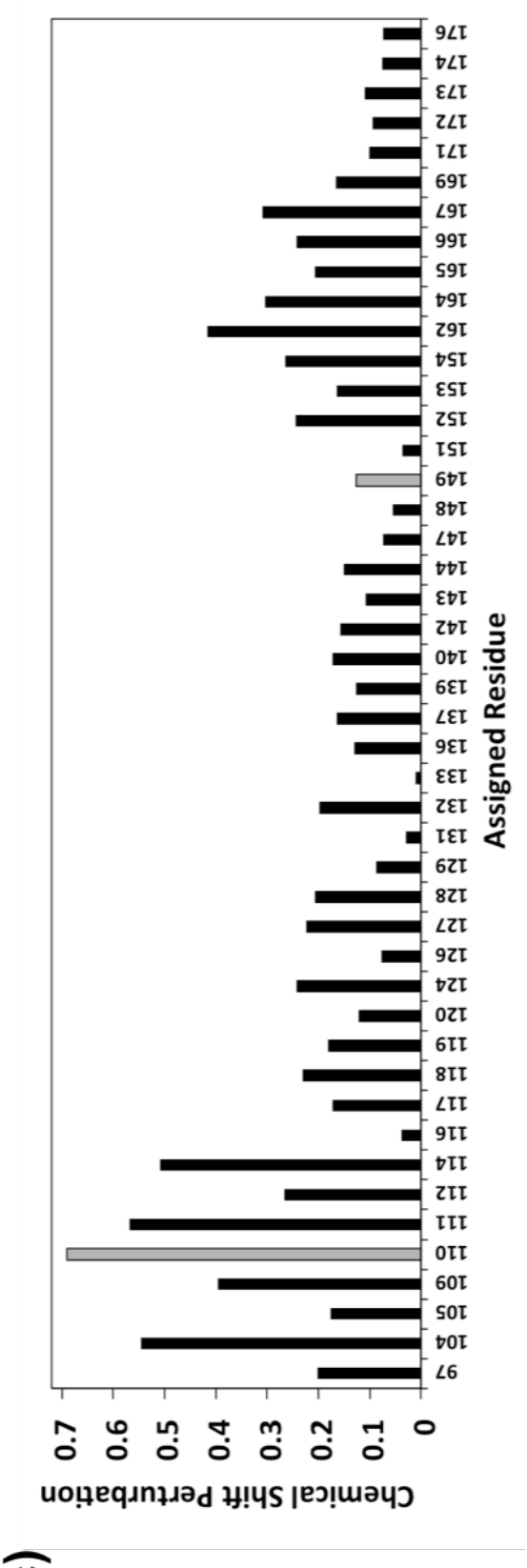
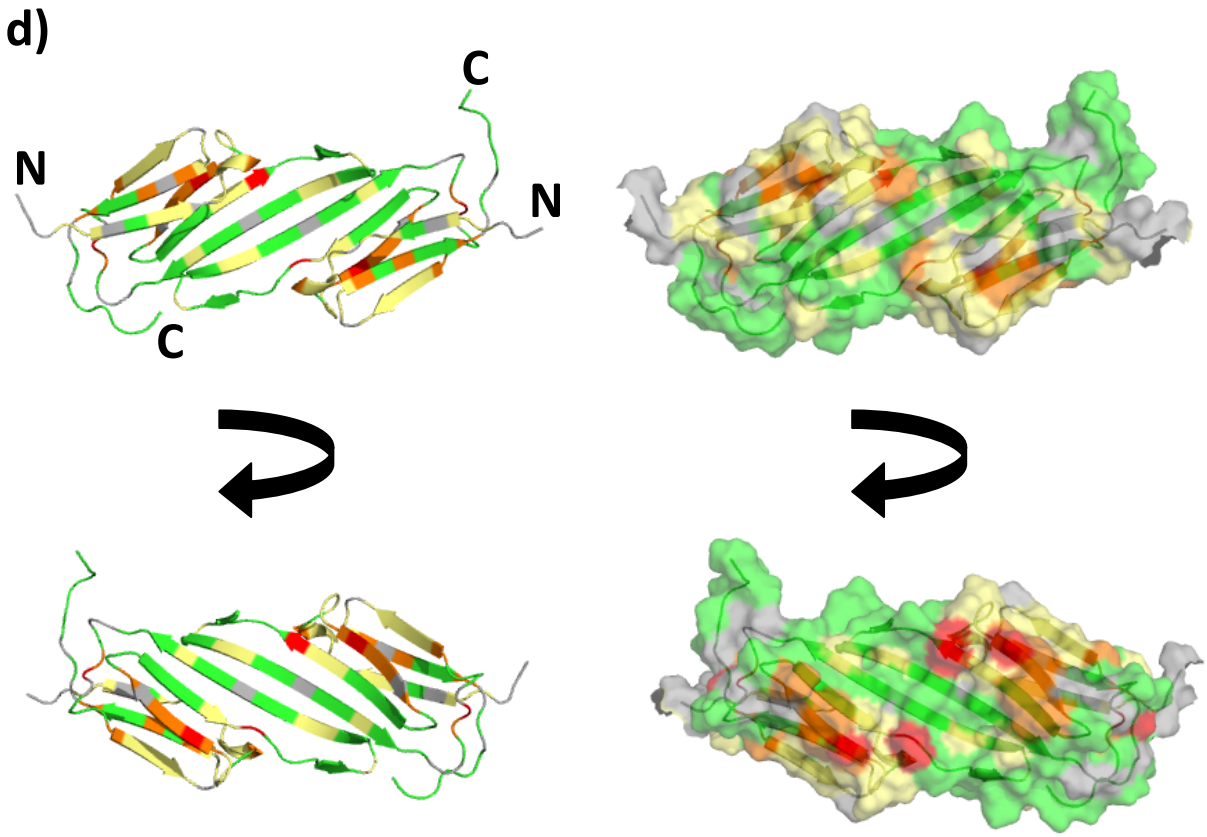


Figure 4.5- Comparison of ACD-only with D3/176del assignments, chemical shifts, and heterogeneity. (a) ^{1}H-^{15}N HSQC-TROSY spectra collected in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 30°C of ACD-only and D3/176del mutants of HSPB1. There are correlations between the two spectra but almost all peaks shift at least slightly. (b) Constructs used for assignment. Underlined residues are assigned under these conditions (more are assigned for ACD-only at lower temperatures). Residues colored red have multiple peaks in the spectrum. (c) Chemical shift perturbations (CSPs) between the two constructs. Gray bars are for residues that show multiple peaks, and the strongest intensity peak was selected for CSP analysis. All peaks have shifted but the strongest changes are observed in the loop preceding β4, β4, and β9. (d) Map of top 25% CSPs on the HSPB1 ACD NMR structure (PDB ID: 2N3J). Gray residues are unassigned for both constructs at these conditions. Green residues are assigned for both constructs and show minimal CSPs. Yellow residues are assigned in ACD-only but unassigned in D3/176del. Orange residues represent the top 25% CSPs between the two constructs. Red residues are assigned to multiple peaks (conformational heterogeneity) in the spectrum.


b)
HSPB1 predicted NTR structure
HSPB1 NMR structure
HSPB1 X-tal structure
D3GXG
ACD-only
10 20 30 40
MTERRVPFSL LRGPDWDPFR DWYPHSRLFD QAFGLPRLPE
50 60 70 80
EWSQWLGGSS WPGYVRPLPP AAIESPAVAA PAYSRALDRQ
MQ
β2
β3
β4
β5
90 100 110 120
LDSGVSEIRH TADRWRVSLD VNHFAPDELT VKTKDGVVEI
LSSGVSEIRH TADRWRVSLD VNHFAPDELT VKTKDGVVEI
β6+7
β8
130 140 150 160
TGKHEERQDE HGYISRCFTR KYTLPPGVDP TQVSSSLSPE
TGKHEERQDE HGYISRCFTR KYTLPPGVDP TQVSSSLSPE
β9
170 180 190 200
GTLTVEAPMP KLATQSNEGT GPGTFESRAQ LGGPEAAKSD
GTLTVEAPMP KLATQS
205
ETAAK


c)
Chemical Shift Perturbation
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Assigned Residue
97
104
105
109
110
111
112
114
116
117
118
119
120
124
126
127
128
129
131
132
133
136
137
139
140
142
143
144
147
148
149
151
152
153
154
162
164
165
166
167
169
171
172
173
174
176



Red- multiple peaks observed in D3/176del
Orange- top 25% CSPs
Yellow- assigned in ACD, not in D3/176del
Green- assigned in both ACD and D3/176del
Gray- unassigned

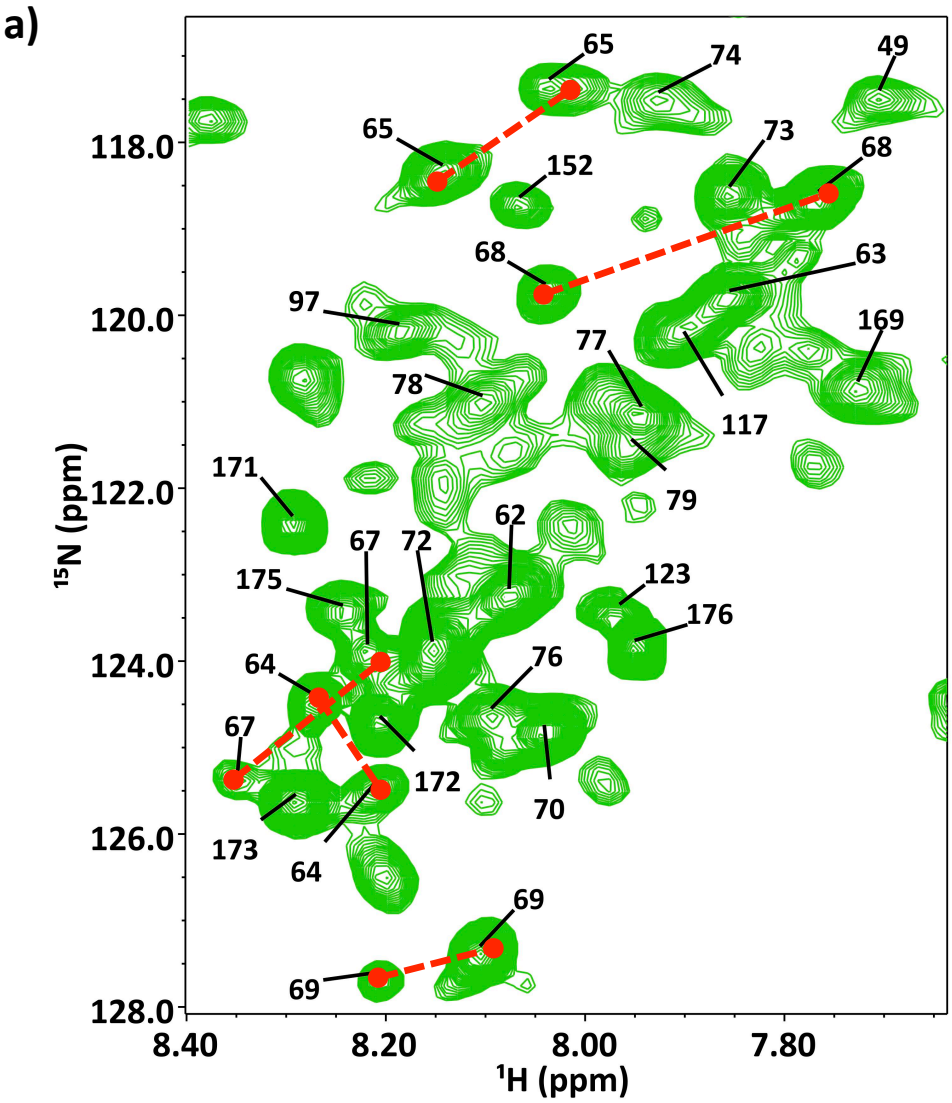
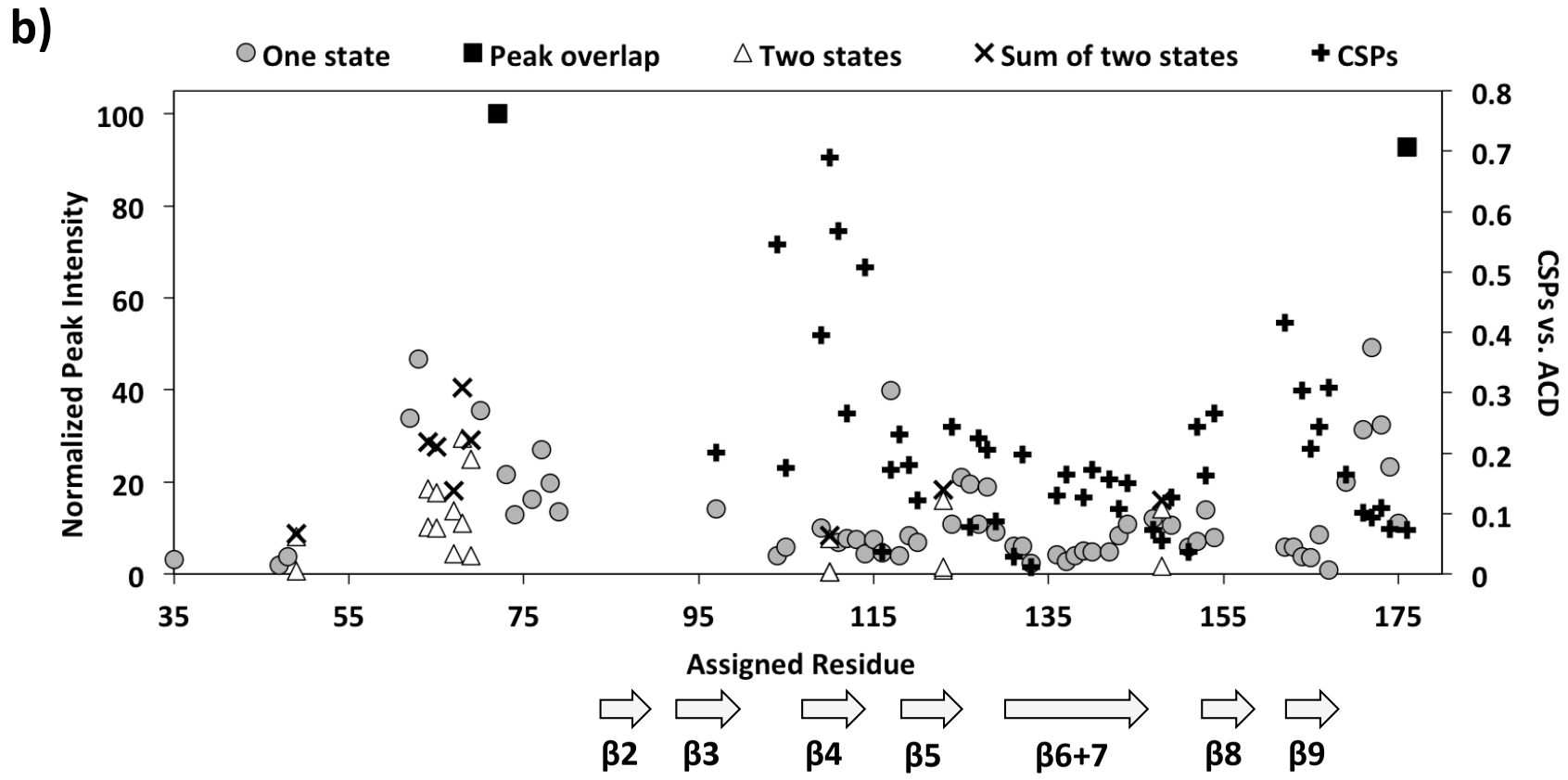


Figure 4.6- Two distinct sets of peaks observed for a region of the NTR in the D3/176del ^{1}H-^{15}N HSQC-TROSY spectrum. (a) The central region of the D3/176del ^{1}H-^{15}N HSQC-TROSY spectrum is shown (same as Figure 3.11), which generally has resonances corresponding to residues in unstructured regions. Dotted red lines connect distinct peaks that correspond to a single residue of the NTR. These residues include E64, S65, A67, V68, and A69. Both conformations observed are relatively disordered due to the resonance positions and intensities. (b) Intensities of assigned peaks were measured from the ^{1}H-^{15}N HSQC-TROSY spectrum and normalized from 0 to 100 between the weakest and strongest peaks shown. Gray circles correspond to residues with only one assigned resonance. White triangles correspond to residues with two or more assigned resonances. Black squares correspond to high intensity peaks that are also overlapping with other peaks, adding to their apparent intensity. Black x marks are the sum of intensities for residues with multiple peaks. Black + marks correspond to CSPs relative to ACD-only (secondary axis). Intensities of peaks in the 62-79 region of the NTR are comparable to those of the flexible CTR.


b)
One state
Peak overlap
Two states
Sum of two states
CSPs
Normalized Peak Intensity
CSPs vs. ACD
Assigned Residue
β2
β3
β4
β5
β6+7
β8
β9

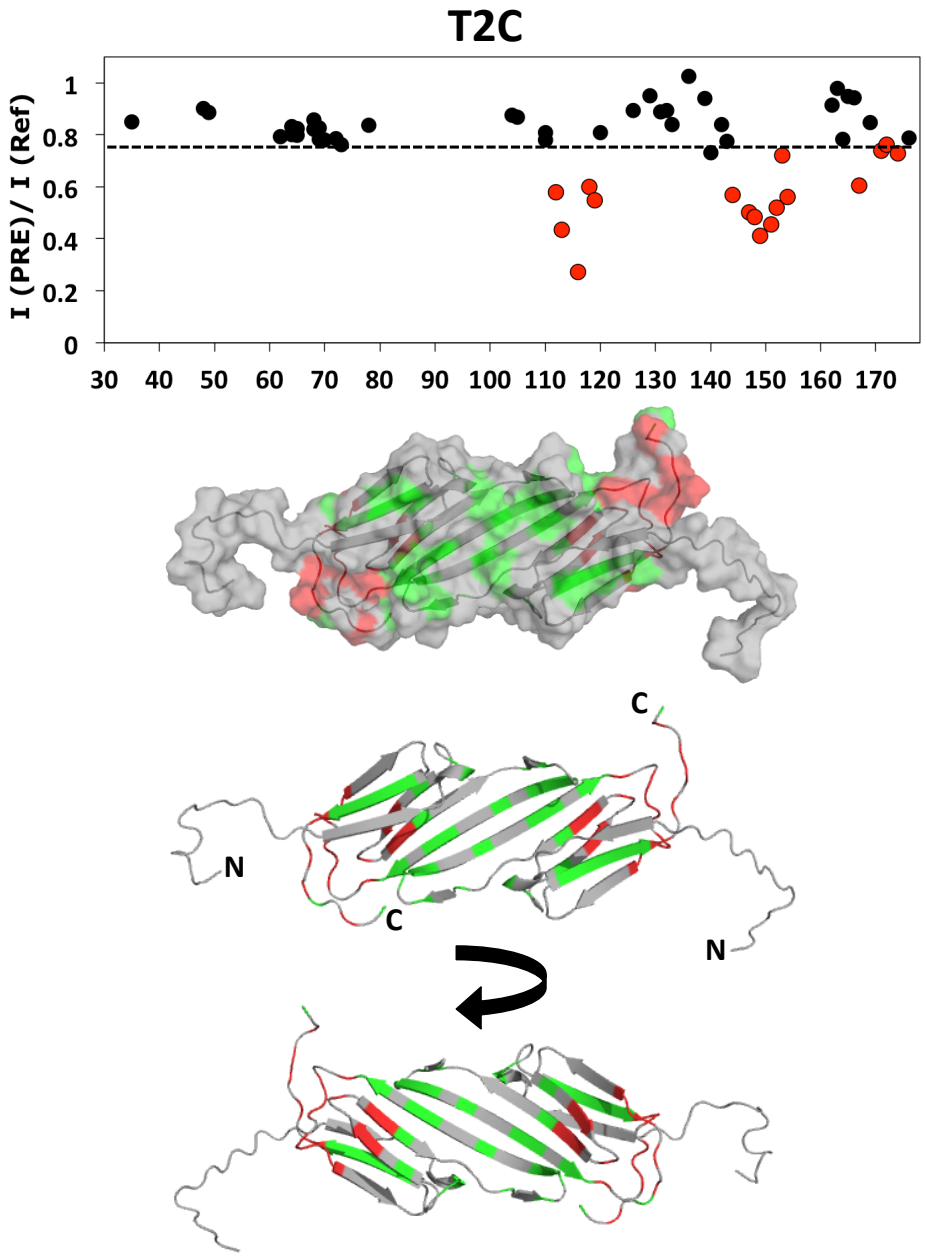
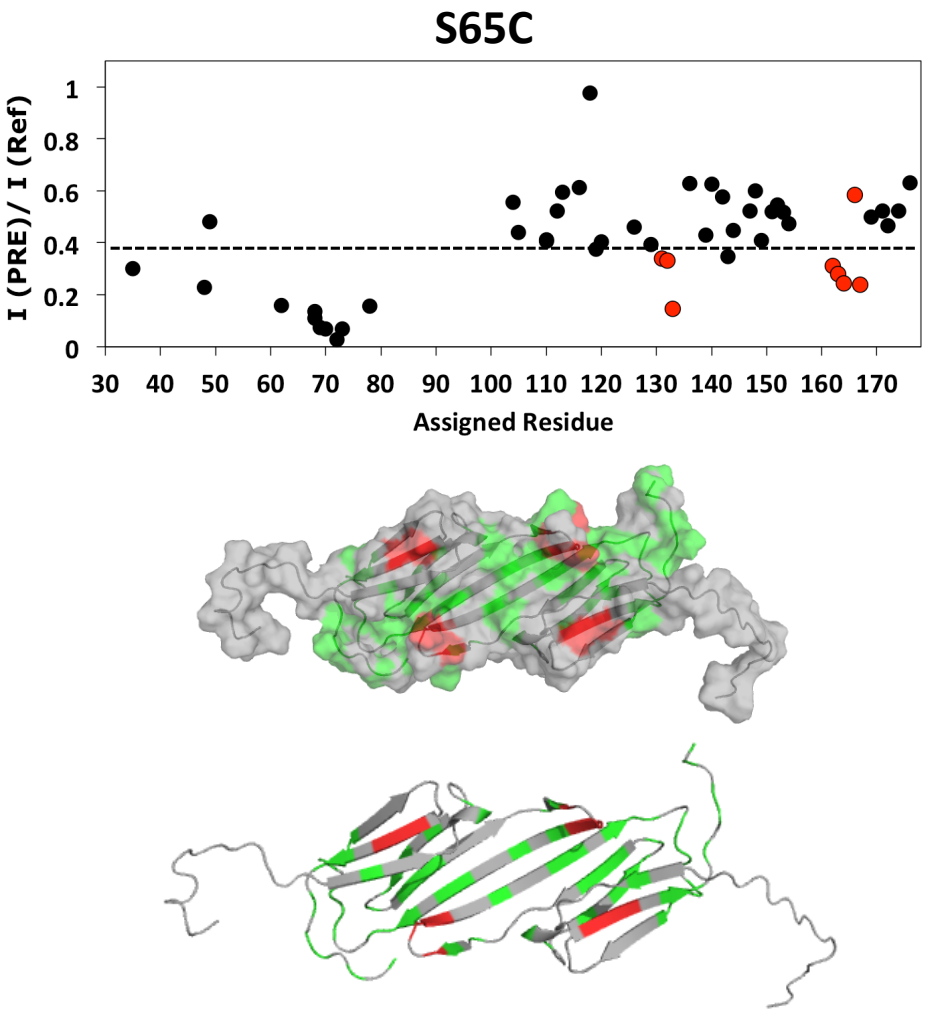
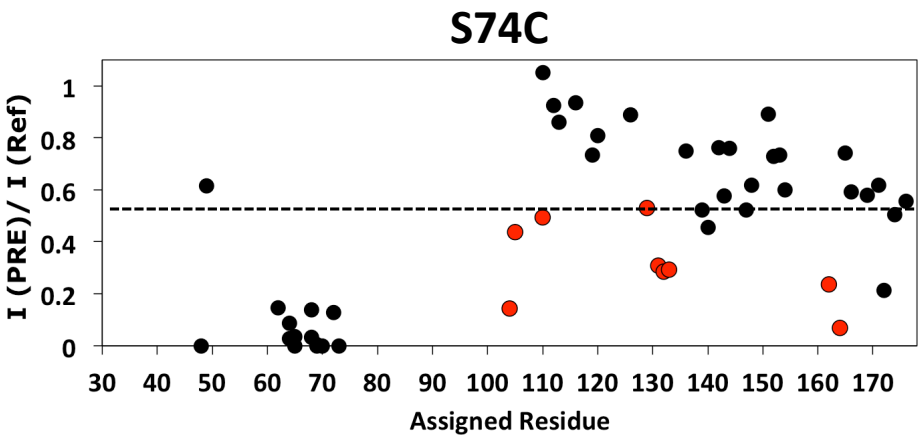
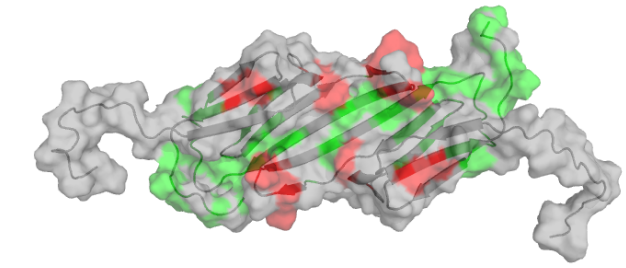
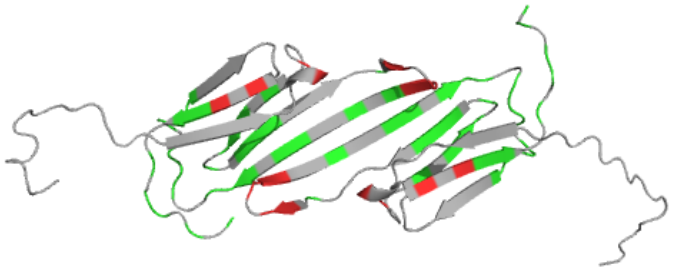


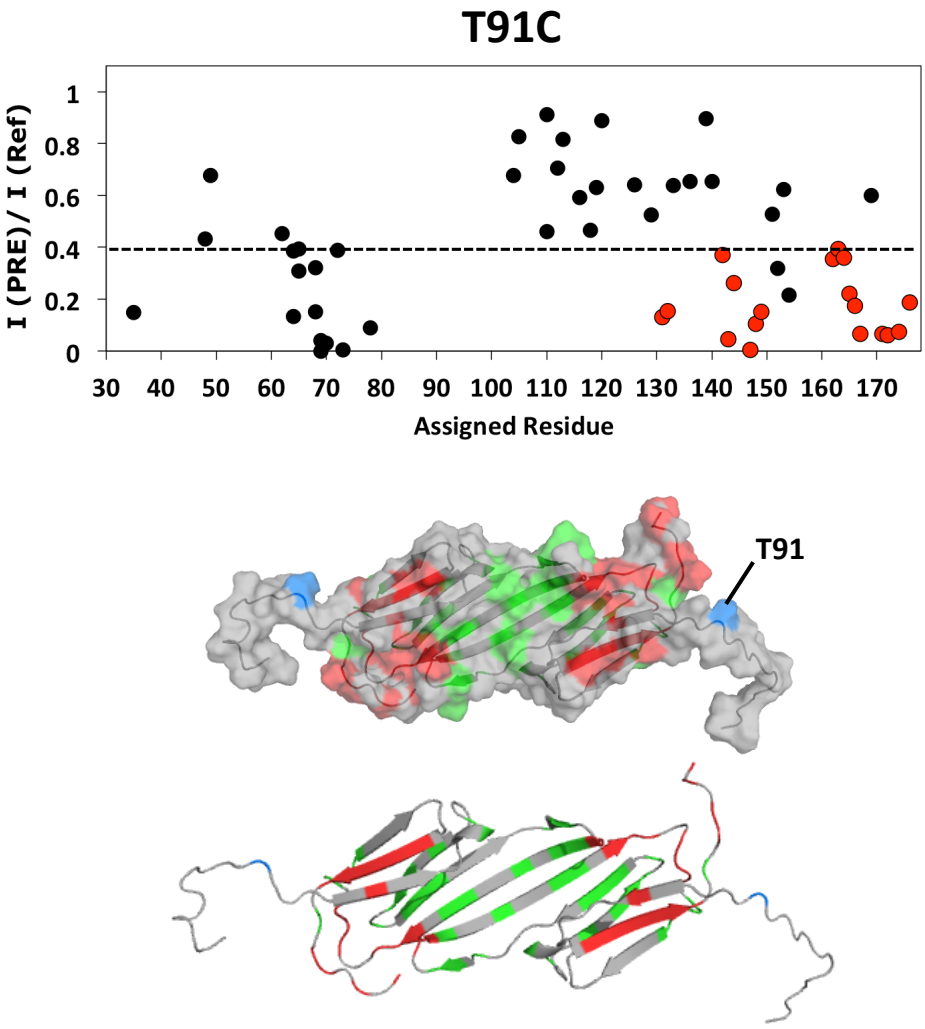
Figure 4.7- Interactions of the NTR with the ACD as seen with PRE probes. The T2C, S65C, S74C, and T91C mutations in the D3/176del/C137S background were well tolerated and spin labels could be introduced. ^{1}H-^{15}N HSQC-TROSY spectra were collected with the active spin label attached and control spectra were collected after, where ascorbate was used to quench the spin label. The ratio of intensities of active spin label spectra to quenched spectra is plotted for residues that gave strong enough peaks for analysis. Results are plotted on the HSPB1 ACD NMR structure (PDB ID: 2N3J) for each spin label position. Residues in gray are not analyzed or unassigned. Residues in green are relatively unaffected. Residues in red show a significant intensity loss and are part of a cluster of affected residues. These residues are highlighted in red on the plots. The blue residues for T91C are position T91 in the ACD structure, where this part of the truncated structure is disordered.


S65C
I (PRE)/ I (Ref)
1
0.8
0.6
0.4
0.2
0
30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Assigned Residue


S74C
I (PRE)/ I (Ref)
1
0.8
0.6
0.4
0.2
0
30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Assigned Residue






T91C
I (PRE)/ I (Ref)
1
0.8
0.6
0.4
0.2
0
30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Assigned Residue
T91

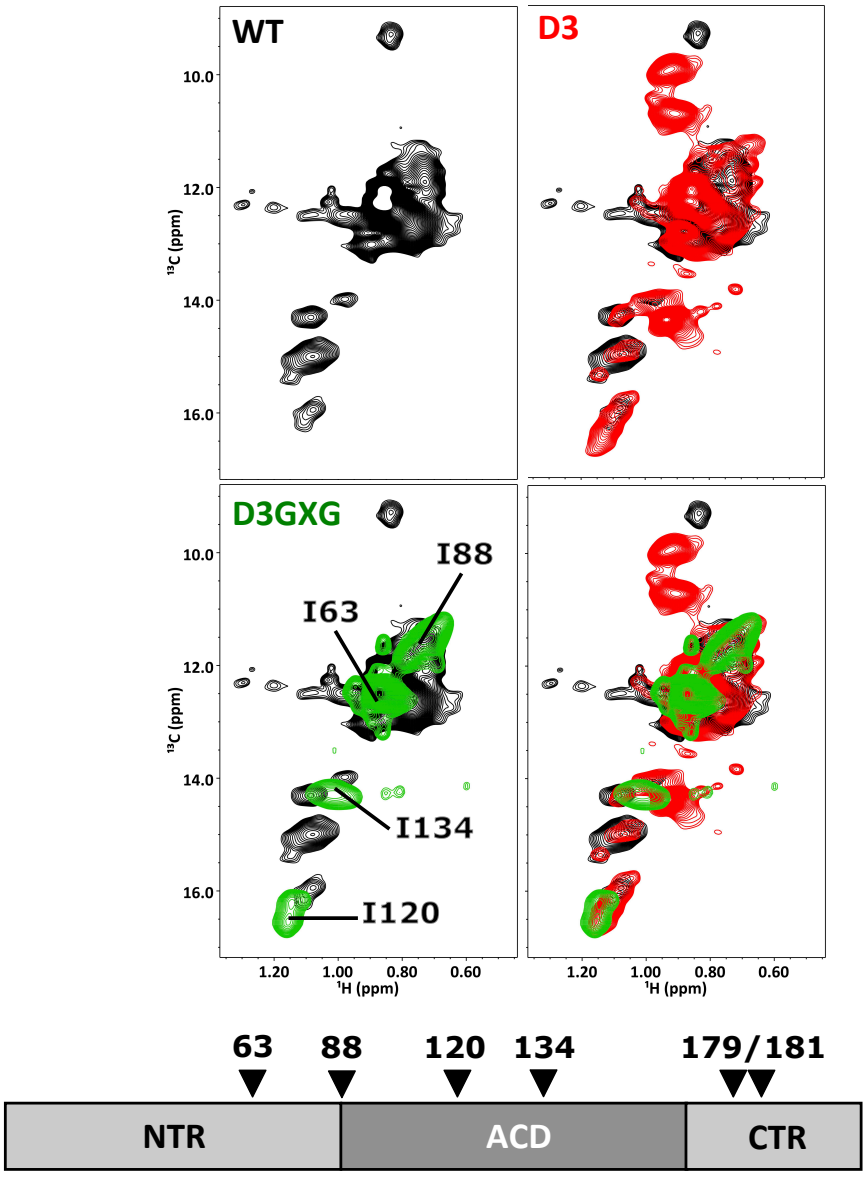


Figure 4.8- ^{1}H-^{13}C HMQC spectra of δ-methyl Ile-labeled WT and D3 oligomers. Samples were prepared in D_2O-based 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and spectra collected at 37°C. Peaks corresponding to δ-methyl groups of the six isoleucines in WT HSPB1 are shown in black. The diagram below the spectra indicates locations of the six isoleucines in the sequence. There are more than six peaks observed, therefore there are multiple slowly exchanging conformations present. Very strong peaks in the center of the spectrum (~12 ppm in ^{13}C) likely correspond to the unbound I179 and I181 methyl groups of the IXI motif in the CTR. D3 oligomers (red) also show peak heterogeneity, but there are additional peaks as positions not observed for WT oligomers. Therefore, some of the isoleucines in D3 are in different conformations from WT. The four isoleucines (I179G and I181G mutations are present) of D3GXG were assigned by mutational analysis and there are only four peaks observed in its spectrum (green).

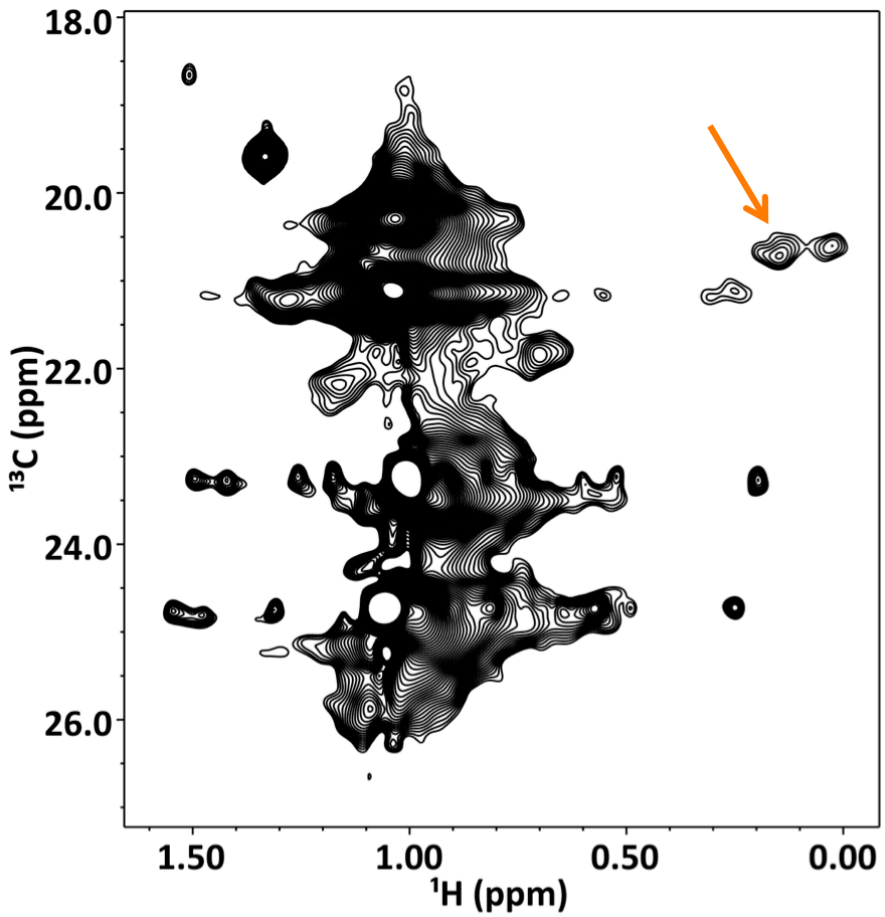


Figure 4.9- ^{1}H-^{13}C HMQC of δ-methyl Leu/γ-methyl Val labeled WT oligomers. The sample was prepared in D_2O-based 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and the spectrum collected at 37°C. Very strong and broad peaks in the center of the spectrum correspond to disordered conformations of the probed methyl peaks of leucines and valines. Unfortunately these peaks overwhelm the spectrum, making it difficult to resolve more dispersed, weak peaks. The arrow indicates peaks possibly attributable to ordered conformations of leucine and valine methyl groups in WT oligomers.

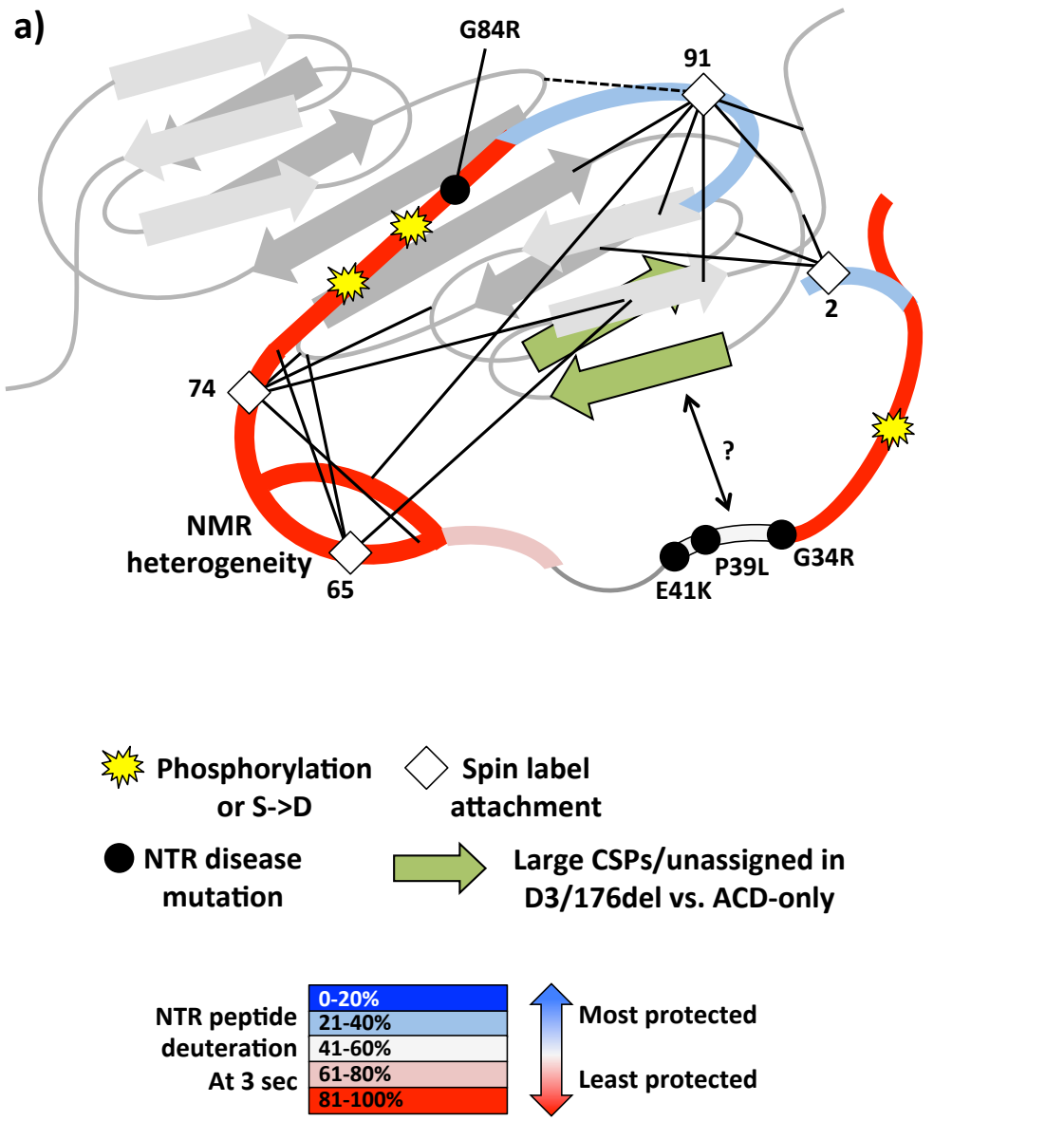
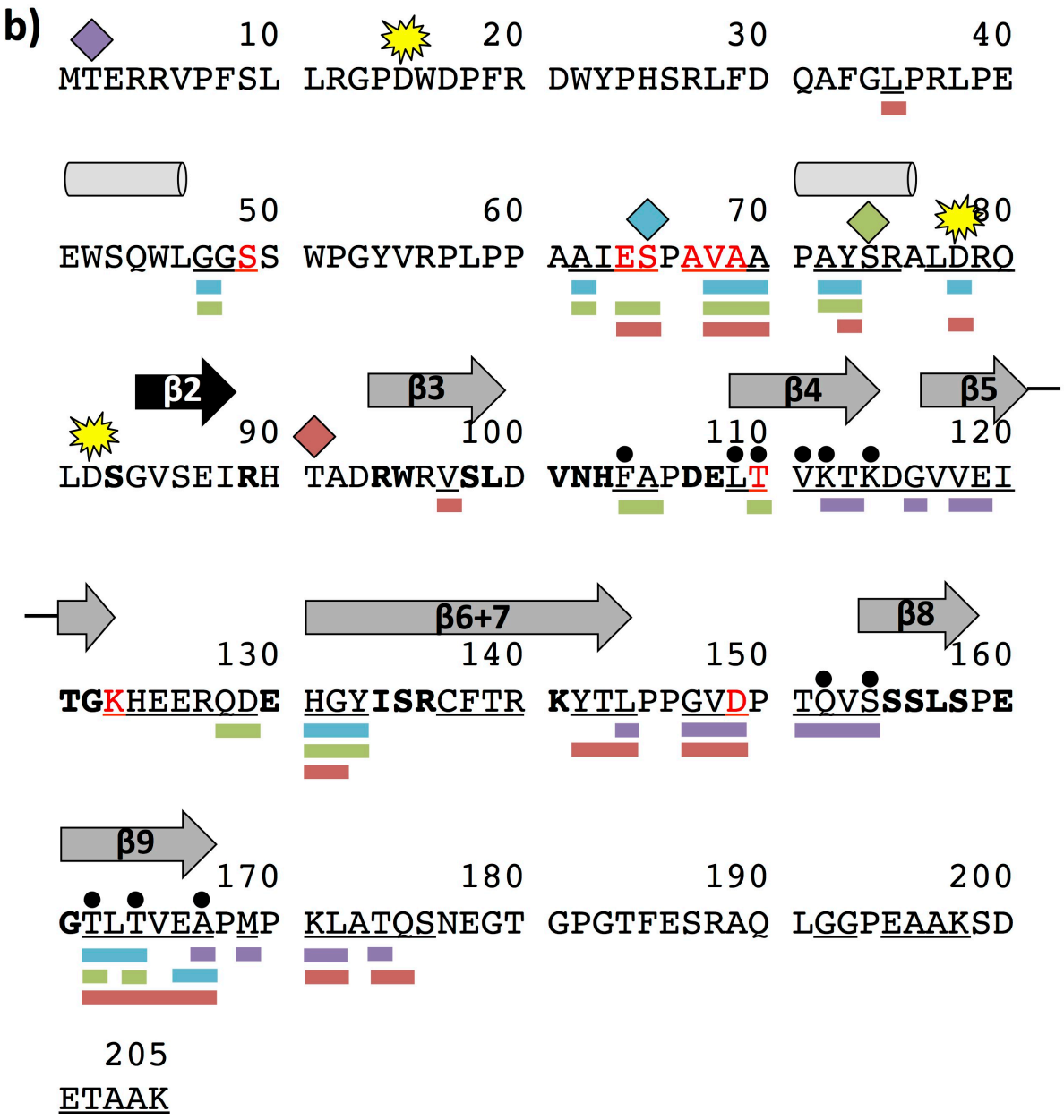
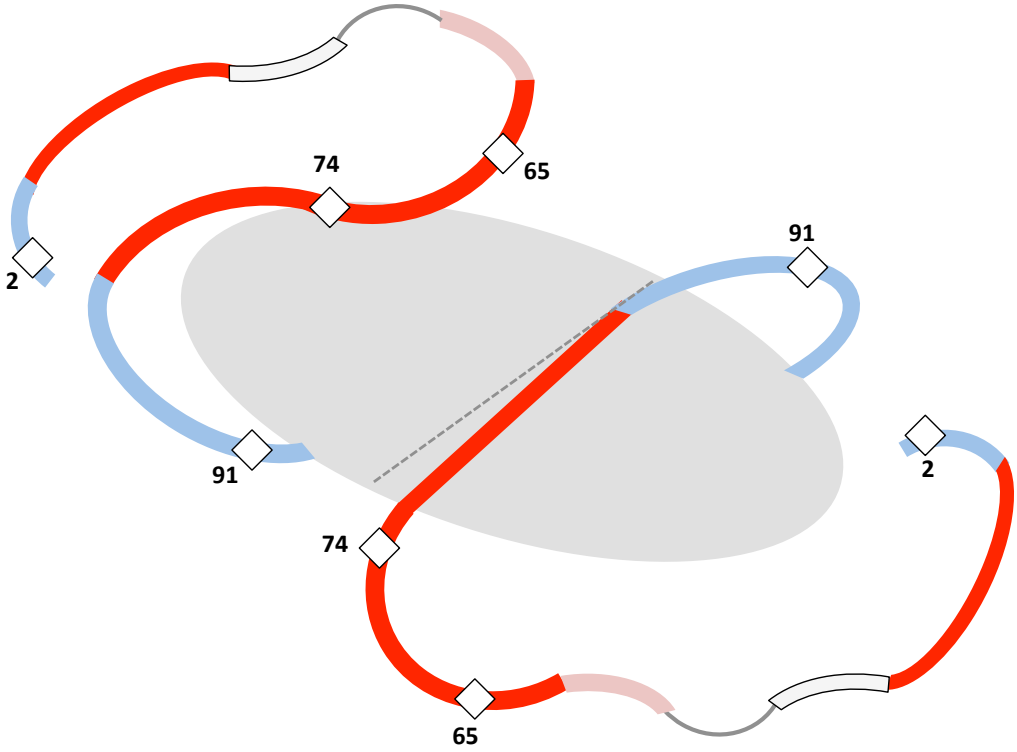


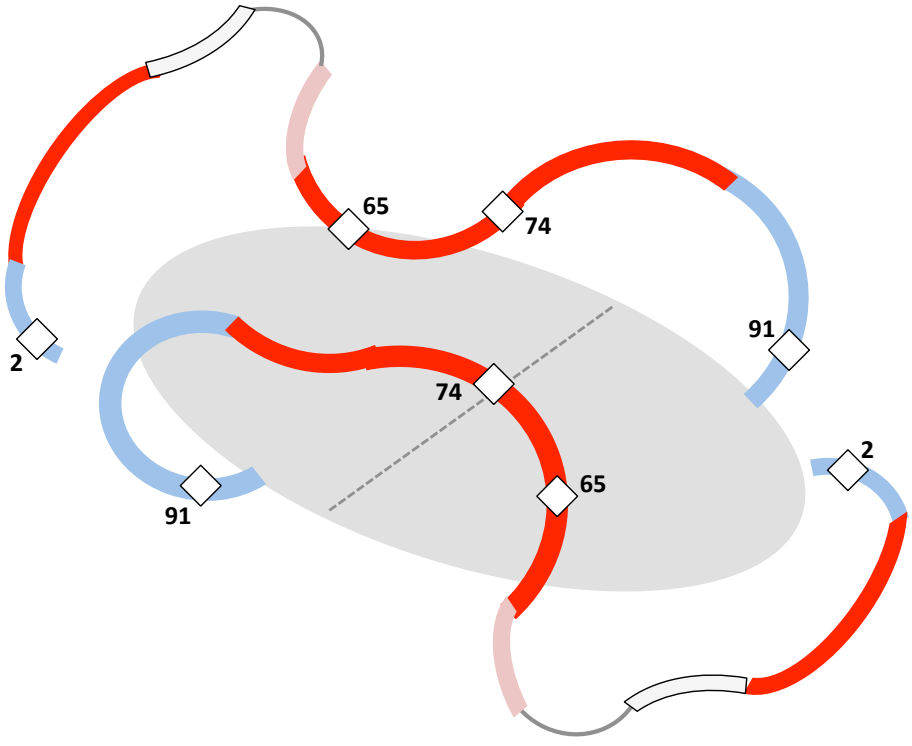
Figure 4.10- Model of structural features observed in D3GXG. (a) The NTR is colored in sections representing the local solvent protection at 3 seconds observed by HDXMS. Heterogeneity observed at the distal part of the NTR is represented by two positions of this region. Heterogeneity observed by NMR of the 60-70 region of the NTR is represented by an additional path. Spin label sites and their observed contacts are represented as diamonds and black lines, respectively. The dotted line emanating from spin label position 91 is presumably an intermolecular interaction due to likely conformational limitations at this site. In green are the β4 and β8 strands, which show large CSPs relative to the ACD-only construct or are unassigned. Presumably part of the NTR interacts with the β4/β8 region but the specific region of the NTR remains undetermined. (b) Properties incorporated into the structural model are displayed along the D3GXG sequence. (c and d) Alternate conformations of the NTR relative to the ACD are shown to highlight a variety of possible intra- and inter-contacts that would satisfy PRE restraints.


b)
10 20 30 40
MTERRVPFSL LRGPDWDPFR DWYPHSRLFD QAFGLPRLPE
50 60 70 80
EWSQWLGGSS WPGYVRPLPP AAIESPAVAA PAYSRALDRQ
β2 β3 β4 β5
90 100 110 120
LDSGVSEIRH TADRWRVSLD VNHFAPDELT VKTKDGVVEI
β6+7 β8
130 140 150 160
TGKHEERQDE HGYISRCFTR KYTLPPGVDP TQVSSSLSPE
β9
170 180 190 200
GTLTVEAPMP KLATQSNEGT GPGTFESRAQ LGGPEAAKSD
205
ETAAK
HSPB1 predicted NTR structure
HSPB1 NMR structure
HSPB1 X-tal structure
Phosphorylation or S->D
Large CSPs vs. ACD-only
Assigned
Unassigned vs. ACD-only
Multiple resonances
Spin label attachment/ contact

c) Intra-protomer NTR-ACD interactions



d) Inter-protomer NTR-ACD interactions



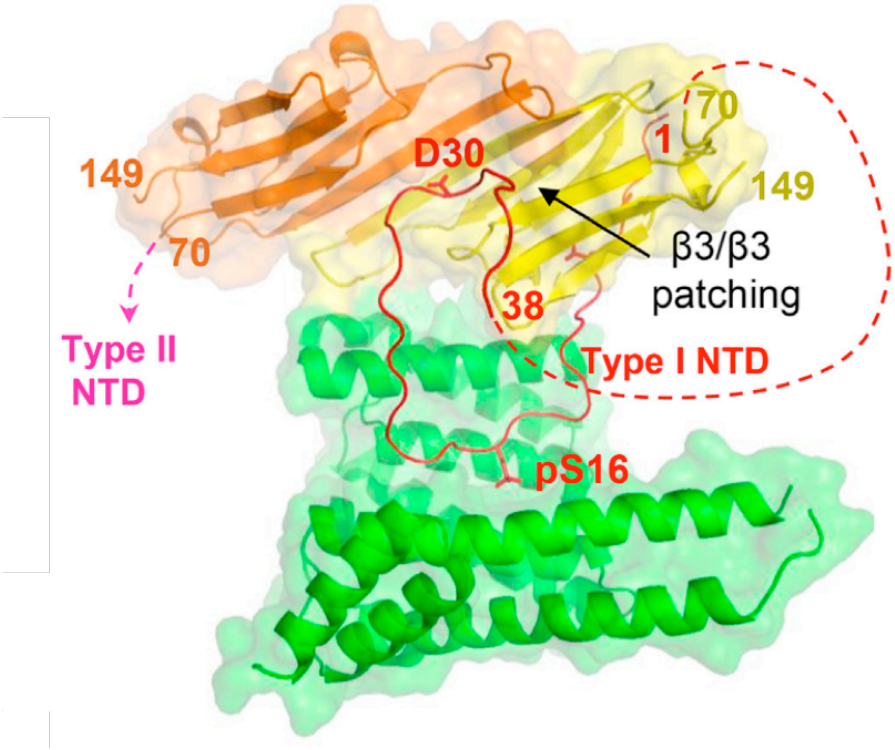


Figure 4.11- Interactions of the NTR with the ACD in HSPB6 dimers. Phosphorylated full-length HSPB6, a predominantly dimeric human sHSP, was crystallized in complex with a binding partner, the 14-3-3 signaling protein (Sluchanko et al. 2017, PDB ID: 5LTW). The distal part of the NTR, which contains an IXI motif, binds the β4/β8 groove. The highly conserved region of the NTR (around residue 30 for most sHSPs) is inserted into the β3/β3 groove.

5. MECHANISMS BY WHICH DISEASE-ASSOCIATED MUTATIONS IN THE NTR AFFECT HSPB1 OLIGOMERIZATION

5.1 Disease-associated mutants have substantial changes in global structure and function

Disease-associated mutations in the NTR of HSPB1 are found in two distinct regions. G34R, P39L, and E41K are clustered in the one highly conserved (amongst sHSPs) region that is predicted to have moderate helical propensity. G84R is located proximal to the ACD and is near two phosphorylation sites for HSPB1. Previous studies show that these mutants have disparate effects on chaperone activity (Muranova et al. 2015, Nefedova et al. 2013). Each of these mutants has also been reported to form larger oligomers than WT, as seen by earlier elution profiles in SEC experiments. Below are data that corroborate these findings on NTR disease mutants in HSPB1 and explore more properties of the mutants.

The ability to inhibit aggregation of the four mutants was compared to WT HSPB1. In a chaperone assay using α-lactalbumin as a model client protein, the mutations have disparate effects on the ability of HSPB1 to delay (lengthen the lag phase) reduction-induced aggregation (Figure 5.1). The client protein begins to aggregate around 20 minutes (increased absorbance due to light scattering of aggregates) under these conditions. WT protein delays aggregation until around 30 minutes. P39L shows a modest increase in chaperone activity relative to WT, delaying aggregation to 35 minutes. The other mutants are all slightly to modestly poorer chaperones relative to WT (do not lengthen lag phase as much), with E41K having the largest effect. These results are consistent with previous reports that G34R, E41K, and G84R are poorer chaperones than WT with other model client proteins, and P39L is either better than or similar to WT in activity.

Each disease mutant was analyzed by size-exclusion chromatography as well as SEC coupled with multi-angle light scattering (SEC-MALS). SEC is a standard method to probe quaternary structure of a protein and separates molecules by their hydrodynamic radius. Figure 5.2a shows SEC elution profiles of each disease mutant injected at 100μM and 20μM (monomer concentration). Each disease mutant elutes earlier than WT, with P39L eluting the earliest, consistent with previous reports (Muranova et al. 2015, Nefedova et al. 2013). Additionally, G84R at 20μM shows a large tail after the main peak, which was also previously observed. Therefore, G84R appears to dissociate at the lower concentration and could be considered a "weaker affinity" oligomer relative to WT. SEC-MALS calculates the average molar mass of the species eluting off the column based on light scattering and protein concentration. One can determine the approximate range of masses for a sHSP complex, and therefore estimate the number of subunits, without relying solely upon SEC elution standards. Figure 5.2b shows molar mass calculations (dotted lines) for WT and disease mutants injected at 100μM (monomer concentration) as they elute off a Superose 6 column on a different setup from that used in Figure 5.2a. All disease mutants elute slightly earlier than WT and have larger average molar masses with an estimated number of subunits shown in Table 5.1. The earlier elution volumes observed are consistent with previous findings. WT oligomers have an average molar mass of 488 kDa (across the top 50% 280 nm absorbance), corresponding to 22 subunits (rounded to nearest even number, as subunits are disulfide linked). P39L forms the largest oligomers, with a molar mass of 927 kDa, corresponding to 40 subunits. G34R and G84R elute similarly and have similar average molar masses of 730 and 751 kDa, corresponding to 32 and 34 subunits for each mutant, respectively. E41K has a slightly smaller molar mass than the other disease mutants of 639 kDa, corresponding to 28 subunits. P39L showed the highest level of polydispersity with the

estimated number of subunits ranging from 34 – 56 over the top 50% of the 280 nm chromatogram. E41K showed the next highest level of polydispersity, ranging from 24 – 38 subunits. G34R showed slightly increased polydispersity (28 – 38 subunits) relative to WT and G84R (18-24 and 30-36 subunits, respectively). It should be noted that there are two caveats with the calculated masses from SEC-MALS in this case. (1) The calculated molar mass at each point is a weight average of all species present. As there is a mixture of oligomeric species eluting at each point, the mass reported here is an overestimate relative to the median mass. (2) Appropriate alignment of MALS and concentration (absorbance or change in refractive index) curves is necessary for an accurate molar mass calculation. Alignment can only be performed on a monodisperse sample, therefore a large but more monodisperse standard (ferritin) was used to align curves for the sHSP mass calculations. Despite these possible sources of error, the relative changes in mass among these disease mutants and WT HSPB1 are reliable and confirm that earlier elution of these mutants is not due to a large change in hydrodynamic radius without an increase in the number of subunits.

To probe changes in secondary structure, each disease mutant was analyzed by far UV circular dichroism spectroscopy (CD). Figure 5.3 shows CD profiles of each disease mutant. The overall shape of the curve is preserved in each mutant relative to WT. However, the overall ellipticity increases (larger negative values centered around 200 nm). The positive 230 nm peak observed for small oligomers (D3 and D3GXG) is not observed for these mutants. The strongest change in ellipticity is observed for P39L. The mixture of secondary structures here is too complex for quantitative analysis, or estimating percentages of each secondary structure type. As random coil gives the weakest ellipticity, it can be inferred that there is not a large increase in

the proportion of disordered regions in each of these mutants. Instead, it is more likely that there is a slightly increase in helical and/or beta-strand content in these disease mutant oligomers.

It has been reported that subunits move into and out of sHSP oligomers relatively rapidly. Therefore measurement of subunit exchange kinetics provides a glimpse of global dynamics. Here "subunits" is used to describe a variety of possible exchanging species- ranging from monomer to a small multimers. Previous studies relied upon Forster Resonance Energy Transfer (FRET) between two different types of fluorophores, each placed on different subunits of a sHSP oligomer. We found that "homo-FRET" (due to substantial overlap of excitation and emission wavelengths) between neighboring identical fluorophores in a sHSP oligomer provides a robust and simpler readout for subunit exchange. Previous exchange studies of HSPB1 used the naturally occurring cysteine at its dimer interface (C137) for attaching fluorophores. However, placement of two fluorophores directly across from each other in the dimerized β-sheet could introduce a strong bias to the ACD-ACD interaction. Following the studies of exchange in HSPB5, an analogous position was selected in HSPB1 in a relatively non-conserved stretch proximal to the IXI motif (S153 in HSPB5, T174 in HSPB1). T174 was mutated to cysteine in cys-free (C137S) versions of WT and each disease mutant. AlexaFluor 488 maleimide was then conjugated at T174C and excess label removed by desalting column. Pre-equilibrated unlabeled and fluorophore-labeled proteins (identical in mutation and concentration) were mixed. The resulting decrease in homo-FRET (quenched state) and increase in fluorescence (dequenched state), was measured as a function of time as labeled subunits exchanged from heavily labeled oligomers into unlabeled oligomers (Figure 5.4). The resulting curves were fit to a single exponential equation to obtain exchange rates. The exchange rate for WT protein was found to be 0.48 per minute at 20μM protein at 37°C (Table 5.2). The calculated rates are consistent with

previous work (Cox et al. 2016; Lelj-Garolla and Mauk 2012). G34R, E41K, and G84R have similar exchange rates under the same conditions of 0.52, 0.45, and 0.42 per minute, respectively. In contrast, P39L oligomers undergo much slower subunit exchange at a rate of 0.15 per minutes at 20 μM and 37°C, around one third the rate of the other proteins. Similar experiments were performed for higher and lower concentrations of each mutant. There does not appear to be a significant change in exchange rate within the concentration range of 5 μM – 40 μM protein. Exchange rates for WT and P39L oligomers were also measured at 25°C. The calculated rates at 25°C were approximately one-tenth of the rates at 37°C (Table 5.2). The lack of dependence on concentration in these experiments suggests that the observed exchange rate is limited by the off-rate of subunits, which is generally concentration independent, as opposed to the on-rate, which is generally concentration dependent. Therefore, the slower exchange rate observed for P39L could be interpreted as a slower off-rate of subunits.

It is evident that P39L stands out amongst these mutants as having the most dramatic change in oligomeric size, a different propensity in chaperone activity, and the only mutant to have an effect on subunit exchange kinetics. However, the lack of clear correlation between any of these behaviors for the other mutants demanded a more detailed analysis of local structural effects.

5.2 Disease-associated mutants have subtle changes in local structure and dynamics

To identify regions of HSPB1 that exhibit changes in dynamics, the disease mutants were analyzed by hydrogen-deuterium exchange mass-spectrometry (HDXMS). This method probes the solvent accessibility of backbone amides at peptide-level resolution. WT/C137S, G34R/C137S, P39L/C137S, E41K/C137S, and G84R/C137S mutants were examined. All conditions were maintained from the experiment shown in Chapter 4 to compare WT/C137S

with D3GXG/C137S. The results in Figure 5.5 are presented in two different formats (a and b) and discussed below.

As in the previous experiment, the ACD of WT HSPB1 is highly protected, relative to the NTR and CTR, with most peptides showing only 50% deuterium exchange after one hour at room temperature (~25°C). The CTR is not protected, as expected for a disordered region, with the peptide including the IXI motif (179-185) being predominantly unprotected. A minor protected population is observed at three seconds that was not resolvable in the previous experiments. This is consistent with findings that only a small population of CTRs in an oligomer are bound to β4/β8 grooves.

The NTR is generally unprotected from solvent, as expected for a disordered region of a protein. Multiple peptides show different conformations, confirmed by several overlapping peptides (Appendix A). The most distal peptide (1-10) exists in two distinct states, one moderately protected and one highly unprotected. Given the challenge of quantitatively deconvoluting multiple states, the exact proportion and protection of each state should be considered only qualitatively. The fact that the protected population of the 1-10 peptide does not substantially decrease in proportion in the early three time points suggests that the exchange between these two states is very slow. The disappearance of the protected population by 4 minutes is most likely due to this structural state becoming fully exchanged (with deuterium). Note that a much longer period of time (~30 minutes) is required for subunit exchange amongst oligomers at this temperature (Table 5.2). Therefore the deuteration behavior observed for the protected state of this peptide is consistent with a local structural feature, such as formation of a helix or intramolecular β-strands. Peptide 11-28 shows slight protection at early time points. Peptide 29-41 also shows two distinct states of protection, similar to peptide 1-10; this peptide

contains the one highly conserved region of the NTR, has moderate helical propensity, and includes several sites of disease-associated mutation. The region consisting of residues 42-53 does not have coverage due to overlap of peptides containing this region with other species in their mass spectra, precluding clear analysis. Peptide 54-62, which encompasses the start of the insertion unique to HSPB1, shows a slight bimodal propensity at the first time point but is predominantly unprotected, representing a highly flexible and solvent accessible region of the NTR. Peptide 63-78 shows a complete lack of protection and indicates high flexibility. Finally, peptide 77-99, which is at the interface of the NTR and ACD and includes the sequence that might form the β2 strand, also shows two distinct states that are defined through the 4 minute time point. This bimodal behavior could be attributed to formation of the β2 strand (more protected state) in one population and a different environment of this sequence from the ACD in the other (less protected state). The deuteration profile presented here highlights regions of particular interest in the NTR where changes in protection might explain further global effects in an HSPB1 oligomer.

The disease mutants G34R and G84R have very similar NTR protection profiles to each other and differ from WT oligomers in similar ways. The 1-10 bimodal peptide shows relatively similar proportions and deuteration levels amongst WT and both mutants. Similarly peptides 11-28, 54-62, and 63-78 do not show substantially different behaviors for either mutant relative to WT. The lack of a bimodal distribution for peptide 54-62 in G84R could be due to an inability to distinguish the small, protected population that is slightly more deuterated (and therefore overlapping with the second state) in this case. Major differences between WT and these two mutants are seen for peptides 29-41 and 77-99. While two distinct states exist in these peptides for the WT protein, G34R and G84R appear to lose their protected states (or a small population

is seen at 3 seconds) and become completely unprotected in these regions. In the ACD, the only change observed is that both G34R and G84R mutants show less protection in the 100-109 peptide, which consists of the loop between β3 and β4, relative to WT. Taken together, these data suggest that in the G34R and G84R mutants, the β2 strand does not form, or at least does not form alongside the β3 strand. As with the D3GXG mutant presented in Chapter 4, it might be that the behavior of peptides 29-41 and 77-87 are coupled.

The E41K mutant shows a very similar deuteration profile to WT. Similar proportions of heterogeneity are observed for the 1-10 and 77-99 peptides as well as the relative deuteration levels of the two populations. Unfortunately, the E41K mutation disfavored pepsin cleavage near the mutation site and therefore matching or similar peptides (i.e. 29-41) could not be analyzed. Therefore, we can only speculate on whether E41K might behave WT-like or differently (loss of heterogeneity as in G34R and G84R) in this region of interest. Given the coupling of the loss of heterogeneity in peptides 29-41 and 77-99 in the G34R and G84R mutants, E41K likely retains heterogeneity in the 29-41 region as the 77-99 region is WT-like. Similarly, peptide 150-163 was not analyzable due to overlap in the mass spectra. Overall, E41K appears to have the most similar local dynamics to WT when compared to other NTR disease mutants.

The P39L mutant differs in its HDX profile from WT in ways different from the other two disease mutants. The 1-10 peptide still shows two distinct states existing in P39L, but the proportion of the protected state has increased slightly. Conversely, the proportion of the protected state of peptide 77-87 has decreased slightly in P39L. Again, as it is difficult to absolutely deconvolute the two states in these peptides and alternate deconvolutions may exist (i.e. smaller proportion of a more protected state vs. larger proportion of a less protected state), it

is more prudent to describe the deuteration profile of P39L in these regions as notably different from WT. The most dramatic change in P39L behavior occurs in the 29-41 peptide. Bimodal behavior is still observed, but the protected state is even more protected from solvent exchange. The simplest explanation is a stabilization of this protected state in the P39L mutant. As the deuteration profiles of the remaining peptides (including the ACD) of P39L appear very similar to their WT counterparts, there is no clear suggestion that other regions are coupled with the protected conformational state of this region. This is the only region to become overall more protected in P39L, and this is not the case for the other disease mutants. Therefore there is likely a connection between the 29-41 region and the greatly reduced subunit exchange kinetics observed for P39L.

5.3 Disease-associated mutations cause disparate global effects in a dimer-prone mutant

To further probe local structural effects of disease-associated mutations in the NTR, each mutation was placed into the D3GXG context. Figure 5.6 shows a chaperone assay in which WT, D3GXG, and each disease mutant in D3GXG are compared. As shown in Chapter 3, D3GXG is a much more potent chaperone than WT, specifically at delaying the aggregation of α-lactalbumin. G34R/D3GXG and G84R/D3GXG retain this similar behavior. E41K/D3GXG is a moderately weaker chaperone than D3GXG but is still much better than WT. P39L/D3GXG is the weakest chaperone of the disease mutant/D3GXG combinations and only delays aggregation slightly past WT. Specifically, the D3GXG mutations do not afford P39L any additional chaperone activity. The relative trends among mutants are in contrast to the disease mutants not in the D3GXG context (Figure 5.1), where P39L is the strongest chaperone.

Disease mutations in the D3GXG context were analyzed by SEC and CD to probe for changes in global and local structure. Interestingly, two of the disease mutations in D3GXG are

not predominantly dimeric. Figure 5.7a shows that P39L/D3GXG and E41K/D3GXG both form very large, polydisperse oligomers that do not readily dissociate at lower concentrations. G34R/D3GXG remains in a small oligomer/dimer form but has a very different elution profile from D3GXG. At higher concentrations there is an earlier eluting shoulder, possibly corresponding to small oligomers. At lower concentrations, the protein elutes much later and broader than D3GXG, even though the protein is under non-reducing conditions (disulfide linked dimer interface). This suggests that G34R/D3GXG (presumably) dimers adopt a more compact conformation that travels through a larger volume in the column. G84R/D3GXG behaves very similarly to D3GXG by SEC, eluting as a sharp peak at both concentrations. Figure 5.7b shows CD profiles of each of these mutants relative to D3GXG and their non-D3GXG counterparts. P39L/D3GXG adopts very similar secondary structure to P39L. E41K/D3GXG has reduced ellipticity, therefore this combination of mutations may destabilize the protein slightly. However, the E41K/D3GXG profile looks highly similar to that of the GXG mutant, which also forms large oligomers. It might be that the E41K mutation "rescues" the effects of the D3 mutations, leaving only GXG-like observable effects on local structure. Although forming small oligomers/dimers, G34R/D3GXG does not adopt D3GXG-like secondary structure and remains similar to G34R. This suggests that the G34R mutation might disfavor the structural feature responsible for producing the positive 230 nm feature observed in D3GXG and other small phosphorylation-mimicking oligomers. G84R/D3GXG completely adopts D3GXG-like secondary structure. Therefore, NTR disease mutants have distinct effects on local structure that then impact oligomerization propensity.

5.4 Sites of local changes in structure and dynamics in disease-associated mutations in D3GXG

Each disease mutant in the D3GXG/C137S context was analyzed by HDXMS (Figure 5.8) comparably to the disease mutants in Figure 5.5. Although different peptides are favored for cleavage due to the D3/GXG mutations (Appendix A), almost entirely identical peptides are presented in Figure 5.8. Peptide 164-185 for D3GXG mutants is presented in contrast to separate 164-178 and 179-185 peptides for WT and disease mutants, as the GXG mutations disfavor cleavage in this region.

The ACD is similarly solvent protected in all mutants, however moderate changes are observed in two regions. In the D3GXG context, G34R and P39L mutants are somewhat less protected in peptides 100-109 and 150-163 relative to D3GXG and their non-D3GXG counterparts. Peptide 100-109 corresponds to the loops between β3 and β4 strands, therefore it is possible that this loop becomes more flexible in these contexts, perhaps due to the absence of a transient interaction with another part of the protein. Similarly, peptide 150-163 corresponds to the β8 strand at the edge of the ACD. As this strand is unobservable by NMR in the D3GXG mutant and is presumed to be interacting with another region of the protein, a loss of solvent protection suggests a loss of this potential interaction.

Of the observable peptides belonging to the CTR, there are no major differences among any of these mutants. The 186-205 peptide remains completely solvent exposed for all cases. The 164-185 peptide, which is relatively large (i.e. more challenging to interpret) and contains the β9 strand, shows slight protection at early time points for D3GXG and G34R/D3GXG. The peptide is completely unprotected for P39L/D3GXG and G84R/D3GXG. It was not possible to analyze the 164-185 peptide for E41K/D3GXG due to spectral overlap.

In the NTR, some differences are observed between D3GXG and each disease mutant in the D3GXG context. The 1-10 peptide shows a high proportion of the protected state at early time points for D3GXG. This high proportion is also observed in E41K/D3GXG and G84R/D3GXG; in the latter case, the unprotected state is not observed or overlaps too much with the protected state in the mass spectra. G34R/D3GXG and P39L/D3GXG, however, showed decreased propensities of the protected state of the 1-10 peptide. Although the 29-41 peptide for P39L/D3GXG could not be analyzed, a similar peptide (29-38) is shown for comparison. While D3GXG and the other disease mutants are largely unprotected in this region, P39L/D3GXG somewhat retains protected, heterogeneous states as observed in P39L-only. For D3GXG and all mutants in this context except for E41K, the 77-99 peptide shows only one relatively unprotected state (an additional protected state is observed in WT, P39L, and E41K oligomers). E41K/D3GXG, however, retains the heterogeneity observed in this region for E41K-only. The remaining NTR peptides (11-28, 54-62, and 63-76) do not show substantial changes in protection among these D3GXG mutants.

In D3GXG and disease/D3GXG mutants that form smaller oligomers, there is generally less protection observed for most peptides. However, peptide 33-41 represents a case of moderately increased protection for D3GXG relative to WT as shown in Figure 5.8b. This peptide could be analyzed for G34R and G84R mutants, which each form large oligomers, and it is similarly unprotected compared to WT. Peptide 33-41 in G84R/D3GXG, which remains a D3GXG-like dimer with a D3GXG-like CD signal, also remains somewhat protected, similar to D3GXG. In contrast, this peptide in P39L/D3GXG, which forms very large oligomers, is highly unprotected. The increased protection observed for this peptide in mutants that form dimers suggests that a new secondary structure feature or interaction might be formed.

As both G34R/D3GXG and G84R/D3GXG form small oligomers/dimers, they were analyzed by NMR spectroscopy. Figure 5.9 shows overlaid ^{1}H-^{15}N HSQC-TROSY spectra of D3/176del and G84R/D3GXG. G84R/D3GXG was not deuterated and there was no analogous spectrum of undeuterated D3GXG on the same magnet, therefore a spectrum of undeuterated D3/176del is used for comparison. The G84R/D3GXG spectrum is highly similar to its non-disease counterpart, suggesting that overall structural features are preserved. This is consistent with SEC and CD profiles of G84R/D3GXG compared to D3GXG. However, a number of peaks shift slightly to moderately, suggesting subtle changes in the chemical environments of the corresponding residues. Of the assigned peaks in D3GXG, these affected peaks correspond to residues at the dimer interface (136-140) and two residues in β9 (163 and 164). A quantitative analysis was not performed due to the many differences between these two samples (length of construct, protein concentration). The shifts observed at the dimer interface are reminiscent of behavior of the ACD-only at low pH, where the intricate charge network of ACD side chains is perturbed. G34R/D3GXG did not yield an ideal spectrum and only a few peaks corresponding to the flexible CTR were observable. This suggests that even at "low" concentrations for NMR (50 μM), self-association of G34R/D3GXG is causing peaks to be broadened out.

5.5 Discussion of structure-function relationship in disease-associated mutants of HSPB1

Multiple properties of NTR disease-associated mutants were examined, including probing their behavior in the D3GXG context. This study provided valuable insights into the local and subsequent global structural effects that could mechanistically explain changes in function. Results for each mutant are described below and summarized in Table 5.3.

G34R- The glycine at position 34 of HSPB1 is conserved among both orthologs and paralogs and is at the end of the small stretch of the NTR that is conserved among sHSPs. In the

HSPB5 oligomeric model (Jehle et al. 2011), this glycine is in a helix modeled into the NTR. Glycines provide a high level of flexibility, as there is no side chain present. The effects of the arginine mutation at this position could be attributed to both a loss in flexibility and the introduction of a large, charged side chain. Relative to WT HSPB1, G34R is a slightly poorer chaperone for the model client protein used here- α-lactalbumin. This result is consistent with poorer activity observed for other client proteins (Muranova et al. 2015). G34R oligomers elute earlier than WT and appear to have a similar stability (not dissociating) in the concentration range examined. The average number of subunits in G34R oligomers increases from 22 to 32 relative to WT oligomers and the polydispersity increases slightly. The additional subunits could be incorporated in numerous ways that would be confirmed best by electron microscopy. Although the polydispersity would inhibit determination of high-resolution structures, the general dimensions could be elucidated. Both WT and G34R oligomers undergo subunit exchange at very similar rates, therefore the global dynamics are not affected by this mutation. However, there are changes in local dynamics. By HDXMS, it is evident that the protected conformations observed in peptides 29-41 and 77-99 are lost in the G34R mutant. As the G34R mutation is within the 29-41 peptide it is not surprising that there appears to be a change in secondary structure. Specifically, any feature that was formed in this region in WT oligomers is disfavored by the G34R mutation. The loss of a protected state in the 77-99 region is an indirect effect that is possibly coupled with the effects seen in 29-41. Specifically, the protected conformations seen in these two regions are likely coupled and only appear as a pair in the subsequent mutants discussed.

The G34R/D3GXG mutant adopts both strong D3GXG-like chaperone activity and forms small oligomers. Therefore, the G34R mutation does not appear to promote oligomerization.

However, the SEC elution profile of G34R/D3GXG suggests there are differences in shape of the molecule relative to D3GXG. Specifically, the G34R/D3GXG dimer (disulfide linked) elutes later than D3GXG, suggesting a more compact average conformation. The later elution could also be due to the mutant interacting with the column resin, as different hydrophobic regions might be exposed. Although G34R/D3GXG forms dimers/small oligomers, it does not adopt D3GXG-like secondary structure as seen by CD. Therefore, the structural feature attributable to the positive 230 nm peak in D3GXG (likely an exciton couplet between tryptophans) is not present in G34R/D3GXG. Relative to G34R, G34R/D3GXG does not show any substantial changes in local dynamics by HDXMS except for moderately decreased protection of peptides 100-109 and 150-163. Loss of protection in these parts of the ACD suggest a loss of transient interactions between the ACD and parts of the NTR (discussed further in Chapter 6). Overall, it is evident that the G34R mutation in the D3GXG context likely adopts different local (secondary and tertiary) structure from D3GXG, but this does not impact chaperone function with the model client examined here.

P39L- The proline at position 39 of HSPB1 is conserved among orthologs but not among paralogs. In fact, this position is a leucine in some human sHSPs. In the HSPB5 oligomeric model (Jehle et al. 2011), this position does not have a proline and is in a helix modeled into the NTR. Prolines are relatively rigid compared to other residues due to being locked in one backbone phi-angle. Therefore prolines are often found capping helices in ordered regions of a protein. Disordered proteins and regions are often enriched in prolines as they do not promote ordered secondary structure but also limit the conformational space, possibly disfavoring aggregation. Mutation of the proline in HSPB1 likely changes the structural propensity of this region. Relative to WT HSPB1, P39L is a moderately better chaperone for the model client

protein used here- α-lactalbumin. In a previous study, P39L was found to have only WT-like activity with other client proteins, but the trend of P39L being a better chaperone compared to the other NTR disease mutants is consistent (Muranova et al. 2015). Of all of the NTR disease mutants, P39L oligomers elute earliest by SEC and have the most substantial increase in the average number of subunits- from 22 subunits in WT to 40. There is a large increase in polydispersity of oligomers observed by SEC for P39L. Also, P39L is the only mutation that alters subunit exchange. Specifically, the rate of exchange is approximately one-third the rate of WT. The only substantial change observed in local dynamics by HDXMS is the increased protection of the protected conformation of peptide 29-41. P39 is within this peptide, thus substitution of the rigid proline likely alters secondary structure propensities in this region. Given the predicted helical structure in this region (secondary structure prediction and modeling of HSPB5 oligomers), a proline to leucine mutation could increase helical propensity. It is likely that the more strongly protected feature in P39L is coupled with the slower subunit exchange. Specifically, a more rigid helix in the 29-41 region could slow the off-rate of subunit exchange in P39L oligomers. It is also possible that the slower exchange rate and/or change in secondary structure in the 29-41 region is responsible for the increased chaperone function.

The P39L/D3GXG mutant adopts neither D3GXG-like chaperone activity nor D3GXG quaternary structure. This mutant remains a similar chaperone to WT, therefore it has not lost its function. The oligomers formed in this context are very large and broad as seen by the SEC elution profile, more so in both aspects than P39L-alone. P39L/D3GXG also retains P39L-like secondary structure. Therefore, the structural feature attributable to the 230 nm peak in the D3GXG CD spectrum is not observable in this mutant. Unfortunately the 29-41 peptide was not observable by HDXMS, but a similar 29-38 peptide shows that two protected states at early time

points are retained (similar to P39L-only). However, the protected conformation observed in peptide 77-99 in WT and P39L oligomers is not observed in P39L/D3GXG. Therefore the D3GXG mutations impact local dynamics of P39L oligomers. Similar to G34R/D3GXG, P39L/D3GXG shows moderately decreased protection of peptides 100-109 and 150-163. Loss of protection in these parts of the ACD suggest a loss of transient interactions between the ACD and parts of the NTR. The increased protection observed in the 29-41 region for P39L and the slower exchange rate likely account for the ability of P39L/D3GXG to form large oligomers. Despite the mutations to abrogate NTR-NTR and ACD-CTR interactions, some type of interaction beyond the canonical ACD-ACD interaction is present in these oligomers.

E41K- The glutamate at position 34 of HSPB1 is conserved among both orthologs and paralogs as either a glutamate or aspartate. In secondary structure prediction of the NTR of HSPB1, this position is at the very start of a helix. In the HSPB5 oligomeric model (Jehle et al. 2011), this position (aspartate) is at the end of a helix modeled into the NTR. As a negative charge is conserved at this position, the E41K mutation could affect the protein by both loss of a negative charge and introduction of a positive charge. Relative to WT HSPB1, E41K is a slightly weaker chaperone for the model client protein used here- α-lactalbumin. This result is consistent with weaker activity observed for other client proteins (Muranova et al. 2015). E41K oligomers elute earlier than WT as seen by SEC, and they exhibit the smallest change in subunits of the disease mutants- increasing from 22 subunits in WT to 28. Subunits exchange at a similar rate in both WT and E41K oligomers. Overall, there are no substantial changes in local dynamics relative to WT oligomers as observed by HDXMS. However, the conserved 29-41 region of particular interest could not be analyzed due to lack of pepsin cleavage at this site.

The E41K/D3GXG mutant forms very large, polydisperse oligomers (more so than for E41K-alone), similar to P39L/D3GXG oligomers. However, the chaperone activity of E41K/D3GXG is intermediate between E41K and D3GXG. CD spectra of E41K/D3GXG reveal a secondary structure similar to that of GXG, which overall has weak ellipticity relative to all other mutants probed. Therefore secondary structure different from that of E41K or D3GXG alone is adopted. The simplest explanation for GXG-like secondary structure is that the E41K mutation "undoes" some of the effects of phosphorylation-mimicking mutations, leaving a net change in global interactions similar to GXG-only. By HDXMS however, there are no dramatic differences observed between E41K and E41K/D3GXG. There is an increased proportion of the protected state of the 1-10 peptide, similar to D3GXG. However, the inability to probe the 29-41 peptide means we cannot rule out a local structural change in this region that might be responsible for the dramatic changes observed by CD and SEC.

G84R- The glycine at position 84 of HSPB1 is conserved among orthologs and appears conserved among paralogs, but there is little homology in this region for a reliable alignment. The effects of the arginine mutation at this position could be attributed to both a loss in flexibility and the introduction of a large, charged side chain. This glycine is in the β2 strand of the crystal structure of HSPB1, and this region is predicted to form a β strand. In HSPB5 oligomeric models and structures of ACD-only constructs, this glycine is also in the region of the β2 strand, which is modeled to form in only certain conformations in the oligomeric model. Glycines provide a high level of flexibility, as there is no side chain present. Relative to WT HSPB1, G84R is a slightly poorer chaperone for the model client protein used here- α-lactalbumin. This result is consistent with poorer activity observed for other client proteins (Nefedova et al. 2013). The average size of G84R oligomers is larger than WT oligomers as

seen by earlier elution and an increase in average number of subunits from 22 to 34. However, at lower concentrations G84R oligomers appear to dissociate and form a long tail by SEC. This weaker global affinity of G84R oligomers is consistent with previous findings (Nefedova et al. 2013). Despite the lowered apparent affinity of G84R oligomers, the subunit exchange rate remains similar to WT. The G84R mutation shows essentially the same effects as G34R on local dynamics as seen by HDXMS. Specifically, peptides 29-41 and 77-99 appear to lose their protected conformations. As the G84R mutation is in the 77-99 region, a direct local effect is the more straightforward explanation for this structural change. The coupled loss of protected structure in both regions mirrors the coupled effects in G34R.

The G84R/D3GXG mutant adopts D3GXG-like chaperone activity, quaternary structure, and secondary structure. SEC elution profiles of this mutant and D3GXG overlay very well, suggesting similar conformations and resulting hydrodynamic radii are adopted. Similarly, the CD spectra of both mutants overlay very well, indicating that D3GXG-like secondary structure is preserved with the G84R mutation present. Overall, there were no major changes observed in local dynamics by HDXMS comparing G84R/D3GXG to D3GXG. In comparison to G84R-only, G84R/D3GXG has altered protection in the 1-10 peptide. Specifically, only one conformation is observed but it is more solvent exposed than the protected state observed in G84R-only. Additionally, the 33-41 peptide that is more protected in D3GXG compared to WT also has increased protection in G84R/D3GXG compared to G84R-only. Therefore, similar local structural changes are still present when G84R is introduced into the D3GXG context, suggesting this region is linked to the unusual CD feature observed at 230 nm. This mutant was also amenable to NMR as it remains dimeric at high concentrations required for NMR (~200μM). Several peaks are shifted in the G84R/D3GXG NMR spectrum relative to D3GXG. Of assigned

peaks, the most affected predominantly correspond to residues at the ACD-ACD dimer interface. These residues are also often affected in response to lowered pH or certain mutations in the ACD as discussed in Chapter 4. Also, oligomers at low pH or containing ACD disease-associated mutations on average form larger but lower affinity oligomers compared to WT. G84R is at the boundary between the NTR and ACD, where the β2 strand might form antiparallel to the β3 strand. If the β2 region were oriented away from the ACD as in the ACD-only NMR structure, it seems unlikely that dimer interface residues would be affected by mutation of this region. Coupled with the lack of assignments for β3 in the D3GXG context (whereas these are assigned in the ACD-only context), this result suggests that the β2 region is likely folding over the ACD, possibly forming a strand along β3 in the D3GXG context. However, the frequency of this structural feature occurring remains to be determined. It is possible that there is space for only a single β2 strand in a D3GXG dimer.

The correlation of several biophysical properties and chaperone activity of disease mutants of HSPB1 fit a general model of binding site availability as discussed in the previous section. Each of the NTR disease mutants shows some chaperone activity that varies between "normal" (only disease mutation present) and D3GXG contexts. In the D3GXG context, each mutant has equal or stronger chaperone activity relative to its non-D3GXG counterpart. Presumably, each D3GXG mutant has increased availability of its β4/β8 groove to clients, due to the IXI to GXG mutation in the CTR. G34R/D3GXG and G84R/D3GXG both form dimers or small oligomers, where presumably the NTR is more available for client interactions. Both of these mutants have similar activity to D3GXG, which is the most potent chaperone explored in this thesis. E41K/D3GXG has increased chaperone activity relative to E41K and decreased activity relative to D3GXG. P39L/D3GXG is still a stronger chaperone than WT but is the

weakest of the disease mutants in the D3GXG context. Similar to E41K/D3GXG, P39L/D3GXG likely has accessible β4/β8 grooves and less accessible NTRs. However, P39L/D3GXG retains its P39L-like secondary structure, suggesting that the local structural effects are very different between P39L and E41K mutations in the D3GXG context. Therefore, client-binding regions along the NTR could vary between these two mutants, despite close proximity of the mutations. The mechanism for effects on chaperone activity in the disease mutants alone (no D3GXG mutations) is less clear. The decreased activity of G34R, E41K, and G84R could be due to decreased accessibility of NTR or β4/β8 groove binding sites. The formation of larger oligomers by these mutants could suggest increased NTR-NTR interactions and therefore limited availability to clients, but this is a highly speculative relationship. In the case of P39L, the increased protection of the protected conformation seen in the 29-41 peptide likely lowers the subunit exchange rate as there are no other major changes in local dynamic observed. The altered structural propensity in this region could directly alter one or more NTR client-binding sites. However, more indirect effects on activity could be due to the altered subunit exchange rate for P39L.

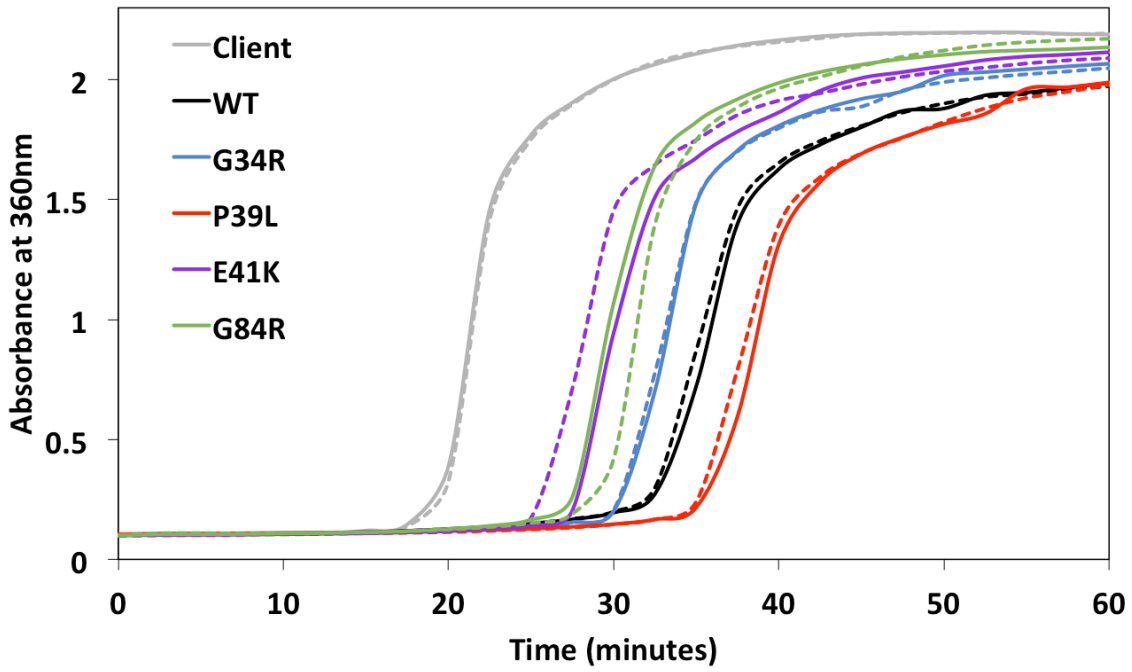


Figure 5.1 - Chaperone assay of NTR disease mutants and WT HSPB1. The chaperone assay was performed in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C. Aggregation of the model client protein α-lactalbumin (600μM) was triggered by reduction of disulfide bonds (addition of 20mM DTT) and subsequent destabilization. Absorbance was then measured as a function of time in the presence or absence of WT, G34R, P39L, E41K, or G84R forms of HSPB1 at 30μM (monomer concentration). Absorbance increases as the client protein aggregates due to light scattering. Duplicates in the same assay are shown as dashed lines. WT moderately delays formation of aggregates (prolonged lag phase). G34R, E41K, and G84R mutants do not delay aggregation as well as WT. P39L delays aggregation moderately more than WT, making it the strongest chaperone for this client.

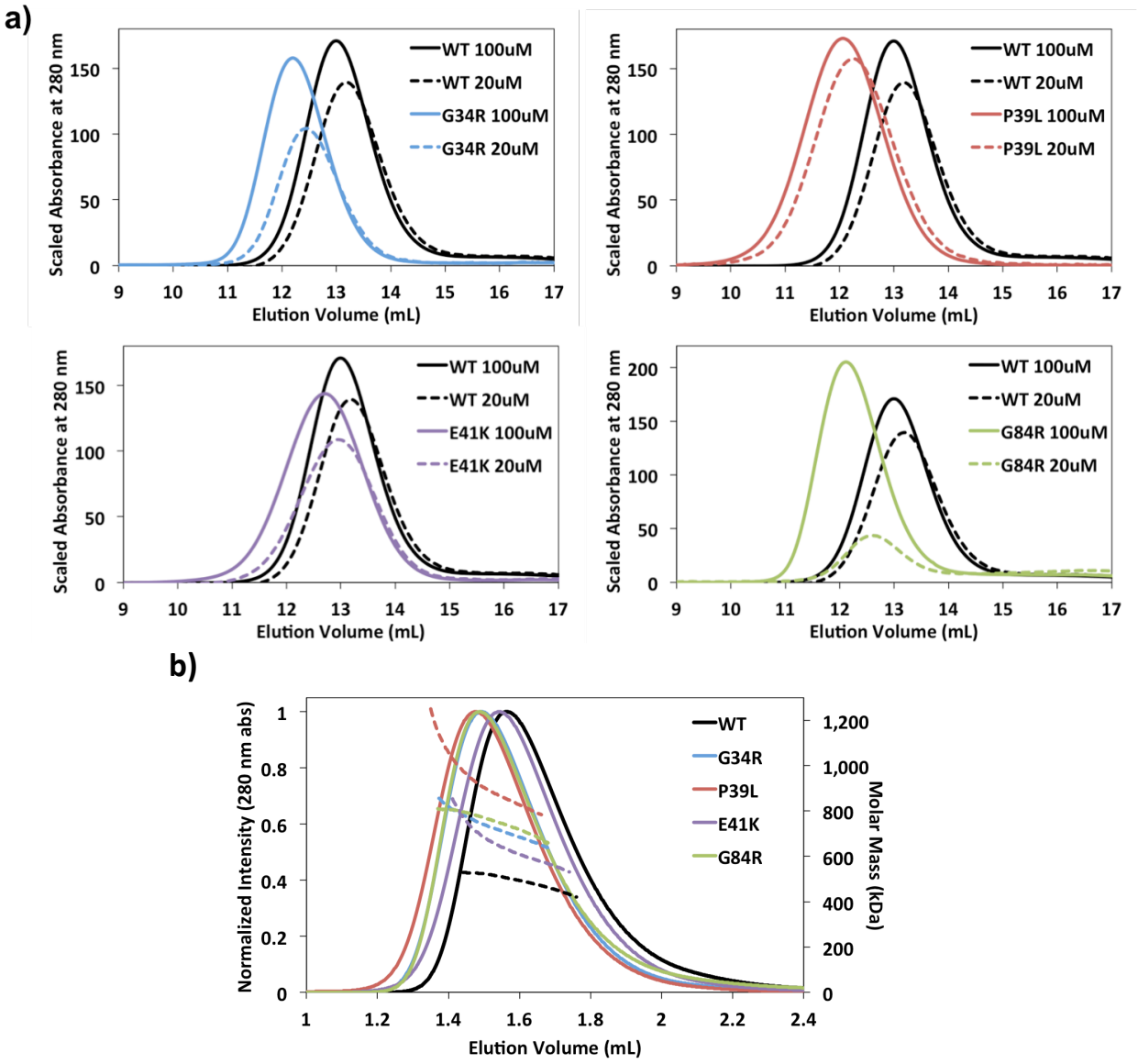


Figure 5.2 - SEC and SEC-MALS profiles of NTR disease mutants and WT HSPB1. (a) WT, G34R, P39L, E41K, and G84R samples were prepared at 100µM and 20µM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and incubated at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. WT traces (identical for Figure 3.1a) are shown for comparison. Each disease mutant elutes earlier than WT. G84R shows oligomer dissociation as seen by the tail after the main peak at 20µM. (b) Identical samples were prepared at 100µM. Samples were then injected onto a SEC column (2.4 mL Superose6) at room temperature. Eluted samples were analyzed by absorbance, multi-angle light scattering (MALS), and change in refractive index. Each disease mutant elutes slightly earlier than WT. Molar mass calculations based on the relationship between light scattering and concentration (from change in refractive index) are plotted in dotted lines across the top 50% of the 280 nm absorbance. All oligomers are polydisperse in that a broad range of masses are present. All disease mutants for larger (by mass) oligomers than WT.

Table 5.1- Molar masses and corresponding number of subunits calculated for WT and NTR disease mutant oligomers based on SEC-MALS. Average masses and mass ranges are reported over the top 50% of the 280 nm absorbance of each chromatogram.

	Average molar mass (kDa)	Estimated # of subunits	Molar mass range (kDa)	Estimated # of subunits
WT	488	22	420 - 530	18 - 24
G34R	730	32	638 - 856	28 - 38
P39L	927	40	783 - 1250	34 - 56
E41K	639	28	531 - 854	24 - 38
G84R	751	34	659 - 809	30 - 36

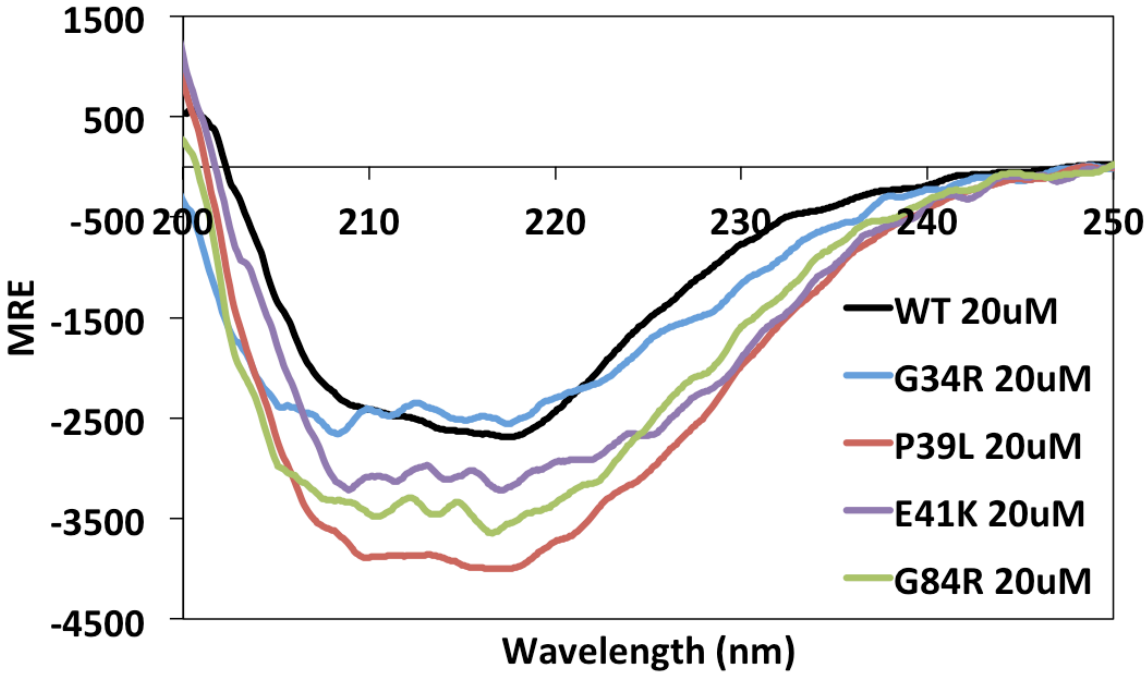


Figure 5.3 - Far-UV CD profiles of WT and NTR disease mutants of HSPB1. 20μM samples of WT and NTR disease mutants were prepared in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer were analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. Each disease mutant shows moderate to no (as for G34R) increases in overall ellipticity. P39L shows the strongest effect relative to WT.

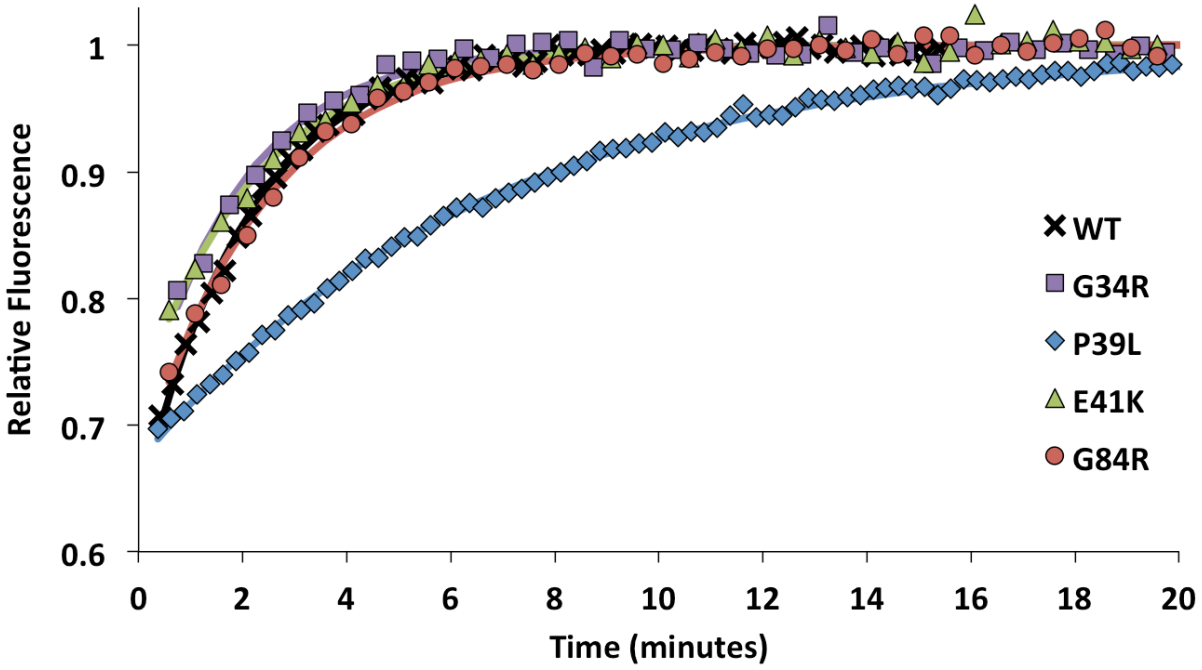


Figure 5.4 - Subunit exchange kinetics of NTR disease mutants and WT HSPB1 via homo-FRET. C137S/T174C mutations were made in WT, G34R, P39L, E41K, and G84R in order to attach AlexaFluor 488. For the representative curves here, 20μM samples were prepared in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and incubated at 37°C for one hour. For each mutant, fluorophore-labeled and unlabeled protein was mixed in a 1:2 ratio and the change in fluorescence was measured with time at 37°C. Fluorophore-labeled subunits are quenched in oligomers containing many other fluorophore-labeled subunits. Exchange out of labeled oligomers results in a release of homo-FRET (or quenching) and a subsequent increase in fluorescence. Data were fit to a single exponential rate function. Only the P39L mutant shows a substantial change in subunit exchange rate.

Table 5.2- Calculated subunit exchange rates for NTR disease mutants and WT HSPB1 at different concentrations and temperatures via homo-FRET.

		37°C			25°C		
		Rate (min^{-1})	Rate error (min^{-1})	R^2	Rate (min^{-1})	Rate error (min^{-1})	R^2
WT	40 uM	0.556	+/- 0.006	0.9985	0.0592	+/- 0.0006	0.9974
	20 uM	0.475	+/- 0.008	0.9963	0.0535	+/- 0.0006	0.9969
	5 uM	0.487	+/- 0.012	0.9915			
G34R	40 uM	0.561	+/- 0.020	0.9856			
	20 uM	0.520	+/- 0.033	0.9584			
	5 uM	0.512	+/- 0.039	0.9316			
P39L	40 uM	0.165	+/- 0.002	0.9978	0.0156	+/- 0.0001	0.9981
	20 uM	0.149	+/- 0.001	0.9983	0.0146	+/- 0.0001	0.9978
	5 uM	0.148	+/- 0.002	0.9952	0.0117	+/- 0.0002	0.9944
E41K	35 uM	0.457	+/- 0.018	0.9854			
	20 uM	0.446	+/- 0.019	0.9804			
	5 uM	0.329	+/- 0.018	0.9483			
G84R	40 uM	0.414	+/- 0.013	0.9807			
	20 uM	0.421	+/- 0.015	0.9834			
	5 uM	0.367	+/- 0.026	0.9200			

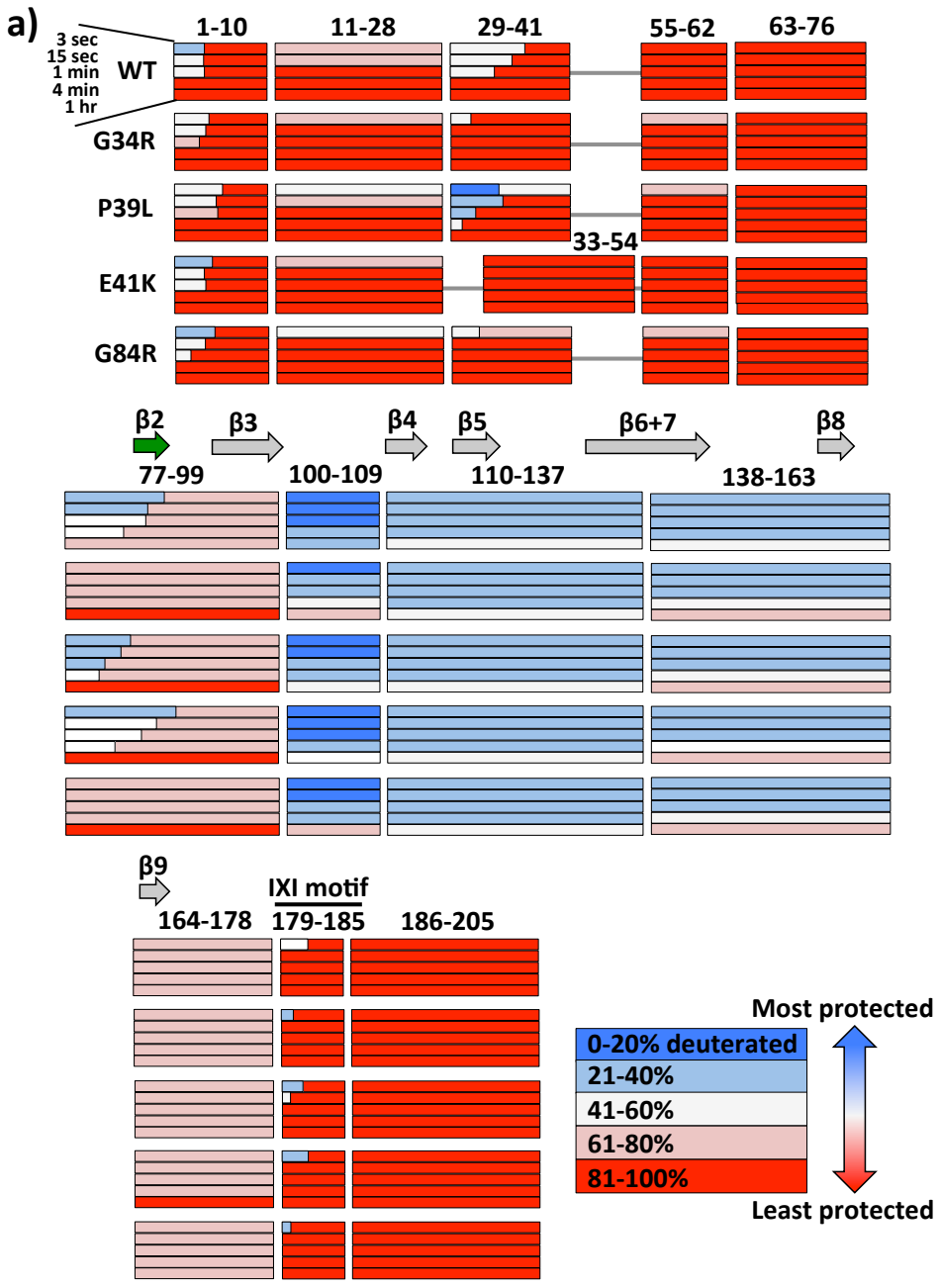
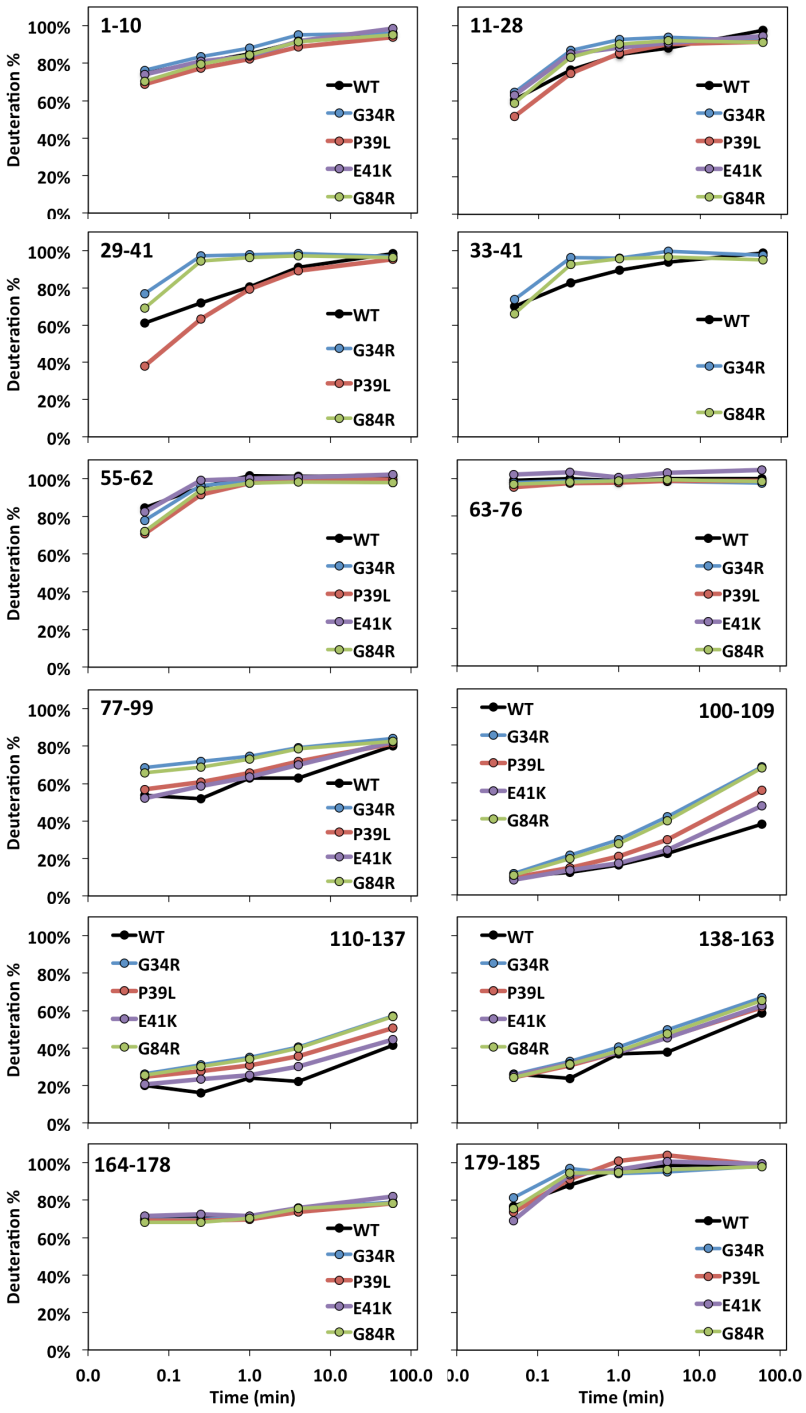


Figure 5.5 - Comparison of solvent protected regions in NTR disease mutants WT HSPB1 using HDXMS. WT/C137S, G34R/C137S, P39L/C137S, E41K/C137S, and G84R/C137S samples were prepared at 200μM in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and equilibrated at room temperature for 3 hours. Exchange experiments were performed at room temperature (~25°C) by diluting the proteins 10X into an identical D_2O based buffer. Peptides analyzed are shown as a series of horizontal bars (e.g. peptides consisting of 1-10, 11-28, etc.). (a) Time points are represented as bars stacked vertically for one peptide. The color-coding represents deuteration level of each peptide (blue- most protected, red- least protected). For certain peptides, the horizontal bar is split into two sub-bars, which represent two distinct

populations (conformations) observed in the mass spectrum for that peptide. The relative lengths of the sub-bars within one peptide bar correspond to relative proportions of the two conformations. (b) The data is alternately presented as quantitative line plots for each peptide, except for the addition of the 33-41 peptide and omission of the 186-205 peptide (completely unprotected in all cases). For peptides that show multiple conformations, the weighted average of deuteration is presented.



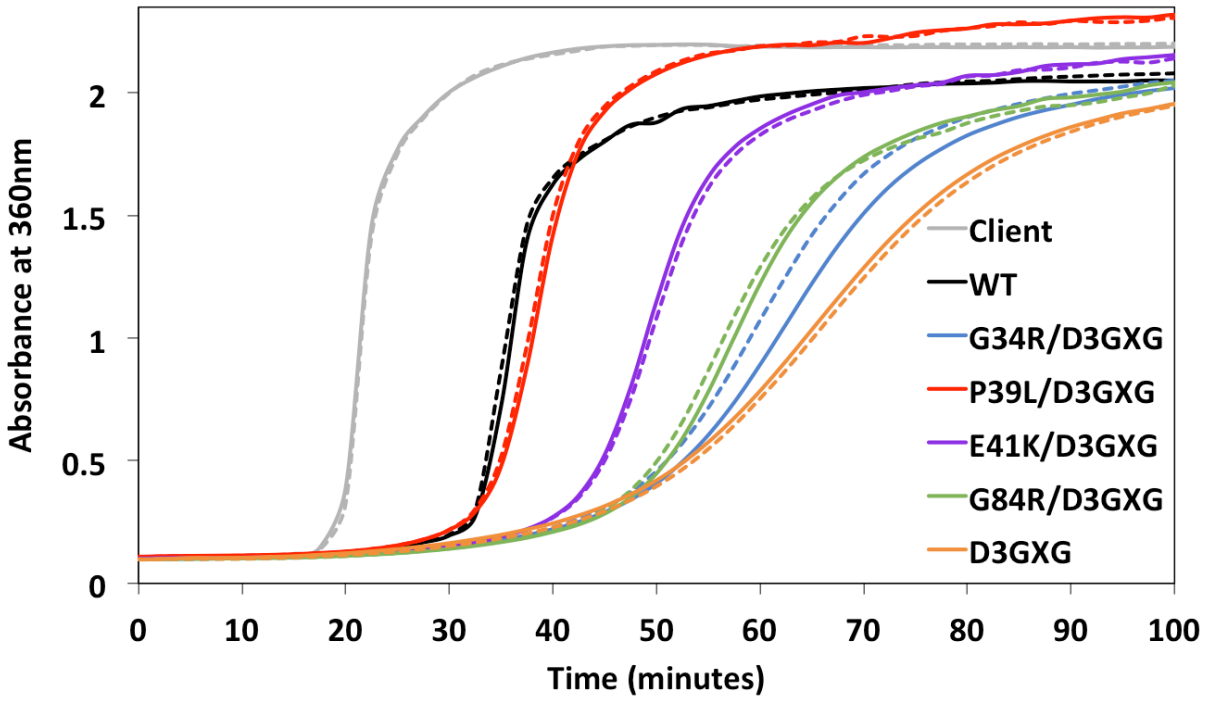


Figure 5.6- Chaperone assay of NTR disease mutants in the D3GXG context of HSPB1. The chaperone assay was performed in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C. Aggregation of the model client protein α-lactalbumin (600μM) was triggered by reduction of disulfide bonds (addition of 20mM DTT) and subsequent destabilization. Absorbance was then measured as a function of time in the presence or absence of WT, D3GXG, G34R/D3GXG, P39L/D3GXG, E41K/D3GXG, or G84R/D3GXG forms of HSPB1 at 30μM (monomer concentration). Absorbance increases as the client protein aggregates due to light scattering. Duplicates in the same assay are shown as dashed lines. WT moderately delays formation of aggregates (prolonged lag phase). D3GXG very strongly delays aggregation. The disease mutants in D3GXG delay aggregation in an intermediate range between WT and D3GXG.

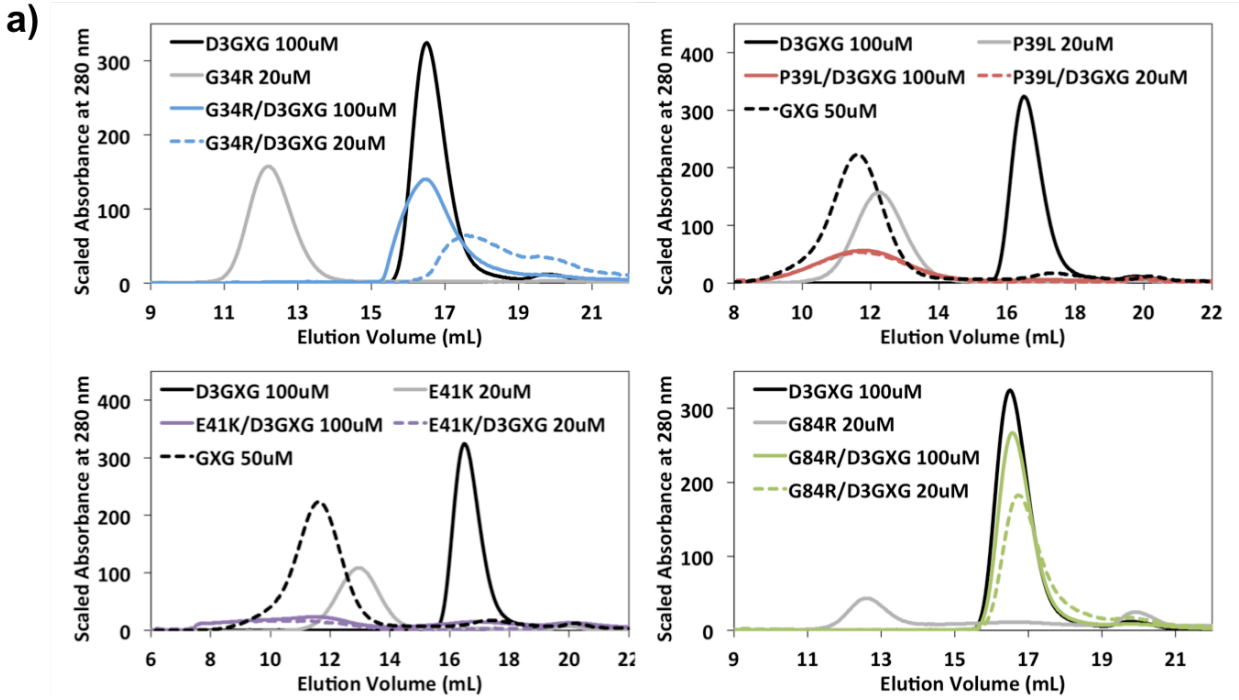
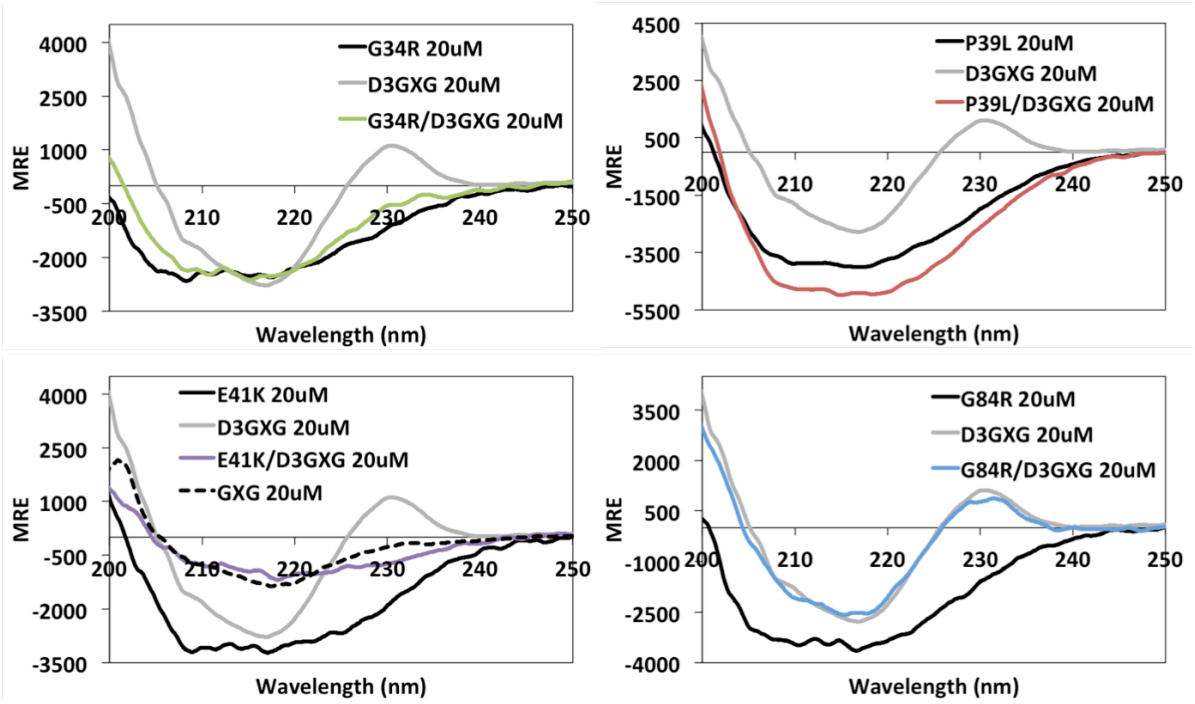


Figure 5.7- SEC and CD profiles of NTR disease mutants in the D3GXG context of HSPB1. (a) G34R/D3GXG, P39L/D3GXG, E41K/D3GXG, and G84R/D3GXG mutants were incubated at 100μM and 20μM (monomer concentration) in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer at 37°C for one hour. Samples were then injected onto a SEC column (24 mL Superose6) at room temperature. G34R/D3GXG and G84R/D3GXG mutants elute late, similar to D3GXG. P39L/D3GXG and E41K/D3GXG mutants elute very early, more consistent with the GXG elution profile. (b) 20μM samples of disease mutants in D3GXG were prepared in 25mM NaPi/50mM NaCl/0.25mM EDTA/pH7.5 buffer and analyzed by far-UV CD spectroscopy at 20°C. Samples were incubated at room temperature for 3 hours prior to measurement. G34R/D3GXG adopts G34R-like secondary structure, despite forming small oligomers. P39L/DGXG adopts P39L-like secondary structure. E41K/D3GXG adopts more GXG-like structure as opposed to E41K or D3GXG. G84R/D3GXG behaves similarly to D3GXG and has the positive 230 nm peak.

b)


G34R 20uM
D3GXG 20uM
G34R/D3GXG 20uM
MRE
Wavelength (nm)
P39L 20uM
D3GXG 20uM
P39L/D3GXG 20uM
MRE
Wavelength (nm)
E41K 20uM
D3GXG 20uM
E41K/D3GXG 20uM
GXG 20uM
MRE
Wavelength (nm)
G84R 20uM
D3GXG 20uM
G84R/D3GXG 20uM
MRE
Wavelength (nm)

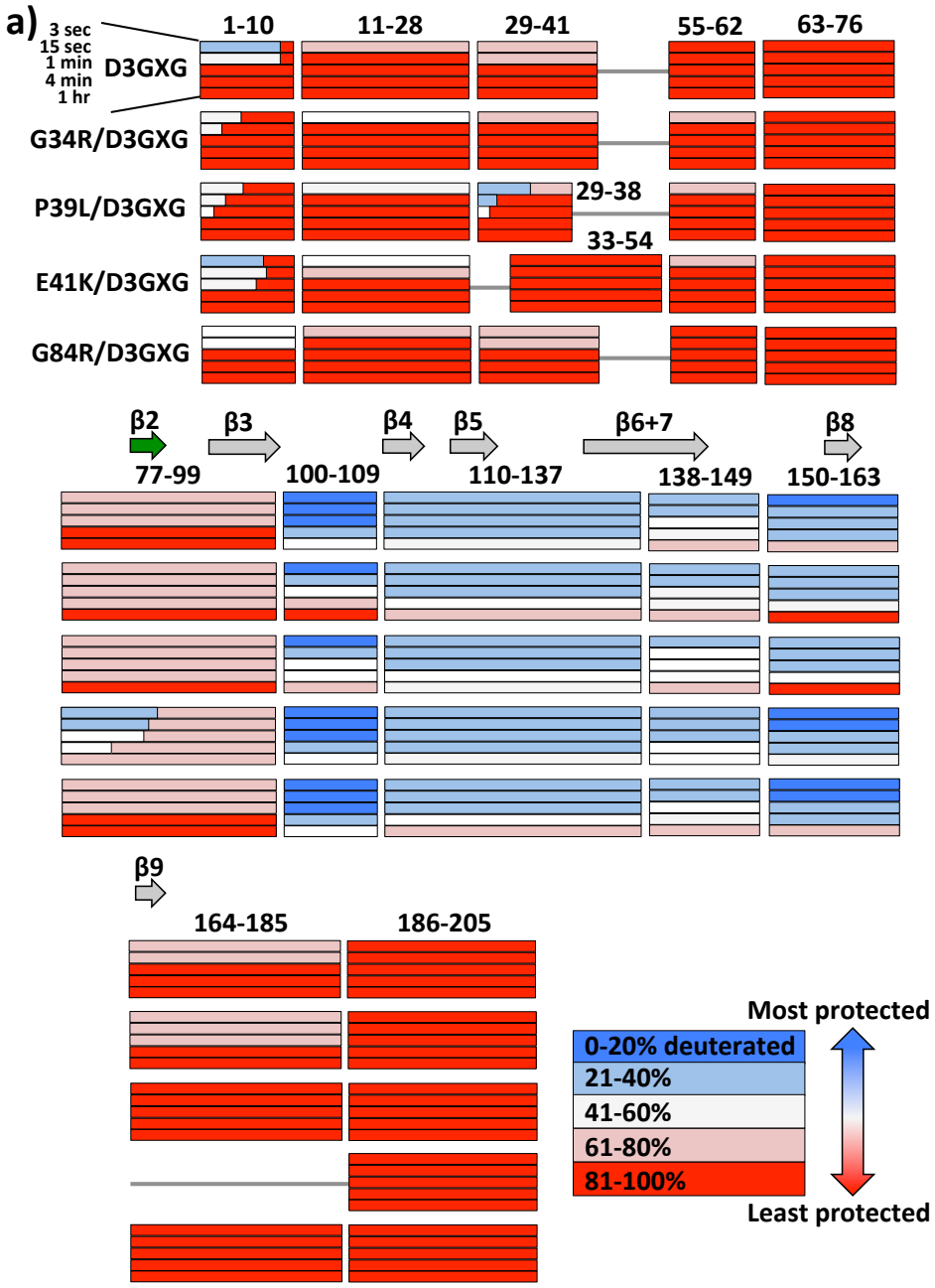
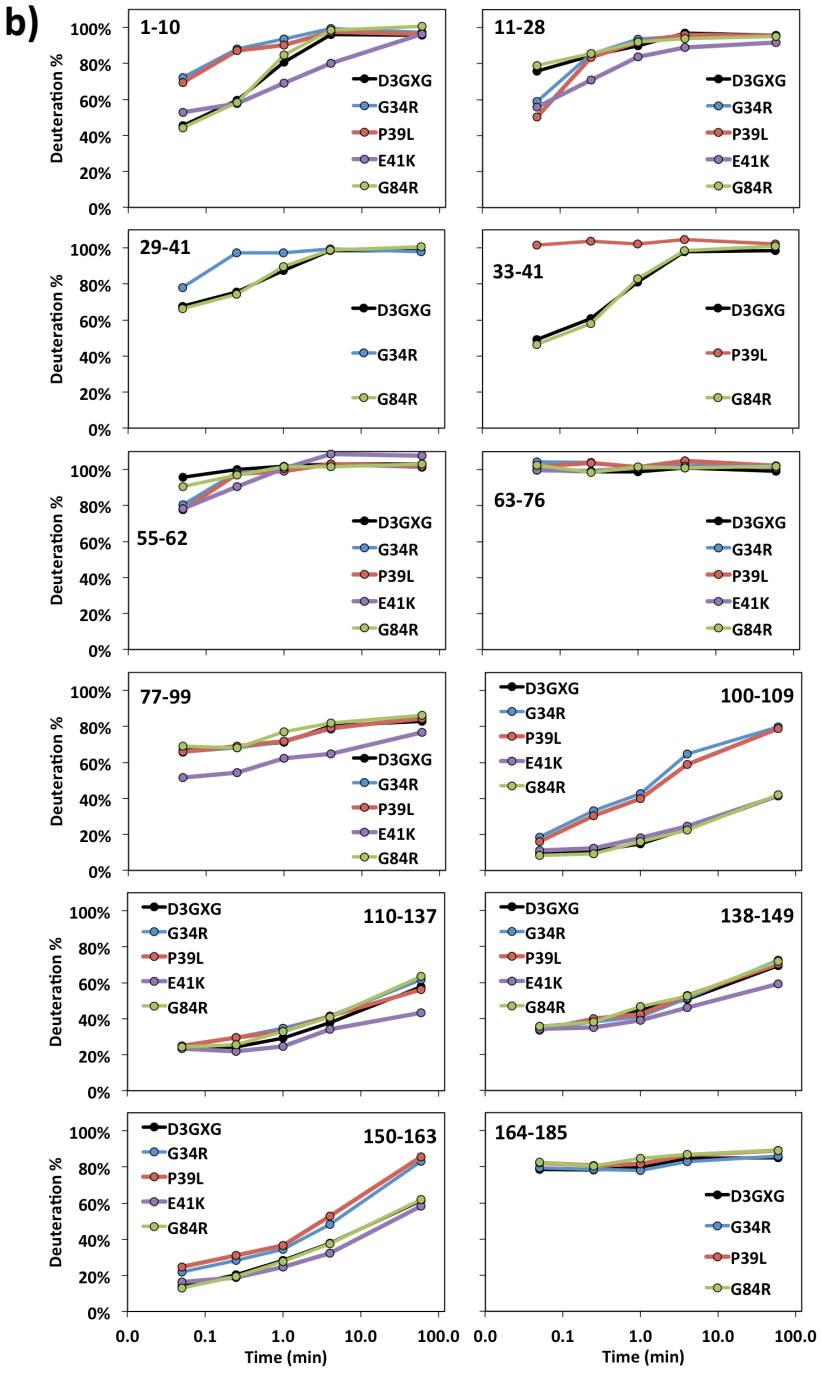


Figure 5.8- Comparison of solvent protected regions in NTR disease mutants in the D3GXG context of HSPB1 using HDXMS. D3GXG/C137S, G34R/D3GXG/C137S, P39L/D3GXG/C137S, E41K/D3GXG/C137S, and G84R/D3GXG/C137S, samples were prepared at 200μM in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer and equilibrated at room temperature for 3 hours. Exchange experiments were performed at room temperature (~25°C) by diluting the proteins 10X into an identical but D_2O based buffer. (a) Peptides analyzed are shown as a series of horizontal bars (e.g. peptides consisting of 1-10, 12-28, etc.). Time points are represented as bars stacked vertically for one peptide. The color-coding represents deuteration level of each peptide (blue- most protected, red- least protected). For

certain peptides, the horizontal bar is split into two sub-bars, which represent two distinct populations (conformations) observed in the mass spectrum for that peptide. The relative lengths of the sub-bars within one peptide bar correspond to relative proportions of the two conformations. (b) The data is alternately presented as quantitative line plots for each peptide, except for the addition of the 33-41 peptide and omission of the 186-205 peptide (completely unprotected in all cases). For peptides that show multiple conformations, the weighted average of deuteration is presented.



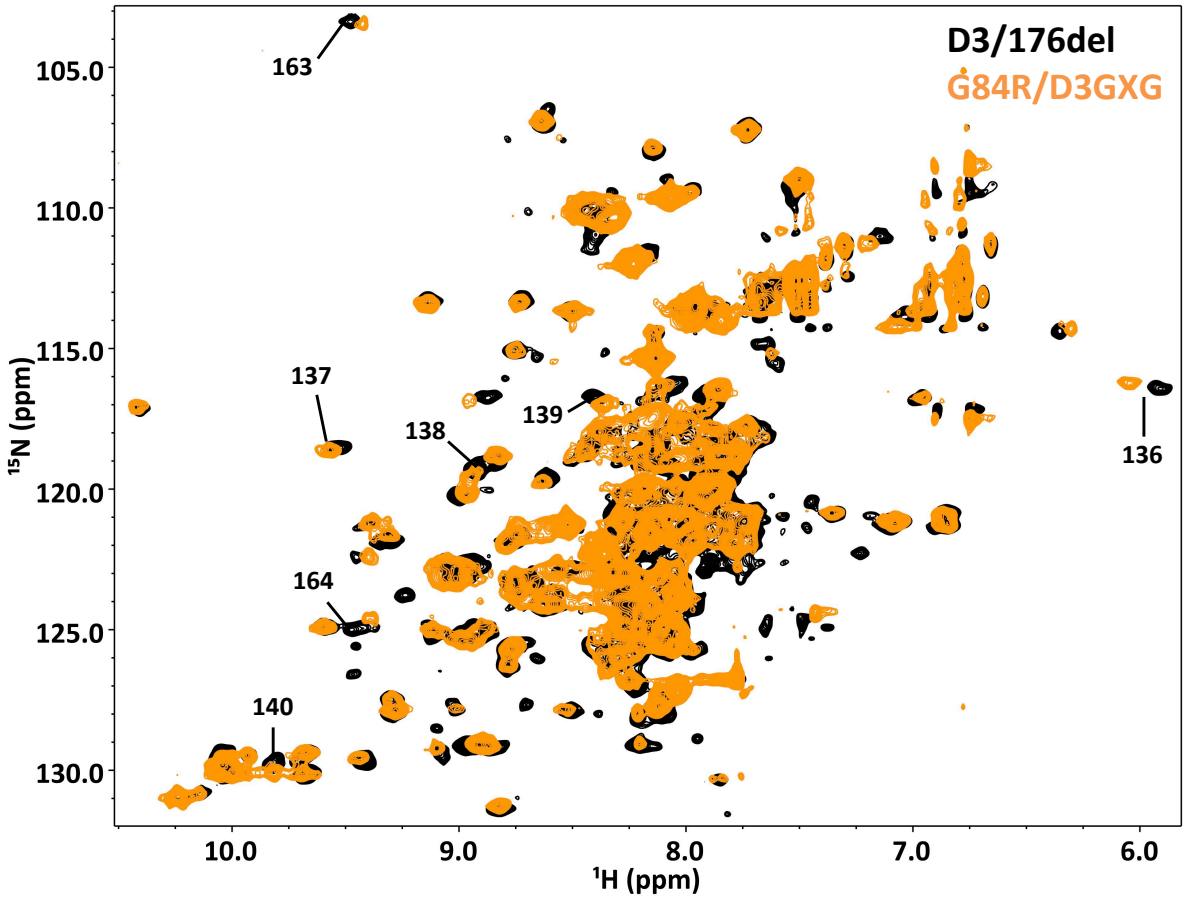


Figure 5.9- Comparison of G84R/D3GXG and D3/176del HSPB1 dimers by ^{1}H-^{15}N HSQC-TROSY spectra. G84R/D3GXG and D3/176del undeuterated samples were prepared in 50mM NaPi/100mM NaCl/0.5mM EDTA/pH7.5 buffer, and spectra were collected at 30°C. While many intensity changes might be due to the differences between D3GXG and D3/176del constructs, many peaks shift in the disease mutant that do not normally shift between the constructs. These peaks are noted and mostly correspond to residues at the ACD-ACD dimer interface.

Table 5.3- Summary of functional, structural, and dynamics changes in NTR disease-associated mutants of HSPB1.

	Property	G34R	P39L	E41K	G84R
	Chaperone activity	Weaker	Better	Weaker	Weaker
	Oligomerization	Larger	Largest	Larger	Larger lower "stability"
	Subunit exchange	Similar	Slower	Similar	Similar
	Local dynamics	29-41 and 77-99 lose secondary protected state	29-41 protected state is more protected	Similar, but 29-41 region is unknown	29-41 and 77-99 lose secondary protected state
In D3GXG context	**Chaperone activity (vs. D3GXG)**	Similar	Weakest	Weaker	Similar
	Oligomerization	Small oligomers	Very large, polydisperse	Very large, polydisperse	Dimers
	Secondary structure	G34R-like	P39L-like	GXG-like	D3GXG-like
	Local dynamics (vs. D3GXG)	Smaller 1-10 protected state, 100-109 and 150-163 less protected	Smaller 1-10 protected state, 29-38 more protected, 100-109 and 150-163 less protected	77-99 retains protected secondary state as seen in E41K-only	Similar

6. CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Summary of novel findings on HSPB1 structure

Historically, many studies of sHSPs have focused on truncated forms of the protein and/or explore primarily one type of intermolecular interaction. These studies are essential "bottom-up" approaches to piece together a working understanding of a highly complex structural problem. With this foundation and technological improvements, newer studies, including the work in this thesis, aim to characterize the myriad states and interactions that occur simultaneously in a sHSP system.

This work demonstrates correlative structural and functional effects of perturbations to each type of known HSPB1 inter-protomer interaction (NTR-NTR, ACD-ACD, and ACD-CTR) and the combination of these perturbations. In general, abrogation of one type of interaction does not necessarily lead to oligomer dissociation. There is intricate interplay among these types of interactions that modulate oligomer size and function. Although some NTR-ACD interactions are proposed in oligomeric models, this work first establishes contacts between particular regions of the NTR with an ACD in the dimeric D3GXG form of HSPB1. The newly defined interaction is likely altered in stress-states (phosphorylation) of the protein and is implicated in effects of NTR disease-associated mutations.

The HDXMS work presented in this thesis provides the first glimpse of local dynamics and structure in specific regions of the NTR in a polydisperse, oligomeric sHSP. Previous studies have utilized sparse restraints from various techniques to develop models of the NTR in oligomers, but definitive properties of each segment (here 10-20 residue peptide) had not previously been described. Multiple regions were identified in the NTR that have moderate levels of solvent protection, presumably due to formation of local structural features (helix or

hairpin) or interactions with other regions of the protein, either intra- or inter-protomer. In WT oligomers, three of these regions had two distinct states with different levels of solvent protection. The findings of conformational heterogeneity are consistent with general observations of heterogeneity discussed in previous oligomeric models (Jehle et al. 2011). Additionally, it was possible to directly compare solvent protection of regions of the NTR between WT oligomers and D3GXG dimers, which represent a stressed (phosphorylation-mimicking), "activated" state of HSPB1. Two of the three heterogeneous regions observed in the NTR of WT oligomers become homogeneous and less protected in the dimeric form. Therefore, these regions are implicated in either direct inter-protomer interactions in an oligomer or indirectly undergo structural transitions upon dissociation. There is also evidence for one region of the NTR (peptide 33-41) becoming more solvent protected in the dimeric state relative to the oligomeric context. Adoption of different structural features among different oligomeric states of a sHSP likely alters hydrophobic surfaces recognized by client proteins.

The only full-length human sHSP dimer structure to be published is the recent crystal structure of HSPB6 in complex with a binding partner, but disordered regions are not resolved (Sluchanko et al. 2017). Using HDXMS, NMR assignments, and PRE probes to characterize the D3GXG, we have developed a working structural model of full-length HSPB1 dimers. Multiple possible orientations of the NTR relative to the ACD are possible but limited by PRE restraints, such as the interaction of position 2 of the NTR with the edge of the ACD. One of the most interesting findings was that part of the NTR (approximately residues 60-70) is in two distinct states, yet both states are highly disordered. Two PRE probes show that this region (at least one state of it) is in close proximity to the ACD. The numerous prolines in the NTR of HSPB1 could both limit typical ordered structure and also restrain disordered regions into more discrete states.

From a functional perspective, the disordered, hydrophobic NTR must be conformationally restrained so as to avoid promoting aggregation. Therefore, the heterogeneous, disordered, but "controlled" structure of the extreme phosphorylation-mimicking dimer demonstrates how a sHSP might have evolved to present numerous, transient binding sites for client proteins.

Most studies of disease-associated mutants of sHSPs have focused on mutations in the ACD, partly because it is possible to examine in greater detail truncated ACD-only forms of sHSPs. Many of these studies, and work on NTR disease mutants (Muranova et al. 2015, Nefedova et al. 2013b), rely on making correlations among only global results. It is technically challenging to identify key structural changes in sHSP disease mutants that help explain their mechanisms for disease. The use of HDXMS in this thesis to examine local dynamic/structural changes in several NTR disease mutants of HSPB1 provides one of the few structural, mechanistic insights into disease-altered states of full-length human sHSPs. The additional use of the D3GXG construct as a background for these disease mutants helped elucidate distinctly different mechanisms of structural changes among the NTR disease mutants. The most interesting result was that certain disease mutants in the D3GXG context formed large oligomers. Thus certain mutations promote local structures in the NTR that favor oligomerization. These varied results among mutants are surprising considering the close proximity in sequence of several NTR disease mutants, and highlights the multiple features and roles of the NTR.

6.2 Future directions

There are many tempting experiments that are ideal next steps for studying HSPB1 structure and interaction with client proteins and are discussed in this section. The next two sections discuss broader applications of this thesis work.

To develop a more detailed structural model of states of the D3GXG dimer, additional spin label experiments and possibly SAXS analysis would be essential. Several sites along the NTR were explored for spin label attachment and PRE experiments. However, many more sites, particularly closer to the conserved region containing sites of disease mutations, should be explored to identify potential contacts between the NTR and the β4/β8 groove. The PRE studies performed here probed both intra- and inter-protomer interactions with no distinction, as all protomers were ^{15}N-labeled and spin-labeled. To probe strictly inter-protomer interactions (NTR with the neighboring ACD), one could measure PRE effects of a mixture of ^{15}N-labeled protein mixed with unlabeled protein that has a spin probe attached. Position 91 examined here is likely conformationally limited and most contacts observed are probably intra-protomer. However, for positions 2, 65, and 74, inter- and intra-protomer interactions are equally possible with the available data. To obtain distance restraints between two positions in the dimer, one could perform DEER-EPR (Double Electron Electron Resonance-Electron Paramagnetic Resonance). This method often utilizes the same spin probe used in this study (MTSL), therefore the necessary constructs with the appropriate probe have been characterized. DEER-EPR provides distance measurements between two spin probes in a protein that is flash-frozen. With a sHSP frozen in multiple states (no exchange to complicate the experiment), it is possible to identify multiple distance restraints due to conformational heterogeneity. SAXS of the dimeric form of HSPB1 could provide an envelope in which to model restraints from NMR chemical shifts and PRE effects. For regions of proteins that are both disordered and extended, SAXS generally averages out the states and will provide no density for those regions. Although there are disordered regions in the NTR of HSPB1, multiple contact points between the NTR and ACD suggest that disordered regions are conformationally limited and not extremely flexible (only

"locally" flexible). Therefore, a comparison of SAXS envelopes for ACD-only and D3GXG might identify density additional to that of the ACD-only in which the NTR can be modeled. Given the structural heterogeneity observed for D3GXG, there would likely be multiple SAXS envelopes generated to accommodate multiple NTR orientations.

To better understand the structural roles of regions of the NTR, further peptide studies could be carried out systematically. The analysis of the NTR peptide proposed to form a hairpin showed formation of non-random coil secondary structure. These results and the recent findings on regions of the NTR interacting with the ACD in the context of HSPB6 (Sluchanko et al. 2017) suggest that there is value in studying isolated pieces of the NTR. For example, peptide 1-10 by HDXMS showed two distinct conformations, but it is unknown whether this is due to a local helix/hairpin formation or interaction with the ACD. Probing secondary structure formation of this region as an isolated peptide would likely answer this question. However, as local structural features and NTR-ACD interactions appear to be context specific, isolated peptide studies should only be used to corroborate findings in a larger context.

To better probe local structural changes in sHSP disease mutants, other known disease-associated mutations could be placed in the D3GXG context. Specifically, particular disease mutants in the ACD discussed in Chapter 4 might alter the NTR-ACD interactions observed in D3GXG. Identifying local changes requires the resulting construct to be NMR-amenable (small), which was not the case for three of the four NTR disease mutations in D3GXG characterized in this thesis. However, if an ACD disease mutation in the D3GXG background results in an oligomeric protein, it is highly likely that NTR-ACD interactions are perturbed to promote an oligomer-favoring state.

The tools and constructs developed in this thesis have already proven useful to current studies by other researchers. Another graduate student has used WT, D3, and D3GXG forms of HSPB1 to probe changes in interactions with a known disordered client protein, tau. The results from the perspective of the client protein have been published (Baughman et al. 2018). Identification of tau binding sites on D3GXG HSPB1 by NMR is in progress, a task that would be intractable without a characterized, dimeric form of HSPB1. Although beyond the scope of this graduate student's work, applying HDXMS to HSPB1 interacting with client proteins could identify structural changes that occur as HSPB1 is performing its function. Transient hydrophobic interactions with clients would not likely be detected by HDXMS, as was the case for non-vertebrate sHSPs (Basha et al. 2013), but indirect changes in structural propensities could be measured. As certain stress-mimicking or disease-associated states of human sHSPs have been shown to interact more tightly with clients (Rajagopal et al. 2015b), more distinct changes might be observed by HDXMS for non-WT states.

6.3 Functional questions in the sHSP field

The heterogeneous properties of sHSPs make them both technically challenging to characterize and also difficult to define by canonical standards for protein function. Conventional protein structure-function studies often rely on one or a few well-ordered structures of a protein and a single, tight, quantifiable interaction with a binding partner. sHSPs have a mixture of ordered and disordered domains that adopt multiple local conformations with higher order effects on oligomer polydispersity. As subunits exchange rapidly among oligomers, it is not possible to capture a homogeneous sample of oligomers. The general role of sHSPs is to interact with a variety of destabilized, aggregation-prone proteins in the cell. Therefore, there must be multiple binding sites on a sHSP to accommodate diverse client proteins. Also, sHSPs

must interact only transiently with client proteins so that clients have frequent chances to refold spontaneously or interact with other types of chaperones. The challenges we face in characterizing sHSP function require a paradigm shift in how we think about heterogeneous protein systems.

The most sought after question in the sHSP field is to determine client binding sites on a sHSP. Only after numerous reports on transient interactions between client proteins and multiple regions of sHSPs has it been generally accepted that multiple binding sites exist. However, formation of polydisperse oligomers greatly complicates clear definition of consistent binding sites in a sample for study. The full-length HSPB1 D3GXG dimer developed in this thesis offers the chance to probe client interactions with a simplified state of a sHSP. However, heterogeneous local structures and disordered regions of D3GXG still pose a challenge to identifying interaction sites. It is possible that NTR-client interactions are so heterogeneous in location and orientation and so transient that for certain sHSP-client pairs no distinct interaction sites can truly be defined. Instead, comparing general preferences of different clients for different binding regions in a sHSP might yield more information, albeit indirectly, on the function of sHSPs.

The second most enticing question is identifying the "functionally active state(s)" of sHSPs. Although there is a third question regarding changes in binding sites in stressed or disease-mutated states of sHSPs, I believe pursuit of answers to both of these questions is the same. This question is particular to oligomeric sHSPs where a broad distribution of oligomeric states coexists. However, it is not even clear what subunits (monomers, dimers, small oligomers) are rapidly exchanging among oligomers. Without defining the structural states present in a sample, it is challenging to identify which states are more or less functional. By controlling the

oligomeric distribution of the system (e.g. phosphorylation-mimicking mutants) and correlating oligomeric size with changes in function, it is possible to make inferences about the activity of on average smaller or larger oligomers. In general for HSPB1 and HSPB5, phosphorylation-mimicking mutants form smaller oligomers and have increased activity measured both *in vitro* and *in vivo*. Although these are the most well-characterized and correlated perturbations, it is not clear if increased function in phosphorylated states is due to exposure of binding sites upon oligomer dissociation, alteration of binding sites due to local structural changes, or the more likely combination of these effects. This thesis work demonstrates distinct conformational changes in the NTR between dimeric and oligomeric states of HSPB1 that could lead to different binding surfaces for clients. For the model client protein used here (α-lactalbumin), the D3 HSPB1 mutant was not a more effective chaperone than WT, while D3GXG was hyperactive. D3 forms small oligomers and adopts similar secondary structure to D3GXG based on CD spectroscopy. The only presumable difference between the two states is that the β4/β8 groove is occupied in D3 and not in D3GXG, suggesting that α-lactalbumin specifically requires both NTR and β4/β8 groove binding sites on HSPB1 in order to interact in a manner that delays aggregation. However, D3 and intermediate phosphorylation mimics are potent chaperones relative to WT for several other model client proteins reported in numerous studies (Hayes et al. 2009, Jovcevsky et al. 2015, Shashidharamurthy et al. 2005). It therefore seems that different clients require exposure of different binding sites on a sHSP. The example of phosphorylated sHSPs highlights the need to characterize interactions with multiple client proteins- there are different preferences for sHSP binding sites among clients, and without this knowledge one might falsely categorize the activity of different states of the sHSP.

6.4 A technical roadmap for characterizing sHSPs

This thesis provides a roadmap for the effective structural characterization of sHSPs. Specifically, presented here are types of mutations and experiments that can aid in studying all regions of a sHSP in various oligomeric states.

The mutations explored here to create a stable dimer of HSPB1 (D3GXG) can be applied to other oligomeric sHSPs. As many sHSPs have phosphorylation sites in their NTR, phosphorylation-mimicking mutations (to Asp or Glu) can be used in these cases to promote dissociation of oligomers by perturbing NTR-NTR interactions. Many sHSPs also have IXI motifs in their CTRs and/or NTRs. Mutation of these IXI motifs to "GXG" can abrogate interaction of this binding motif with the β4/β8 groove of the ACD. While many studies previously have utilized one of these types of mutation, the combinatorial effects can be pursued as was done for HSPB1. Generation of a smaller, more monodisperse form of a sHSP can aid in structural characterization via almost any method (except for electron microscopy, which requires larger proteins/complexes). Also, a smaller, monodisperse form of a sHSP would aid in more quantitative analyses of interactions with client proteins. It is much less challenging to probe affinities and stoichiometries of sHSP-client interactions if there are not also homo-interactions that complicate the experiment.

The most substantial and applicable finding here is the value of hydrogen-deuterium exchange mass spectrometry (HDXMS) for studying sHSPs. This method has been applied to oligomeric, monodisperse plant sHSPs (Basha et al. 2013), but it has not been expanded upon for other sHSPs. HDXMS of HSPB1 provided information on local dynamics of 10-20 residue long peptides throughout the protein. As the experiment is not limited by protein size, it was possible to rigorously compare local dynamics and inferred structure between the D3GXG dimer and WT

oligomer forms. This approach draws the focus of structure elucidation away from only well-ordered regions and towards a more complete depiction of a protein, including ordered and disordered regions. Also, it is possible to detect different protected states of a region and therefore make inferences about multiple conformations in a protein. This method could be applied to any other sHSP, in any oligomeric or disease-mutated state, under almost any conditions, and with any potential client protein. With a small amount of D_2O, anyone can prepare HDXMS samples. Although a pepsin column was used here to limit overlapping pepsin peptides in the spectra, commonly available pepsin can also be used to digest the protein immediately after quenching the reaction. Thanks to software advancements in predicting peptide sequences from tandem MS data, the initial peptide mapping of a protein is far less laborious. As MS is a widespread technique used in almost any chemistry/biochemistry research setting, most institutions likely have a liquid chromatography/mass spectrometer setup with the required sensitivity and resolution for HDXMS. However, in order to limit back exchange (deuterons back to protons), it is recommended that the chromatography part of the setup be chilled to 1-4°C. There are multiple programs freely available for analyzing HDXMS data (used here- Guttman et al. 2013). Although the experiments and analysis are time and labor intensive, they do not require the researcher to be an expert in mass spectrometry.

This work also presents the first NMR analysis of a full-length sHSP dimer in solution. A few solution NMR studies have probed either the truncated ACD or large oligomers, and in the latter case only the flexible CTR is generally observable. Solid-state NMR can provide structural information on sHSP oligomers as well, but this is an even more specialized approach than standard solution NMR. The experiments presented here on the D3GXG dimer of HSPB1 required a large amount of D_2O and ample time on high field magnets with cryo-probes, which

can be cost prohibitive. Also, expertise in NMR (self or from a mentor) is necessary to carry out the experiments and interpret the more complicated results associated with sHSPs. However, there are numerous academic research groups that have these resources and expertise. As described above, D3GXG-like mutants of other oligomeric sHSPs could be amenable to these same NMR experiments if the protein remains in a small form at high concentrations. Also, several sHSPs natively exist as monomers or dimers, as is the case for human HSPB6 and HSPB8. Assignment of the NTR of sHSPs not only aids in structural understanding but also provides a method for determining binding sites of client proteins at residue-level resolution. Even if much of the disordered NTR is unassignable, as is the case for D3GXG HSPB1, the use of spin labels and the PRE effect throughout the NTR allows for determination of approximate distances between specific regions of the NTR and ACD. Therefore, if the resources and expertise are available, solution NMR of full-length sHSPs in dimeric/monomeric forms is a worthwhile pursuit.

REFERENCES

Alderson TR, Benesch JLP, and Baldwin AJ (2017) Proline isomerization in the C-terminal region of HSP27. Cell Stress Chap. 4:639-51

Aqualina JA, Shrestha S, Morris AM, and Ecroyd H (2013) Structural and functional aspects of hetero-oligomers formed by the small heat-shock proteins αB-crystallin and HSP27. J. Biol. Chem. 288:13602-9

Arrigo AP, Simon S, Gibert B, et al. (2007) Hsp27 (HspB1) and alphaB-crystallin (HspB5) as therapeutic targets. FEBS Letters 581:3665-74

Bagneris C, Bateman OA, Naylor CE, Cronin N, Boelens WC, Keep NH, and Slingsby C (2009) Crystal structures of alpha-crystallin domain dimers of alphaB-crystallin and Hsp20. J. Mol. Biol. 392:1242-52

Baldwin AJ, Lioe A, Hilton GR, et al. (2011) The polydispersity of αB-crystallin is rationalized by an interconverting polyhedral architecture. Structure 19:1855-63

Baldwin AJ, Walsh P, Hansen DF, et al. (2012) Probing dynamic conformations of the high molecular weight αB-crystallin heat shock protein ensemble by NMR spectroscopy. JACS 134:15343-50

Baranova EV, Weeks SD, Beelen S, Bukach, OV, Gusev NB, and Strelkov SV (2011) Three-dimensional structure of α-crystallin domain dimers of human small heat shock proteins HSPB1 and HSPB6. J Mol. Biol. 411:110-22

Basha E, Jones C, Blackwell AE, et al. (2013) An unusual dimeric small heat shock protein provides insight into the mechanism of this class of chaperone. J. Mol. Biol. 425: 1683-96

Baughman HER, Clouser AF, Klevit RE, and Nath A (2018) HspB1 and Hsc70 chaperones engage distinct tau species and have different inhibitory effects on amyloid formation. J Biol. Chem. doi: 10.1074/jbc.M117.803411

Braun N, Zacharias M, Peschek J, et al. (2011) Multiple molecular architectures of the eye lens chaperone αB-crystallin elucidated by a triple hybrid approach. PNAS 108:20491-6

Chalova AS, Sudnitsyna MV, Strelkov SV, and Gusev NB (2014) Characterization of human small heat shock protein HspB1 that carries C-terminal domain mutations associated with hereditary motor neuron diseases. Biochim. Biophys. Acta. 12:2116-26

Clouser AF and Klevit RE (2017) pH-dependent structural modulation is conserved in the human small heat shock protein HSPB1. Cell Stress Chap. 4:569-75

Cox D, Selig E, Griffin MDW, Carver J, and Ecroyd H (2016) Small heat-shock proteins prevent α-synuclein aggregation via transient interactions and their efficacy is affected by the rate of aggregation. J. Biol. Chem. 43:22618-29

Datskevich PN, Nefedova VV, Sudnitsyna MV, and Gusev NB (2012) Mutations of small heat shock proteins and human congenital diseases. Biochemistry (Moscow) 77:1500-14

Delbecq S, Jehle S, and Klevit RE (2012) Binding determinants of the small heat shock protein, αB-crystallin: recognition of the 'IxI' motif. EMBO 31:4587-94

Echaniz-Laguna A, Geuens T, Petiot P, et al. (2017) Axonal neuropathies due to mutations in small heat shock proteins: clinical, genetic, and functional insights into novel mutations. Hum. Mutat. 38:556-568

Eidenschink L, Kier BL, Huggins KNL, and Andersen NH (2009) Very short peptides with stable folds: Building on the interrelationship of Trp/Trp, Trp/cation, and Trp/backbone-amide interaction geometries. Proteins 2:308-22

Evgrafov OV, Mersiyanova I, Irobi J, et al. (2004) Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy. Nature Genetics 71:1660-8

Grishina IB and Woody RW (1994) Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators. Faraday Discuss. 99:245-62

Gusev NB, Bogatcheva NV, and Marston SB (2002) Structure and properties of small heat shock proteins (sHsp) and their interaction with cytoskeleton proteins. Biochemistry 67:511-9

Guttman M, Weis DD, Engen JR, and Lee KK (2013) Analysis of overlapped and noisy Hydrogen/Deuterium exchange data. J. Amer. Soc. Mass Spectrom. 24:1906-12

Haley DA, Bova MP, Huang QL, et al. (2000) Small heat-shock protein structures reveal a continuum from symmetric to variable assemblies. J. Mol. Biol. 298:261-72

Hayes D, Napoli V, Mazurkie A, Stafford WF, and Graceffa P (2009) Phosphorylation dependence of Hsp27 multimeric size and chaperone function. J. Biol. Chem. 284:18801-7

Hochberg GKA, Ecroyd H, Liu C, Cox D, Cascio D, Sawaya M, Collier MP, Stroud J, Carver JA, Baldwin AJ, Robinson CV, Eisenberg DS, Benesch JL, and Laganowsky A (2014) The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. PNAS 111:E1562-70

Houlden H, Laura M, Wavrant-De Vrieze F, et al. (2008) Mutations in the HSP27 (HSPB1) gene cause dominant, recessive, and sporadic distal HMN/CMT type 2. Neurology 71:1660-8

Hyberts SG, Milbradt AG, Wagner AB, et al. (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. J. Biomol. NMR 52: 315-27

Jehle S, Vollmar BS, Bardiaux B, et al. (2011) N-terminal domain of αB-crystallin provides a conformational switch for multimerization and structural heterogeneity PNAS 108:6409-14

Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol. Biol. 278:313-52

Jovcevski B, Kelly MA, Rote AP, Berg T, Gastall HY, Benesch JLP, Aquilina JA, and Ecroyd H (2015) Phosphomimics destabilize Hsp27 oligomeric assemblies and enhance chaperone activity. Cell Chem. & Biol. 22:186-95

Kennaway CK, Benesch JLP, Gohlke U, et al. (2005) Dodecameric structure of the small heat shock protein Acr1 from Mycobacterium tuberculosis. J. Biol. Chem. 280:33419-25

Kim KK, Kim R, and Kim SH (1998) Crystal structure of a small heat-shock protein. Letters to Nature 394:595-9

Laganowsky A, Benesch JL, Landau M, et al. (2010) Crystal structures of truncated αA and αB crystallins reveal structural mechanisms of polydispersity important for eye lens function. Protein Science 19:1031-43

Landry J, Chretien P, Laszlo A, and Lambert H (1991) Phosphorylation of HSP27 during development and decay of thermotolerance in Chinese hamster cells. J. Cell Physiology 147:93-101

Lanneau D, Brunet M, Frisan E, et al. (2008) Heat shock proteins: essential proteins for apoptosis regulation. J Cell. Mol. Med. 12:743-61

Larsen JK, Yamboliev IA, Weber LA, and Gerthoffer WT (1997) Phosphorylation of the 27-kDa heat shock protein via p38 MAP kinase and MAPKAP kinase in smooth muscle. Am. J. Physiol. 273:L930-40

Lelj-Gorolla B and Mauk AG (2012) Roles of the N- and C-terminal sequences in Hsp27 self-association and chaperone activity. Protein Science 21:122-33

McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 4:404-5

McDonald E, Bortolus M, Koteiche H, and Mchaourab H (2012) Sequences, structure, and dynamic determinants of Hsp27 (HspB1) equilibrium dissociation are encoded by the N-terminal domain. Biochemistry 51:1257-68

Mchaourab HS, Lin YL, and Spiller BW (2012) Crystal structure of an activated variant of small heat shock protein Hsp16.5. Biochemistry 51:5105-12

Muranova LK, Weeks SD, Strelkov SV, and Gusev NB (2015) Characterization of mutants of human small heat shock protein HspB1 carrying replacements in the N-terminal domain and associated with hereditary motor neuron diseases. PLOS ONE DOI:10.1371/journal.pone.0126248

Nefedova VV, Sudnitsyna MV, Strelkov SV, and Gusev NB (2013) Structure and properties of G84R and L99M mutants of human small heat shock protein HspB1 correlating with motor neuropathy. Arch. Bioch. Bioph. 538:16-24

Peschek J, Braun N, Rohrberg J, et al. (2013) Regulated structural transitions unleash the chaperone activity of αB-crystallin. PNAS 110:3780-9

(a) Rajagopal P, Liu Y, Shi L, Clouser AF, and Klevit RE (2015) Structure of the α-crystallin domain from the redox-sensitive chaperone, HSPB1. J. Biomol. NMR 63:223-8

(b) Rajagopal P, Tse E, Borst AJ, Delbecq SP, Shi L, Southworth DR, and Klevit RE (2015) A conserved histidine modulates HSPB5 structure to trigger chaperone activity in response to stress-related acidosis. eLife DOI: 10.7554/eLife.07304

Ritossa F (1996) Discovery of the heat shock response. Cell Stress Chap. 2:97-98

Rogalla T, Ehrnsperger M, Preville X, et al. (1999) Regulation of Hsp27 oligomerization, chaperon function, and protective activity against oxidative stress/tumor necrosis factor α by phosphorylation. J. Biol. Chem. 274:18947-56

Rossor AM, Davidson GL, Blake J, et al. (2012) A novel p.Glu175X premature stop mutation in the C-terminal end of HSP27 is a cause of CMT2. J. Peripheral Nervous System 17:201-5

Ruschak AM and Kay LE. (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. J. Biomol. NMR 46:75-87

Shashidharamurthy R, Koteiche HA, Dong J, and Mchaourab HS (2005) Mechanism of chaperone function in small heat shock proteins. J. Biol. Chem. 280:5281-9

Shemetov AA, Seit-Nebi AS, and Gusev NB (2011) Phosphorylation of human small heat shock proteins HspB8 (Hsp22) by ERK1 protein kinase. Mol. Cell. Biochem. 355:47-55

Sluchanko NN, Beelan S, Kulikova AA, Weeks SD, Antson AA, Gusev NB, and Strelkov SV (2017) Structural basis for the interaction of a human small heat shock protein with the 14-3-3 universal signaling regulator. Cell Structure 25:305-16

Tang G, Peng L, Baldwin PR, et al. (2007) EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157:38-46

Van Montfort RLM, Basha E, Friedrich KL, et al. (2001) Crystal structure and assembly of a eukaryotic small heat shock protein. NSMB 8:1025-30

Vargas-Roig LM, Gago FE, Tello O, et al. (1998) Heat shock protein expression and drug resistance in breast cancer patients treated with induction chemotherapy. International J. of Cancer 79:468-75

Weeks SD, Baranova EV, Heirbaut M, Beelen S, Shkumatov AV, Gusev NB, and Strelkov SV (2013) Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6. J. Struct. Biol. 185:342-54

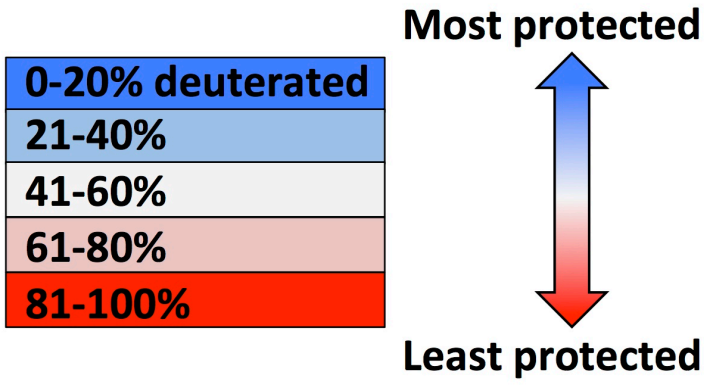
White HE, Orlova EV, Chen S, et al. (2006) Multiple distinct assemblies reveal conformational flexibility in the small heat shock protein Hsp26. Structure 14:1197-204.

Xi J-H, Bai F, Gross J, et al. (2008) Mechanism of Small Heat Shock Protein Function *In Vivo*. J. Biol. Chem. 283:5801-14

Zhang R, Zhang F, Li X, et al. (2014) A novel transgenic mouse model of Chinese Charcot-Marie-Tooth disease type 2L. Neural Regen. Res. 4:413-19

APPENDIX A- HDXMS peptides

The following tables show the deuteration levels of each peptide analyzed at the indicated five time points for the following mutants- WT/C137S, G34R/C137S, P39L/C137S, E41K/C137S, G84R/C137S, D3GXG/C137S, G34R/D3GXG/C137S, P39L/D3GXG/C137S, E41K/D3GXG/C137S, and G84R/D3GXG/C137S. Not all peptides could be analyzed for each mutant due to changes in cleavage sites or overlap in the mass spectra. For peptides demonstrating bimodal behavior in their mass spectra (two conformational populations for one peptide), the relative fractions of each population are shown in the columns labeled "fraction" and the deuteration levels of each population are shown in the columns labeled "population". The average deuteration level for bimodal peptides (weighted for the fractions of each population) is shown in the fifth "Avg" column. For peptides demonstrating only one conformation, the deuteration levels at each time point are listed in the "Avg" column. The deuteration levels are color coded for quick visual analysis based on the following color scheme, consistent with the main text figures.



Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
1-10	3 sec	0.39	0.61	0.41	0.96	0.75	0.37	0.63	0.41	0.97	0.76	0.52	0.48	0.43	0.96	0.69	0.41	0.59	0.39	0.99	0.74	0.43	0.57	0.40	0.93	0.70
	15 sec	0.38	0.62	0.51	0.99	0.80	0.34	0.66	0.54	0.98	0.84	0.45	0.55	0.53	0.97	0.77	0.33	0.67	0.46	0.98	0.81	0.32	0.68	0.51	0.93	0.80
	1 min	0.35	0.65	0.58	0.99	0.85	0.27	0.73	0.62	0.98	0.88	0.46	0.54	0.65	0.97	0.82	0.35	0.65	0.58	0.98	0.84	0.16	0.84	0.56	0.90	0.84
	4 min	1.00	0.00	0.91	0.00	0.91	1.00	0.00	0.95	0.00	0.95	1.00	0.00	0.89	0.00	0.89	1.00	0.00	0.92	0.00	0.92	1.00	0.00	0.91	0.00	0.91
	1 hr	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.94	0.00	0.94	1.00	0.00	0.99	0.00	0.99	1.00	0.00	0.95	0.00	0.95
1-11	3 sec	0.38	0.62	0.32	0.94	0.70	0.34	0.66	0.32	0.93	0.73	0.44	0.56	0.31	0.89	0.64	0.41	0.59	0.31	0.97	0.70	0.38	0.62	0.28	0.88	0.65
	15 sec	0.34	0.66	0.38	0.97	0.77	0.25	0.75	0.39	0.96	0.82	0.36	0.64	0.39	0.94	0.74	0.34	0.66	0.40	0.99	0.79	0.30	0.70	0.38	0.94	0.77
	1 min	0.33	0.67	0.50	0.96	0.81	0.23	0.77	0.51	0.97	0.87	0.29	0.71	0.46	0.94	0.80	0.36	0.64	0.53	0.98	0.82	0.29	0.71	0.50	0.95	0.82
	4 min	0.27	0.73	0.62	0.97	0.87	1.00	0.00	0.96	0.00	0.96	0.22	0.78	0.53	0.95	0.86	1.00	0.00	0.90	0.00	0.90	1.00	0.00	0.92	0.00	0.92
	1 hr	1.00	0.00	0.95	0.00	0.95	1.00	0.00	0.97	0.00	0.97	1.00	0.00	0.94	0.00	0.94	1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.95	0.00	0.95
2-10	3 sec	0.46	0.54	0.33	0.95	0.67	0.40	0.60	0.31	0.96	0.70	0.50	0.50	0.31	0.93	0.62	0.46	0.54	0.30	0.98	0.67	0.43	0.57	0.27	0.92	0.64
	15 sec	0.42	0.58	0.40	0.97	0.73	0.37	0.63	0.44	0.98	0.78	0.42	0.58	0.40	0.95	0.72	0.41	0.59	0.41	0.98	0.74	0.34	0.66	0.36	0.94	0.74
	1 min	0.39	0.61	0.49	0.96	0.78	0.31	0.69	0.56	0.98	0.85	0.52	0.48	0.58	0.99	0.78	0.42	0.58	0.52	0.97	0.78	0.44	0.56	0.57	1.00	0.81
	4 min	1.00	0.00	0.90	0.00	0.90	1.00	0.00	0.94	0.00	0.94	0.47	0.53	0.71	0.98	0.85	1.00	0.00	0.89	0.00	0.89	1.00	0.00	0.91	0.00	0.91
	1 hr	1.00	0.00	0.95	0.00	0.95	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.94	0.00	0.94	1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.97	0.00	0.97
2-11	3 sec	0.47	0.53	0.26	0.94	0.62	0.44	0.56	0.28	0.93	0.64	0.56	0.44	0.29	0.90	0.56	0.47	0.53	0.25	0.93	0.61	0.54	0.46	0.26	0.89	0.56
	15 sec	0.44	0.56	0.34	0.97	0.69	0.38	0.62	0.40	0.97	0.75	0.51	0.49	0.40	0.96	0.68	0.44	0.56	0.36	0.97	0.70	0.43	0.57	0.36	0.94	0.69
	1 min	0.39	0.61	0.43	0.97	0.76	0.27	0.73	0.49	0.95	0.83	0.50	0.50	0.52	0.97	0.75	0.41	0.59	0.48	0.95	0.75	0.33	0.67	0.45	0.94	0.78
	4 min	0.36	0.64	0.60	0.98	0.84	1.00	0.00	0.95	0.00	0.95	0.40	0.60	0.64	0.98	0.84	1.00	0.00	0.88	0.00	0.88	0.26	0.74	0.63	0.96	0.88
	1 hr	1.00	0.00	0.95	0.00	0.95	1.00	0.00	0.97	0.00	0.97	1.00	0.00	0.94	0.00	0.94	1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.96	0.00	0.96
11-28	3 sec					0.61					0.65					0.52					0.63					0.59
	15 sec					0.77					0.87					0.75					0.85					0.83
	1 min					0.85					0.93					0.85					0.89					0.90
	4 min					0.88					0.94					0.90					0.91					0.92
	1 hr					0.97					0.92					0.92					0.95					0.91
12-28	3 sec					0.58					0.61					0.48					0.62					0.55
	15 sec					0.75					0.85					0.71					0.83					0.81
	1 min					0.85					0.91					0.83					0.87					0.88
	4 min					0.87					0.92					0.88					0.89					0.90
	1 hr					0.92					0.91					0.89					0.93					0.89
29-41	3 sec	0.67	0.33	0.45	0.93	0.61	0.17	0.83	0.45	0.83	0.77	0.40	0.60	0.15	0.54	0.38						0.23	0.77	0.41	0.77	0.69
	15 sec	0.51	0.49	0.48	0.96	0.72	1.00	0.00	0.97	0.00	0.97	0.44	0.56	0.36	0.84	0.63						1.00	0.00	0.94	0.00	0.94
	1 min	0.46	0.54	0.58	1.00	0.81	1.00	0.00	0.98	0.00	0.98	0.21	0.79	0.39	0.90	0.79						1.00	0.00	0.96	0.00	0.96
	4 min	1.00	0.00	0.91	0.00	0.91	1.00	0.00	0.98	0.00	0.98	0.09	0.91	0.41	0.94	0.89						1.00	0.00	0.97	0.00	0.97
	1 hr	1.00	0.00	0.99	0.00	0.99	1.00	0.00	0.97	0.00	0.97	1.00	0.00	0.95	0.00	0.95						1.00	0.00	0.96	0.00	0.96

Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
29-43	3 sec											0.18	0.82	0.42	0.94	0.85						0.22	0.78	0.45	0.80	0.72
	15 sec											0.10	0.90	0.43	0.93	0.88						1.00	0.00	0.95	0.00	0.95
	1 min											0.06	0.94	0.37	0.93	0.90						1.00	0.00	0.97	0.00	0.97
	4 min											1.00	0.00	0.95	0.00	0.95						1.00	0.00	0.97	0.00	0.97
	1 hr											1.00	0.00	0.96	0.00	0.96						1.00	0.00	0.96	0.00	0.96
33-41	3 sec					0.70					0.74															0.66
	15 sec					0.83					0.96															0.93
	1 min					0.89					0.96															0.96
	4 min					0.94					1.00															0.96
	1 hr					0.99					0.98															0.95
33-54	3 sec																				0.86					
	15 sec																				0.97					
	1 min																				0.98					
	4 min																				1.00					
	1 hr																				1.02					
42-62	3 sec	0.38	0.62	0.57	0.92	0.79	0.15	0.85	0.49	0.83	0.78	0.32	0.68	0.32	0.71	0.59						0.22	0.78	0.48	0.79	0.72
	15 sec	0.19	0.81	0.62	0.94	0.88	0.05	0.95	0.46	0.97	0.95	0.20	0.80	0.45	0.89	0.80						1.00		0.94	0.00	0.94
	1 min	1.00	0.00	0.97	0.00	0.97	1.00		0.98	0.00	0.98	1.00		0.94	0.00	0.94						1.00		0.97	0.00	0.97
	4 min	1.00	0.00	1.00	0.00	1.00	1.00		0.99	0.00	0.99	1.00		0.98	0.00	0.98						1.00		0.98	0.00	0.98
	1 hr	1.00	0.00	1.09	0.00	1.09	1.00		0.98	0.00	0.98	1.00		0.98	0.00	0.98						1.00		0.97	0.00	0.97
54-62	3 sec						0.27	0.73	0.53	0.84	0.76	0.31	0.69	0.36	0.76	0.64					0.78					0.70
	15 sec						1.00	0.00	0.97	0.00	0.97	1.00	0.00	0.88	0.00	0.88					0.95					0.94
	1 min						1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.95	0.00	0.95					0.99					0.99
	4 min						1.00	0.00	0.99	0.00	0.99	1.00	0.00	0.98	0.00	0.98					1.03					1.00
	1 hr						1.00	0.00	0.99	0.00	0.99	1.00	0.00	0.99	0.00	0.99					1.03					0.99
55-62	3 sec					0.85					0.78					0.71					0.82					0.72
	15 sec					0.95					0.96					0.91					0.99					0.94
	1 min					1.02					0.99					0.98					1.00					0.98
	4 min					1.01					0.99					1.00					1.01					0.98
	1 hr					1.02					0.98					0.99					1.02					0.98
55-76	3 sec										0.95					0.89					1.00					-
	15 sec										1.00					0.96					1.04					0.96
	1 min										1.00					0.97					1.03					0.98
	4 min										1.01					0.99					1.05					0.99
	1 hr										1.00					0.99					1.06					0.98

Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
63-73	3 sec					1.00					0.97					0.94										0.96
	15 sec					1.01					0.99					0.98										0.99
	1 min					1.00					0.98					0.99										1.00
	4 min					1.01					0.99					1.00										1.01
	1 hr					1.01					0.98					1.00										1.00
63-76	3 sec					0.99					0.98					0.95					1.02					0.97
	15 sec					1.00					0.99					0.98					1.03					0.98
	1 min					0.99					0.99					0.98					1.01					0.99
	4 min					1.00					0.98					0.99					1.03					0.99
	1 hr					1.00					0.98					0.99					1.05					0.98
63-84	3 sec										0.98										0.98					
	15 sec										0.98										1.01					
	1 min										0.98										1.00					
	4 min										0.98										1.00					
	1 hr										0.97										1.01					
63-99	3 sec										0.80	0.24	0.76	0.47	0.77	0.70										
	15 sec										0.83	0.23	0.77	0.57	0.80	0.75										
	1 min										0.84	0.32	0.68	0.65	0.83	0.77										
	4 min										0.88	0.27	0.73	0.69	0.87	0.82										
	1 hr										0.90	1.00	0.00	0.88	0.00	0.88										
77-84	3 sec					0.91					0.98					0.94					0.95					
	15 sec					0.97					0.98					0.96					0.98					
	1 min					0.96					0.97					0.96					0.97					
	4 min					0.98					0.97					0.98					0.99					
	1 hr					0.98					0.96					0.97					1.00					
77-87	3 sec	0.52	0.48	0.62	0.99	0.80					0.97	0.21	0.79	0.44	0.94	0.84					0.84					0.97
	15 sec	0.46	0.54	0.67	1.00	0.85					0.97	0.28	0.72	0.63	0.96	0.87					0.88					0.96
	1 min	0.48	0.52	0.72	1.00	0.86					0.97	0.24	0.76	0.69	0.96	0.90					0.88					0.97
	4 min	0.35	0.65	0.76	1.00	0.92					0.97	0.29	0.71	0.76	0.99	0.92					0.94					0.98
	1 hr	1.00	0.00	0.98	0.00	0.98					0.96	1.00	0.00	0.96	0.00	0.96					0.99					0.96
77-99	3 sec	0.47	0.53	0.37	0.68	0.54					0.69	0.29	0.71	0.32	0.67	0.57	0.51	0.49	0.35	0.69	0.52					0.66
	15 sec	0.39	0.61	0.33	0.64	0.52					0.72	0.26	0.74	0.38	0.69	0.61	0.42	0.58	0.40	0.72	0.59					0.69
	1 min	0.38	0.62	0.45	0.74	0.63					0.75	0.18	0.82	0.40	0.71	0.66	0.35	0.65	0.44	0.74	0.63					0.73
	4 min	0.28	0.72	0.43	0.71	0.63					0.79	0.15	0.85	0.46	0.77	0.72	0.23	0.77	0.47	0.77	0.70					0.79
	1 hr	1.00	0.00	0.80	0.00	0.80					0.84	1.00	0.00	0.81	0.00	0.81	1.00	0.00	0.82	0.00	0.82					0.83

Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
87-99	3 sec	0.11	0.89	0.08	0.30	0.28					0.33	0.99	0.02	0.28	0.27	0.28					0.26					
	15 sec	0.57	0.43	0.23	0.45	0.32					0.38	0.40	0.60	0.16	0.42	0.32					0.33					
	1 min	0.16	0.84	0.12	0.44	0.39					0.45	0.82	0.18	0.36	0.60	0.40					0.39					
	4 min	0.13	0.87	0.20	0.50	0.46					0.57	0.20	0.80	0.25	0.56	0.50					0.47					
	1 hr	1.00		0.60	0.00	0.60					0.68	1.00	0.00	0.63	0.00	0.63					0.62					
88-95	3 sec					0.49					0.59					0.53					0.50					0.58
	15 sec					0.54					0.62					0.56					0.55					0.61
	1 min					0.57					0.64					0.61					0.60					0.63
	4 min					0.63					0.68					0.66					0.65					0.69
	1 hr					0.74					0.83					0.78					0.79					0.82
88-99	3 sec					0.24					0.25					0.23					0.22					0.25
	15 sec					0.29					0.31					0.28					0.28					0.31
	1 min					0.36					0.38					0.36					0.34					0.40
	4 min					0.43					0.50					0.46					0.43					0.52
	1 hr					0.56					0.64					0.59					0.60					0.62
100-109	3 sec					0.10					0.11					0.09					0.08					0.11
	15 sec					0.12					0.21					0.15					0.13					0.19
	1 min					0.16					0.29					0.21					0.17					0.27
	4 min					0.22					0.42					0.30					0.24					0.40
	1 hr					0.38					0.68					0.56					0.48					0.68
110-132	3 sec					0.22					0.28					0.26					0.23					0.27
	15 sec					0.24					0.33					0.29					0.26					0.32
	1 min					0.27					0.38					0.33					0.26					0.36
	4 min					0.31					0.42					0.38					0.32					0.41
	1 hr					0.45					0.52					0.50					0.49					0.52
110-137	3 sec					0.20					0.26					0.25					0.20					0.25
	15 sec					0.16					0.31					0.28					0.23					0.30
	1 min					0.24					0.35					0.31					0.25					0.34
	4 min					0.22					0.40					0.36					0.30					0.40
	1 hr					0.42					0.57					0.51					0.44					0.57
133-149	3 sec					0.23					0.23					0.22					0.25					0.22
	15 sec					0.25					0.27					0.26					0.29					0.25
	1 min					0.28					0.30					0.28					0.31					0.28
	4 min					0.32					0.35					0.33					0.36					0.33
	1 hr					0.39					0.52					0.43					0.43					0.50

Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
138-149	3 sec					0.33					0.31					0.30					0.31					0.30
	15 sec					0.36					0.36					0.35					0.37					0.36
	1 min					0.39					0.41					0.39					0.40					0.39
	4 min					0.46					0.49					0.46					0.47					0.47
	1 hr					0.55					0.65					0.58					0.59					0.63
138-163	3 sec					0.26					0.26					0.24					0.25					0.24
	15 sec					0.24					0.33					0.31					0.32					0.31
	1 min					0.37					0.40					0.38					0.38					0.38
	4 min					0.38					0.50					0.46					0.46					0.48
	1 hr					0.59					0.67					0.62					0.63					0.65
150-163	3 sec										0.16					0.14										0.15
	15 sec										0.23					0.21										0.21
	1 min										0.32					0.29										0.30
	4 min										0.43					0.38										0.41
	1 hr										0.63					0.58										0.62
150-164	3 sec										0.12															0.11
	15 sec										0.19															0.17
	1 min										0.27															0.25
	4 min										0.37															0.35
	1 hr										0.57															0.56
164-172	3 sec										0.41					0.41					0.43					0.41
	15 sec										0.42					0.41					0.44					0.42
	1 min										0.44					0.43					0.46					0.45
	4 min										0.53					0.50					0.54					0.55
	1 hr										0.62					0.61					0.63					0.62
164-178	3 sec					0.70					0.70					0.69					0.71					0.68
	15 sec					0.70					0.70					0.69					0.72					0.68
	1 min					0.71					0.71					0.70					0.71					0.70
	4 min					0.74					0.76					0.74					0.76					0.75
	1 hr					0.79					0.79					0.78					0.82					0.78
165-178	3 sec										0.76					0.75										
	15 sec										0.76					0.76										
	1 min										0.77					0.77										
	4 min										0.83					0.81										
	1 hr										0.86					0.86										

Peptide	Time	WT/C137S					G34R/C137S					P39L/C137S					E41K/C137S					G84R/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
		1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
179-185	3 sec	0.45	0.55	0.50	0.99	0.77	0.20	0.80	0.36	0.93	0.81	0.33	0.67	0.36	0.92	0.74	0.41	0.59	0.34	0.94	0.69	0.15	0.85	0.30	0.83	0.75
	15 sec	1.00		0.88	0.00	0.88	1.00		0.97	0.00	0.97	0.13	0.87	0.45	0.98	0.91	1.00		0.94	0.00	0.94	1.00		0.94	0.00	0.94
	1 min	1.00		0.96	0.00	0.96	1.00		0.94	0.00	0.94	1.00		1.01	0.00	1.01	1.00		0.96	0.00	0.96	1.00		0.95	0.00	0.95
	4 min	1.00		0.99	0.00	0.99	1.00		0.95	0.00	0.95	1.00		1.04	0.00	1.04	1.00		1.01	0.00	1.01	1.00		0.96	0.00	0.96
	1 hr	1.00		0.99	0.00	0.99	1.00		0.98	0.00	0.98	1.00		0.99	0.00	0.99	1.00		1.00	0.00	1.00	1.00		0.98	0.00	0.98
179-205	3 sec										0.93					0.92					0.93					0.91
	15 sec										0.98					0.96					0.98					0.96
	1 min										0.98					0.97					1.00					0.97
	4 min										0.99					0.98					1.00					0.98
	1 hr										0.98					0.98					1.01					0.97
186-201	3 sec					0.98					0.97					0.97					1.01					0.97
	15 sec					0.98					0.96					0.96					1.01					0.96
	1 min					0.96					0.96					0.96					0.98					0.97
	4 min					0.98					0.96					0.97					1.00					0.97
	1 hr					0.98					0.95					0.96					1.01					0.96
186-205	3 sec					0.98					0.95					0.95					0.99					0.94
	15 sec					0.99					0.97					0.97					1.02					0.96
	1 min					0.97					0.98					0.97					1.00					0.98
	4 min					0.99					0.97					0.98					1.01					0.98
	1 hr					0.99					0.97					0.98					1.02					0.97
187-205	3 sec					0.98					0.95					0.95					0.99					0.94
	15 sec					0.99					0.97					0.97					1.02					0.97
	1 min					0.97					0.97					0.97					1.00					0.98
	4 min					0.99					0.97					0.98					1.01					0.98
	1 hr					0.99					0.97					0.98					1.02					0.97

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
1-10	3 sec	0.85	0.15	0.38	0.87	0.45	0.44	0.56	0.41	0.96	0.72	0.46	0.54	0.41	0.94	0.70	0.68	0.32	0.37	0.87	0.53					0.44
	15 sec	0.86	0.14	0.53	0.99	0.59	0.24	0.76	0.51	1.00	0.88	0.26	0.74	0.54	0.99	0.87	0.70	0.30	0.43	0.92	0.58					0.58
	1 min	1.00	0.00	0.81	0.00	0.81	1.00	0.00	0.94	0.00	0.94	0.14	0.86	0.51	0.97	0.90	0.59	0.41	0.52	0.93	0.69					0.85
	4 min	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.99	0.00	0.99	1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.80	0.00	0.80					0.98
	1 hr	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.97	0.00	0.97	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.96	0.00	0.96					1.01
1-11	3 sec	0.89	0.11	0.33	0.89	0.39	0.48	0.52	0.38	0.93	0.67	0.45	0.55	0.34	0.89	0.64						0.84	0.16	0.32	0.71	0.38
	15 sec	0.83	0.17	0.45	0.93	0.53	0.27	0.73	0.48	1.00	0.86	0.26	0.74	0.46	0.99	0.85						0.87	0.13	0.47	0.87	0.52
	1 min	0.40	0.60	0.59	0.87	0.76	1.00	0.00	0.94	0.00	0.94	0.12	0.88	0.41	0.97	0.90						1.00	0.00	0.79	0.00	0.79
	4 min	1.00	0.00	0.96	0.00	0.96	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.99	0.00	0.99						1.00	0.00	0.97	0.00	0.97
	1 hr	1.00	0.00	0.96	0.00	0.96	1.00	0.00	0.98	0.00	0.98	1.00	0.00	0.98	0.00	0.98						1.00	0.00	1.01	0.00	1.01
2-10	3 sec	0.65	0.35	0.28	0.61	0.39	0.50	0.50	0.35	0.94	0.64	0.51	0.49	0.34	0.93	0.63	0.79	0.21	0.32	0.89	0.44					0.36
	15 sec	0.79	0.21	0.44	0.89	0.53	0.30	0.70	0.49	0.99	0.84	0.29	0.71	0.46	1.00	0.84	0.74	0.26	0.36	0.88	0.49					0.50
	1 min	1.00	0.00	0.77	0.00	0.77	1.00		0.94	0.00	0.94	0.16	0.84	0.47	0.96	0.88	0.69	0.31	0.47	0.93	0.61					0.80
	4 min	1.00	0.00	0.96	0.00	0.96	1.00		0.99	0.00	0.99	1.00	0.00	0.98	0.00	0.98	1.00		0.76	0.00	0.76					0.98
	1 hr	1.00	0.00	0.96	0.00	0.96	1.00		0.97	0.00	0.97	1.00	0.00	0.97	0.00	0.97	1.00		0.95	0.00	0.95					1.01
2-11	3 sec	0.75	0.25	0.23	0.60	0.32	0.48	0.52	0.28	0.87	0.58	0.45	0.55	0.26	0.85	0.58	0.77	0.23	0.27	0.93	0.42					0.27
	15 sec	0.70	0.30	0.34	0.75	0.47	0.28	0.72	0.43	0.98	0.82	0.27	0.73	0.39	0.97	0.82	0.74	0.26	0.32	0.98	0.49					0.43
	1 min	0.56	0.44	0.58	0.89	0.72	1.00		0.94	0.00	0.94	0.11	0.89	0.34	0.95	0.88	0.64	0.36	0.50	1.00	0.68					0.74
	4 min	1.00	0.00	0.95	0.00	0.95	1.00		1.00	0.00	1.00	1.00	0.00	0.98	0.00	0.98	1.00		0.76	0.00	0.76					0.95
	1 hr	1.00	0.00	0.96	0.00	0.96	1.00		0.97	0.00	0.97	1.00	0.00	0.98	0.00	0.98	1.00		1.02	0.00	1.02					1.01
11-28	3 sec					0.76					0.59					0.50					0.56					0.79
	15 sec					0.84					0.85					0.84					0.71					0.86
	1 min					0.90					0.94					0.92					0.84					0.92
	4 min					0.97					0.95					0.96					0.89					0.94
	1 hr					0.96					0.96					0.96					0.92					0.95
12-28	3 sec					0.81					0.54					0.44					0.52					0.84
	15 sec					0.88					0.83					0.80					0.68					0.88
	1 min					0.87					0.92					0.90					0.82					0.93
	4 min					0.90					0.94					0.94					0.87					0.92
	1 hr					0.88					0.94					0.94					0.90					0.93
29-38	3 sec											0.55	0.45	0.26	0.71	0.46										
	15 sec											0.20	0.80	0.35	0.92	0.80										
	1 min											0.13	0.87	0.43	0.97	0.90										
	4 min											1.00		0.98	0.00	0.98										
	1 hr											1.00		0.99	0.00	0.99										

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
29-41	3 sec					0.68					0.78															0.66
	15 sec					0.76					0.97															0.74
	1 min					0.88					0.97															0.89
	4 min					0.98					0.99															0.99
	1 hr					0.99					0.98															1.01
29-43	3 sec																									0.58
	15 sec																									0.69
	1 min																									0.88
	4 min																									0.98
	1 hr																									1.01
33-41	3 sec					0.49										1.01										0.46
	15 sec					0.61										1.04										0.58
	1 min					0.81										1.02										0.83
	4 min					0.98										1.05										0.99
	1 hr					0.98										1.02										1.01
33-54	3 sec																				0.89					
	15 sec																				0.94					
	1 min																				0.97					
	4 min																				1.00					
	1 hr																				1.02					
42-62	3 sec					0.62					0.82	0.22	0.78	0.42	0.77	0.69										0.59
	15 sec					0.74					1.00	0.10	0.90	0.55	0.96	0.92										0.73
	1 min					0.90					1.00	1.00	0.00	1.00	0.00	1.00										0.91
	4 min					1.00					1.01	1.00	0.00	1.03	0.00	1.03										0.99
	1 hr					1.01					1.02	1.00	0.00	1.05	0.00	1.05										1.02
54-62	3 sec					0.88					0.75					0.77					0.68					0.79
	15 sec					0.88					0.95					0.98					0.79					0.88
	1 min					0.96					0.97					0.99					0.97					0.96
	4 min					1.13					1.00					1.04					1.03					0.99
	1 hr					1.05					1.00					1.04					0.98					1.02
55-62	3 sec					0.96					0.80					0.78					0.78					0.91
	15 sec					1.00					0.99					0.97					0.90					0.97
	1 min					1.02					0.99					0.99					1.01					1.01
	4 min					1.03					1.03					1.03					1.09					1.01
	1 hr					1.03					1.01					1.02					1.08					1.03

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
63-73	3 sec																									1.02
	15 sec																									0.99
	1 min																									1.01
	4 min																									1.01
	1 hr																									1.02
63-76	3 sec					1.00					1.04					1.02					1.00					1.02
	15 sec					0.99					1.04					1.04					0.99					0.98
	1 min					0.99					1.01					1.01					1.01					1.02
	4 min					1.01					1.04					1.05					1.02					1.01
	1 hr					0.99					1.01					1.02					1.02					1.02
63-84	3 sec															1.00										
	15 sec															1.00										
	1 min															1.00										
	4 min															1.01										
	1 hr															0.99										
63-87	3 sec																				0.97					
	15 sec																				0.94					
	1 min																				0.96					
	4 min																				1.02					
	1 hr																				1.01					
63-99	3 sec					0.84					0.81					0.81										0.84
	15 sec					0.80					0.83					0.82										0.81
	1 min					0.81					0.84					0.84										0.89
	4 min					0.88					0.88					0.89										0.91
	1 hr					0.89					0.92					0.91										0.94
77-87	3 sec					0.98					1.00					0.99	0.27	0.73	0.48	1.00	0.86					1.01
	15 sec					0.98					1.00					1.00	0.31	0.69	0.66	0.98	0.88					0.96
	1 min					0.97					0.98					0.98	1.00		0.93	0.00	0.93					1.01
	4 min					0.99					1.00					1.01	1.00		0.97	0.00	0.97					1.00
	1 hr					0.97					0.98					0.98	1.00		1.03	0.00	1.03					1.00
77-95	3 sec																0.35	0.65	0.36	0.84	0.68					
	15 sec																0.32	0.68	0.42	0.84	0.70					
	1 min																0.25	0.75	0.51	0.87	0.78					
	4 min																1.00		0.79	0.00	0.79					
	1 hr																1.00		0.89	0.00	0.89					

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
77-99	3 sec					0.67					0.66					0.66	0.45	0.55	0.31	0.69	0.52					0.69
	15 sec					0.69					0.68					0.69	0.41	0.59	0.34	0.68	0.54					0.68
	1 min					0.71					0.72					0.72	0.39	0.61	0.43	0.74	0.62					0.77
	4 min					0.80					0.78					0.79	0.24	0.76	0.39	0.73	0.65					0.82
	1 hr					0.83					0.85					0.84	1.00		0.77	0.00	0.77					0.86
87-99	3 sec					0.35					0.32					0.32										
	15 sec					0.38					0.37					0.38										
	1 min					0.46					0.44					0.44										
	4 min					0.60					0.44					0.58										
	1 hr					0.66					0.72					0.71										
88-95	3 sec					0.58					0.60					0.63					0.54					0.62
	15 sec					0.60					0.64					0.66					0.51					0.57
	1 min					0.59					0.64					0.66					0.57					0.66
	4 min					0.72					0.73					0.77					0.61					0.74
	1 hr					0.77					0.89					0.87					0.71					0.85
88-99	3 sec					0.28					0.24					0.26					0.26					0.28
	15 sec					0.31					0.29					0.32					0.27					0.30
	1 min					0.40					0.37					0.39					0.34					0.45
	4 min					0.56					0.53					0.55					0.42					0.58
	1 hr					0.63					0.68					0.68					0.54					0.67
100-109	3 sec					0.10					0.19					0.16					0.11					0.08
	15 sec					0.11					0.33					0.31					0.12					0.09
	1 min					0.15					0.43					0.40					0.18					0.16
	4 min					0.23					0.65					0.59					0.25					0.23
	1 hr					0.42					0.80					0.79					0.41					0.42
110-132	3 sec					0.27					0.27					0.30					0.25					0.28
	15 sec					0.29					0.34					0.35					0.24					0.29
	1 min					0.34					0.39					0.39					0.28					0.36
	4 min					0.41					0.45					0.47					0.33					0.41
	1 hr					0.49					0.54					0.54					0.47					0.53
110-137	3 sec					0.24					0.25					0.25					0.23					0.24
	15 sec					0.24					0.30					0.29					0.22					0.25
	1 min					0.29					0.35					0.33					0.25					0.33
	4 min					0.38					0.41					0.41					0.34					0.41
	1 hr					0.58					0.62					0.56					0.43					0.63

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
133-149	3 sec	1.00	0.00	0.25	0.00	0.25	1.00	0.00	0.25	0.00	0.25	1.00	0.00	0.24	0.00	0.24										0.26
	15 sec	1.00	0.00	0.26	0.00	0.26	1.00	0.00	0.27	0.00	0.27	1.00	0.00	0.28	0.00	0.28										0.27
	1 min	1.00	0.00	0.31	0.00	0.31	1.00	0.00	0.30	0.00	0.30	1.00	0.00	0.30	0.00	0.30										0.33
	4 min	1.00	0.00	0.36	0.00	0.36	1.00	0.00	0.37	0.00	0.37	1.00	0.00	0.37	0.00	0.37										0.41
	1 hr	0.23	0.77	0.32	0.73	0.64	0.67	0.33	0.57	0.89	0.68	0.73	0.27	0.50	0.87	0.61										0.79
138-149	3 sec					0.36					0.34					0.33					0.34					0.36
	15 sec					0.38					0.38					0.40					0.35					0.38
	1 min					0.45					0.41					0.42					0.39					0.47
	4 min					0.51					0.51					0.53					0.46					0.53
	1 hr					0.69					0.73					0.70					0.59					0.72
138-163	3 sec					0.25					0.29					0.27					0.24					0.24
	15 sec					0.30					0.36					0.35					0.27					0.29
	1 min					0.37					0.41					0.40					0.32					0.38
	4 min					0.45					0.53					0.53					0.39					0.46
	1 hr					0.66					0.81					0.80					0.59					0.68
150-163	3 sec					0.14					0.22					0.25					0.16					0.13
	15 sec					0.20					0.28					0.31					0.19					0.19
	1 min					0.28					0.34					0.37					0.24					0.28
	4 min					0.38					0.48					0.53					0.32					0.38
	1 hr					0.61					0.83					0.86					0.58					0.62
164-172	3 sec					0.42					0.41					0.44					0.44					0.42
	15 sec					0.42					0.42					0.45					0.44					0.43
	1 min					0.46					0.44					0.48					0.45					0.48
	4 min					0.58					0.57					0.59					0.50					-
	1 hr					0.63					0.65					0.67					0.64					0.64
164-185	3 sec					0.79					0.79					0.82										0.83
	15 sec					0.78					0.79					0.81										0.80
	1 min					0.80					0.78					0.82										0.85
	4 min					0.85					0.83					0.87										0.87
	1 hr					0.85					0.86					0.89										0.89
186-201	3 sec										1.00					1.02										1.01
	15 sec										1.01					1.01										0.96
	1 min										0.99					1.01										1.00
	4 min										1.01					1.03										0.99
	1 hr										0.99					1.00										1.01

		D3GXG/C137S					G34R/D3GXG/C137S					P39L/D3GXG/C137S					E41K/D3GXG/C137S					G84R/D3GXG/C137S				
		Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population			Fraction		Population		
Peptide	Time	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg	1	2	1	2	Avg
186-205	3 sec					0.98					0.98					0.99					1.02					0.99
	15 sec					0.97					1.00					1.01					1.00					0.96
	1 min					0.97					0.99					1.00					1.00					1.01
	4 min					1.00					1.00					1.03					1.02					1.00
	1 hr					0.98					0.99					1.01					1.02					1.02
187-205	3 sec					0.98					0.98					0.99										0.99
	15 sec					0.97					1.00					1.01										0.97
	1 min					0.97					0.99					1.01										1.01
	4 min					1.00					1.01					1.03										1.00
	1 hr					0.98					0.99					1.01										1.02

APPENDIX B- NMR assignments

The following table lists backbone and $^{13}C_B$ nuclear chemical shifts (CSs) assigned to particular residues in D3/176del or D3/GXG. CSs are listed also for peaks that have not been assigned to specific residues or groups, numbered according to archived ^{1}H-^{15}N-HSQC-TROSY peak-lists for D3/176del (first column). Residue assignments of the CTR of D3/GXG are highlighted in green and do not have peak numbers. For sequentially assigned residues that are assigned to two peaks (e.g. residues 64 and 65), alternative conformations are highlighted in gray or not highlighted.

The ^{1}H-^{15}N-TROSY-based triple resonance experiments used to determine these chemical shifts are listed in the top row. The “i", “i-1”, and “i+1” nomenclature refers to nuclei in the current, preceding, or following residue in the sequence, respectively. Occasionally multiple peaks were observed for one ^{1}H-^{15}N correlation and are listed in the same cell. This behavior could be due either to different nuclei having overlapping CSs or to one nucleus experiencing multiple, similar chemical environments. For unassigned peaks, residue classifications are listed based on characteristic CSs and/or intensity changes observed in ILV-unlabeled ^{1}H-^{15}N-HSQC-TROSY spectra (see Chapter 2). Unlabeling of ILV residues results in dramatic intensity losses for these residues and partial intensity losses for A residues.

Additionally, relative peak intensities (normalized from 0 to 100, with 100 being the highest intensity peak) from a ^{1}H-^{15}N-HSQC-TROSY spectrum and peak intensity ratios from PRE experiments (spin-labeled/quenched) are listed for assigned and unassigned peaks.

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes	Intensity (of 100)	T2C	S65C	S74C	T91C
38	L	35	G	34	8.686	127.056		176.137		53.648	44.840				3.1	0.85	0.30		0.15
97	G	47	L	46	7.812	110.904		177.654		44.580	53.730				1.9				
73	G	48	G	47	8.139	111.260		177.081		44.550					3.8	0.90	0.23	0.00	0.43
103	S	49	G	48	7.705	117.399		174.815		58.083	44.512	63.547			8.2	0.89	0.48	0.62	0.68
187	S	49	G	48	8.276	115.063				58.540	45.200	64.805			0.6				
79	A	62	A	61	8.076	123.165	119.145 119.631	177.481		51.884		18.340		CA(i)/CA(i-1) overlap	33.9	0.79	0.16	0.15	0.45
93	I	63	A	62	7.855	119.714	123.802	177.522	176.153	60.677	51.865	37.824	18.326		46.7				
60	E	64	I	63	8.270	124.398	117.724	176.152		55.998	60.669	29.435	37.801		18.5	0.80		0.03	0.39
64	E	64	I	63	8.210	125.401	116.672	175.939	175.288	55.690	60.477	29.893	37.794		10.2	0.83		0.09	0.13
72	S	65	E	64	8.143	118.262		176.000		55.877		62.837	29.564	CA(i)/CA(i-1) overlap	17.7	0.82		0.03	0.31
84	S	65	E	64	8.036	117.270		175.257		55.275	55.690	63.499	29.878		10.0	0.80	0.46	0.00	0.39
62	A	67	P	66	8.218	123.807	117.930	176.720	177.563	52.444	62.576	18.339	31.162		13.7				
55	A	67	P	66	8.349	125.291	119.083	175.621		52.176		18.205			4.4				
101	V	68	A	67	7.766	118.568	126.721	177.704	175.567	61.541	52.194	31.907	18.198		29.5	0.86	0.11	0.03	0.15
83	V	68	A	67	8.040	119.639	127.048	177.585	175.749	61.321	52.195	32.129	18.224		11.0	0.82	0.14	0.14	0.32
75	A	69	V	68	8.106	127.297	124.158	175.577	176.795	51.659	61.489	18.391	31.864		25.0	0.78	0.07	0.00	0.04
63	A	69	V	68	8.207	127.597	124.318	175.707	176.781	51.630	61.227	18.415	32.082		4.0	0.83		0.00	0.00
82	A	70	A	69	8.041	124.738		176.760	175.456	49.999	51.648	17.278	18.466		35.6	0.78	0.07	0.00	0.03
70	A	72	P	71	8.152	123.776	122.585 118.028	176.648	177.563	51.973	62.649	18.288	31.013 40.621		100.0	0.79	0.03	0.13	0.39
94	Y	73	A	72	7.855	118.512	116.841	177.583	175.832	57.467	52.272	37.790	18.140		21.7	0.76	0.07	0.00	0.00
92	S	74	Y	73	7.927	117.420		175.946		57.810	57.370	63.970			12.9				
78	A	76	R	75	8.093	124.559		176.041		52.215	55.703	18.205	31.174		16.2				
89	L	77	A	76	7.944	121.055		177.706		54.961	52.260	41.407	18.243		27.1				
77	D	78	L	77	8.103	120.929		177.080		53.918	55.024	40.429			19.8	0.84	0.16		0.09
183	R	79	D	78	7.954	121.432		176.164		55.671	54.023				13.6				
69	V	97	R	96	8.186	120.099		173.252		58.823	55.202	33.930	32.277		14.2		0.55		0.06
32	F	104	H	103	8.816	118.449		177.856 174.061		57.038		40.298	30.268	CA(i)/CA(i-1) overlap	4.0	0.88	0.56	0.14	0.68
13	A	105	F	104	9.251	126.763		173.548		49.722	57.073	17.766	40.305		5.8	0.87	0.44	0.44	0.83
114	L	109	E	108	6.791	120.264		174.730		52.655	55.761	45.977			9.9				
37	T	110	L	109	8.723	112.675		176.257		61.070	52.630	71.389	46.007		7.6	0.81	0.41	1.05	0.91
158	T	110	L	109	8.833	112.832		177.610		61.791	52.270	71.132			0.4	0.78	0.41	0.49	0.46
175	T	110	L	109	8.566	113.183		178.726		61.220	51.760	70.473			0.4				
24	V	111	T	110	8.954	124.389		172.474		61.255		32.202	71.443	CA(i)/CA(i-1) overlap	6.8				
14	K	112	V	111	9.229	127.338		175.630	175.200	54.691	61.280	36.823	32.364		7.6	0.58	0.52	0.92	0.70

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes	Intensity (of 100)	T2C	S65C	S74C	T91C
16	T	113	K	112	9.115	112.747		175.218		59.500	54.685	69.986	36.663		7.5	0.43	0.59	0.86	0.82
41	K	114	T	113	8.630	123.575		173.360		55.625	59.438	33.660	70.022		4.5		0.50	0.91	0.82
28	D	115	K	114	8.846	124.646		175.603		55.701		39.471	33.683	CA(i)/CA(i-1) overlap	7.5				
42	G	116	D	115	8.619	106.432		175.597		45.433	55.771				4.6	0.27	0.61	0.94	0.59
184	V	117	G	116	7.896	120.051		173.554	175.818	59.741	45.479	31.910			39.8				
18	V	118	V	117	9.081	124.586		175.656		60.787	59.768	33.082			4.1	0.60	0.98	1.38	0.47
11	E	119	V	118	9.401	128.945		173.408	175.079	53.889	60.761	32.851	32.003		8.4	0.55	0.37	0.73	0.63
5	I	120	E	119	9.668	129.810		175.170		59.606	53.915	40.788	32.749		6.9	0.81	0.41	0.81	0.89
171	K	123	G	122	7.093	122.642		174.056	174.842	57.763	44.040	30.464			0.9				
177	K	123	G	122	7.966	123.343		172.094		56.322	43.449	33.988			15.9				
195	K	123	G	122	7.534	120.599		178.837		58.680	44.980	28.155			1.4				
143	H	124	K	123	8.947	122.209		174.598	175.107	53.637	56.329	32.088	33.810		10.7				
23	E	125	H	124	8.972	122.554		175.084 172.858	176.642	54.518	53.641 56.854	29.361	32.085 65.811		21.0				
36	E	126	E	125	8.727	121.165		176.583 174.493	176.097	57.953	54.526 59.457	29.470	71.275		19.6	0.89	0.46	0.89	0.64
26	R	127	E	126	8.900	124.440		176.078	174.716	53.863	57.957	32.447	29.559		10.9				
40	Q	128	R	127	8.691	122.709		174.736	176.349	55.216	53.827	29.156	32.450		18.9		0.44	0.49	0.66
25	D	129	Q	128	8.931	128.568		176.322		51.871	55.218	41.095	29.121		9.1	0.95	0.39	0.53	0.53
95	H	131	E	130	7.849	116.188		176.911		55.457	58.646	32.369			6.1	0.89	0.34	0.31	0.13
122	G	132	H	131	7.461	108.704		175.814		45.872					6.0	0.89	0.33	0.28	0.15
46	Y	133	G	132	8.497	122.420		170.927		56.594	45.885	40.969			2.3	0.84	0.15	0.29	0.64
112	R	136	S	135	5.868	115.677		172.626		55.099					4.2	1.03	0.63	0.75	0.65
7	C	137	R	136	9.517	118.037		175.201		56.339	55.136	41.150			2.7				
27	F	138	C	137	8.882	118.769		173.301	176.223	58.020		40.832			4.1				
56	T	139	F	138	8.353	116.144		173.703		62.180	57.965	71.529	40.974		5.0	0.94	0.43	0.52	0.90
4	R	140	T	139	9.795	129.181		172.705		52.920	62.116				4.8	0.73	0.63	0.46	0.65
12	Y	142	K	141	9.297	121.297		175.920		55.991	54.082	41.252	35.739		4.8	0.84	0.58	0.76	0.37
22	T	143	Y	142	8.979	119.860		175.624		62.775	56.001	68.673	41.245		8.3	0.78	0.35	0.58	0.05
68	L	144	T	143	8.188	128.597		174.515		52.232	62.759	40.464	68.650		10.8	0.57	0.45	0.76	0.26
51	G	147	P	146	8.424	109.763		177.567		44.678	63.152		30.580		12.0	0.50	0.52	0.52	0.00
128	V	148	G	147	7.072	120.922		174.646		62.427	44.662	30.825			14.2	0.48	0.60	0.62	0.10
186	V	148	G	147	6.992	120.727		174.646		62.400	44.690	30.975			1.7				
29	D	149	V	148	8.830	130.817		174.961		49.873	62.408	41.030	30.767		10.7	0.41	0.41	1.10	0.15
54	T	151	P	150	8.359	109.390		178.959		64.129		68.652			5.9	0.45	0.52	0.89	0.53
81	Q	152	T	151	8.066	118.634		175.287		54.449	64.258	28.247	68.446		7.1	0.52	0.55	0.73	0.32
113	V	153	Q	152	6.810	120.830		175.312	171.911	62.223	54.447	30.852	28.227		14.0	0.72	0.52	0.73	0.62

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative Intensity (of 100)	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes		T2C	S65C	S74C	T91C
6	S	154	V	153	9.574	124.424		175.540		56.849	62.186	65.689	30.817		8.0	0.56	0.47	0.60	0.21
129	T	162	G	161	7.025	113.554		179.952		60.517					5.9	0.91	0.31	0.24	0.35
8	L	163	T	162	9.459	135.077		173.899		53.244	60.524	44.115			5.9	0.98	0.28		0.39
10	T	164	L	163	9.454	124.487		174.631		61.217	53.321	70.002			3.8	0.78	0.24	0.07	0.36
31	V	165	T	164	8.818	128.574		174.314		60.630		32.583			3.5	0.95		0.74	0.22
20	E	166	V	165	9.023	121.858		173.023	174.200	53.973	60.693	34.692	32.451		8.5	0.94	0.58	0.59	0.17
116	A	167	E	166	9.373	120.529				50.082		21.403			1.0	0.60	0.24		0.07
102	M	169	P	168	7.727	120.767		177.902 176.283	174.339	51.955	62.364 58.319	31.280	27.724		19.9	0.85	0.50	0.58	0.60
58	K	171	P	170	8.293	122.322	124.064	176.423		55.623	62.403	31.881		CB(i)/CB(i-1) overlap	31.3	0.74	0.52	0.62	0.07
65	L	172	K	171	8.206	124.636	124.973	176.573	177.242	54.486	55.606	41.472	31.895		49.2	0.76	0.47	0.21	0.06
57	A	173	L	172	8.290	125.539	112.720	177.179 174.329	177.883	52.081	54.460	18.269	41.436		32.4				
91	T	174	A	173	7.939	113.293	122.748	177.829		61.312	52.075	69.328	18.370		23.2	0.73	0.52	0.51	0.07
61	Q	175	T	174	8.242	123.342	123.172	174.437		55.364					11.1				
87	S	176	Q	175	7.950	123.761	123.178	175.031	178.595	59.713	55.358	64.163	28.880		92.9	0.79	0.63	0.56	0.19
	G	192	L	191	8.310	109.850	108.672	177.894	174.283	44.803									
	G	193	G	192	8.066	109.252		174.280	172.229	44.187									
	E	195	P	194	8.508	120.763	124.553	177.552	176.491	56.406	63.224	29.083	31.280						
	A	196	E	195	8.056	125.039	122.806	176.576	177.471	52.111	56.397	18.380	29.175						
	A	197	A	196	8.062	123.298	120.640	177.517	177.753	52.053		18.353							
	K	198	A	197	8.153	121.113	120.652	177.788	176.788	55.816	52.032	32.118	18.309						
	E	201	D	200	8.356	121.835	114.456	176.550		56.665	54.073	29.237	40.649						
	T	202	E	201	8.136	114.969	126.135	176.973	174.394	61.875	56.636	69.432	29.239						
	A	203	T	202	8.040	126.662	123.844	174.408	177.199	52.015	61.920	18.503							
	A	204	A	203	8.098	124.386	125.794	177.217	176.721	52.073		18.401							
	K	205	A	204	7.776	126.324	125.795	176.734	181.382	57.196	52.084	32.813	18.333						
0					10.362	116.620		179.550		58.424					2.3	0.70	0.36	0.83	1.08
1	W sidechain				10.197	130.547									1.8	0.62	0.23	0.02	0.12
2	W sidechain	likely W95			10.043	129.260									3.1	0.72	0.15	0.01	0.01
3	W sidechain				9.965	129.695									3.3	0.64	0.00	0.09	0.27
9	LA				9.450	126.121		170.455		52.542	57.189				3.0			0.34	0.42
15	LA				9.200	123.253		172.358	177.925	53.916		35.074			2.8	0.83	0.32		
17	NDL				9.087	126.032		176.097		53.521	54.404	39.529			5.2		0.30		1.29
19	likely W95				9.054	128.811				54.900	57.000				2.7	0.78	0.36		0.07
21	not ILV				9.003	127.369		174.243		53.580					3.2	0.74	0.28	0.45	0.60
30	not ILV				8.835	116.160		176.283		58.530					4.0	0.77	0.00	0.00	0.08

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes	Intensity (of 100)	T2C	S65C	S74C	T91C
33	HQRKEW		HQRKEW		8.767	125.785		175.273	175.949	56.388	54.670	29.553	32.107		14.3	0.32	0.58	1.01	0.67
34	HQRKEW		HQRKEW		8.765	125.281		175.006		54.754		31.978	31.040	CA(i)/CA(i-1) overlap	7.4	0.19	0.56	0.93	0.80
35	likely T				8.735	114.587		178.829		61.890					1.4	0.92	0.40	0.84	0.81
39	HQRKEW		HQRKEW		8.689	125.063		174.972	175.497	54.659	54.929	31.995	30.867		5.3				
43	likely S				8.574	119.262		173.290		57.235	56.447	66.255	34.596		5.3	0.85	0.48	0.46	0.17
44	not ILV				8.556	123.595		173.168		55.509	59.499	33.665			3.8				
45	G		HREFSDL		8.529	107.155		175.608		45.538	55.898				4.7		0.23		0.62
47	HQRKEW		HQRKEW	peak 33	8.477	127.436		174.871		58.640	56.379	32.179	29.697		9.0	0.53	0.27	1.06	0.63
48	S		likely S		8.463	113.132		172.197		56.530		65.069		CA(i)/CA(i-1) overlap	7.3	0.73	0.49	0.68	0.49
49	QEWHRKM	likely E3	T	likely T2	8.447	125.485		173.871		55.037	61.848	31.109	69.165		6.6		0.46	0.49	0.51
50	not ILV				8.426	118.137		177.189		54.064	52.597				3.2	0.90	0.26	0.05	0.19
52	G				8.388	110.654				45.094					4.2	0.84	0.18	0.00	0.04
53	S				8.378	117.639		173.420		57.725	57.950	63.251			7.3	0.79	0.36	0.27	0.63
59	QEWHRK				8.284	120.637		176.561		55.631		28.292		CA(i)/CA(i-1) overlap	15.7	0.84	0.07	0.20	0.05
66	likely T				8.216	121.784		178.148		61.010					2.4		0.72		0.42
67	not ILV	likely FYA	ITVP		8.201	126.404		176.385		54.618	60.387				11.5	0.92	0.26	0.00	
71	QEWHR		NDL		8.142	121.856		174.551		54.110	51.967	29.036	39.136		7.9	0.83			0.00
74	NDYF		ITVP		8.120	107.578		176.181		54.005	65.078	38.774			4.3	0.92	0.59	0.82	0.90
76					8.101	125.529		175.469				29.136	17.324		1.8				0.18
80	NDL				8.075	121.480		176.155		53.841		40.265			4.4				
85	QEWHR				8.014	122.337		176.264	177.311	56.777	56.280	29.653			5.6				0.06
86	not ILV				7.980	125.336		174.923		52.544					3.7	0.73	0.00		0.41
88	G		FDL		7.946	122.073		174.910		46.535			38.860		2.3				
90	not ILV				7.938	118.784		177.255		57.372					2.3		0.37	0.38	0.20
96					7.825	120.266				58.900	55.400	32.862	30.221		19.0				
98	V		G		7.789	119.252		174.238		61.884	45.057	31.500			6.0				
99					7.776	120.314		177.164		59.753		31.331			9.9				
100	not ILV				7.770	121.639									4.7	0.80	0.34		1.13
104	not ILV				7.680	114.384		174.982		56.870					2.8	0.99	0.79	0.00	0.49
105	G		HREFSDL		7.685	106.680		179.855		46.813	56.620				9.2	0.85	0.52	0.78	0.71
106	FYDL				7.633	120.421		174.494		57.686	54.970	41.143			2.2				
107	not ILV		ITVP		7.626	124.395		175.998		54.353	62.231	30.683			6.1	0.73	0.64	0.45	
108	not ILV		ITVP		7.598	112.436		175.028		55.037	61.824	28.311			7.5	0.80	0.46		0.77
109	QEWHR				7.542	120.464		178.741		58.499	57.080	29.160			1.9				
110	NDL				7.517	121.257		173.199	177.228	52.909	54.820	41.306			0.9				

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative Intensity (of 100)	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes		T2C	S65C	S74C	T91C
111	not ILV				7.507	112.783									10.8	0.74	0.56	0.46	0.11
115	ILV				9.414	121.962									1.6	0.54			0.90
117					9.360	118.449									0.5				
118	S				9.345	114.889		173.004		56.635	55.620	66.648			0.5				
119	not ILV				8.668	109.780									2.3		0.36	0.56	0.38
120	not ILV				8.771	106.975									1.9	0.69			
121	not ILV				9.887	129.093									0.3	0.72	0.36	0.46	0.00
123	G				7.093	110.300		172.593		43.453					4.0	0.86	0.31	0.50	0.51
124	not ILV				7.282	111.336		180.321				29.204			4.3	0.77	0.13	0.12	0.32
125	not ILV				7.376	111.726		180.136				27.693			4.0	0.69	0.37	0.77	0.58
126	not ILV				7.433	113.150		179.964							5.7	0.41	0.50	0.97	0.12
127	not ILV				7.178	121.650		179.568		57.590					4.9	0.85	0.45	0.83	0.59
130	not ILV				6.754	112.688				51.947					16.9	0.75	0.50	0.44	0.07
131	not ILV				6.675	113.084		172.100				28.306			5.3	0.52	0.55	0.72	0.05
132	not ILV				6.638	111.256		180.281							3.9	0.74	0.20	0.01	0.32
133	not ILV				6.341	114.314									1.6	0.83		0.01	0.75
134	not ILV				8.158	116.594									2.4	0.52	0.07	0.22	0.13
135	not ILV				8.074	116.037						42.483	24.250		4.2	0.91		0.04	
136	G				6.905	116.136		181.442		42.929					4.4	0.89	0.53	0.26	0.74
137			A		8.565	121.154				59.669	52.200	36.165	16.574		1.1				
138	ILV		ITVP		8.449	121.077		174.666		55.601	61.141	37.871		CA(i)/CA(i-1) overlap	1.3			0.63	
139	ND				8.969	126.946		175.632		55.043	52.710	42.487			0.8				
140	QEWHR				8.435	127.860		174.71		56.213	57.520	30.423			0.7				
141	IV		HQRKEW	peak 148	8.577	125.086		176.075		62.155	55.990	36.664	32.526		0.7				
142	QEWHR				8.474	123.494		176.673		57.550	58.460	29.297			0.7				
144	ND				8.625	118.612		177.450		52.267	57.660	43.141			0.7				
145	A				8.363	118.960		179.156		54.459	58.390	17.244			0.5				
146					8.214	117.384		174.786		55.471	54.530 44.840	31.068			0.8				
147	LA				8.360	127.416				51.520					1.1	0.74			0.99
148	HQRKEW		G		7.601	119.774		172.326	176.181	56.024	45.200	32.570			2.4				
149	not ILV				7.905	122.489		177.116				28.925	64.078		1.6				
150	QEWHR				7.500	118.353		177.521		57.880		29.100			1.4				
151	QEWHR				7.242	118.841				56.661	58.970	28.325			0.5				
152	KYLV		G		7.075	119.831		173.289		51.863	46.370	33.513			0.8				
153	NDL				7.017	116.997		175.528		54.839		40.733		CA(i)/CA(i-1) overlap	1.2				
154	ILVA				7.399	114.158		177.169							3.0				

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes	Intensity (of 100)	T2C	S65C	S74C	T91C
155	not ILV	likely YK	G		8.241	113.912		171.597		54.512	44.690	39.211			0.6				
156	G				8.565	114.311		171.596		45.060					0.6				
157	T				7.182	105.274				59.335		74.426			0.6				
159	T				8.794	112.477		175.064		59.834	58.557 54.6	70.712			0.4				
160	not ILV				8.233	116.485		177.958		56.928	54.460				0.6				
161	not ILV				9.626	128.927									0.4	0.67	0.41	0.00	0.48
162	not ILV				10.127	130.544									0.9	0.61	0.48		
163	ILVA				8.365	120.019		172.675		53.578		31.402			0.4				
164	G				8.410	111.979		174.239		46.190					0.4				
165	not ILV				8.149	114.534				51.884					0.7	0.48	0.00		
166	not ILV				8.710	115.954		177.141		55.480	56.720				0.5				
167					8.573	116.358									0.3				
168	S				8.945	116.116		174.607		57.614	54.530	66.429			0.2				
169	not ILV				6.738	117.133									0.6				
170					6.772	121.492		179.007							0.7				
172					7.435	122.336		177.933		58.795					0.4				
173	not ILV				7.752	129.812									1.3	0.96	0.62		0.52
174	not ILV				7.621	118.749		176.878		59.410					1.8				
176	not ILV				8.372	116.957		177.855		57.150	57.790	40.097			0.7				
178	G		HREFS		7.912	112.730		176.470		45.149	58.680				0.8				
179	FDLN		PV		7.387	123.880				58.569	63.117	39.267	30.876		2.5				
180	FYDL		QEWHR		7.378	123.098		176.434		57.797	55.480	41.615	28.515		1.0				
181	S				6.771	117.684				54.501		61.661			0.7				
182					7.975	120.840				54.925	52.200				14.2				
185	IV				6.869	120.970				62.250		30.871	58.638		1.3				
188	NDL				7.800	117.014		175.593		55.330	54.600	41.895			1.4				
189					8.732	121.660		175.794		57.590		40.345			3.4	0.96	0.56	0.76	0.74
190					9.922	125.021		172.469	174.004 175.763	54.670	59.700	30.875	32.521		0.2				
191			G		7.893	120.041	124.804	173.554	175.818	59.920	45.420	34.666			32.7				
192	S		likely S		8.972	122.554		175.084 172.858	176.642	56.878	57.359	65.390		CB(i)/CB(i-1) overlap	23.1		0.62	0.71	
193	S				7.903	117.499		175.946		55.716	55.180	64.950			15.6				
194	QEWHR				7.503	118.552		177.122		56.570	51.610	27.620			1.4				
196					8.290	125.539		177.179 174.329	177.883	55.180	59.850	30.770			27.5				
197					8.842	111.707		173.892							0.2				
198					8.119	108.658		175.101							0.3				

	Residue (i)		Residue (i-1)		1H-15N HSQC		NNH	HNCO	HNCACO	HNCA/HNCOCA		HNCB/HNCACB			Relative Intensity (of 100)	PRE Intensity Ratio			
Peak #	Type	Sequence #	(i-1)	Sequence #	HN (i)	N (i)	N (i+1)	CO (i-1)	CO (i)	CA (i)	CA(i-1)	CB (i)	CB(i-1)	Notes		T2C	S65C	S74C	T91C
199	ILVA				8.035	108.380									0.4				
200	not ILV				7.955	108.837									1.2	0.58			
201					7.966	109.540									0.8				
202					8.281	111.079		175.624							0.3				
203					7.954	111.721		177.477							0.6				
204					7.633	107.785									0.6				
205	ILVA				8.336	114.647									0.4		0.07		0.22
206					8.564	114.924		176.854							0.4				
207	not ILV				7.579	114.952									1.0				
208					7.553	115.768		174.842							0.8				
209					7.356	116.793									0.5				
210					7.159	116.832									0.6				
211					7.349	118.979		173.234							0.4				
212	not ILV				7.330	120.975		174.671							1.5	0.66		0.11	
213					7.178	123.140		174.021							0.6				
214	not ILV				7.741	122.506									1.2				
215					7.819	123.777									3.2				
216	ILVA				8.113	128.597		174.515							1.4				
217	not ILV				8.738	130.571									0.8				