# Immature SOD1: protein folding, misfolding and aggregation

by

Harmeen Kaur Deol

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Chemistry

Waterloo, Ontario, Canada, 2022

**Examining Committee Membership**

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

| | |
|---|---|
| External Examiner | Dr. Natalie Goto<br>Associate Professor |
| Supervisor(s) | Dr. Elizabeth M. Meiering<br>Professor |
| Internal Member | Dr. Thorsten Dieckmann<br>Associate Professor and Associate Chair of Graduate Studies and Research |
| | Dr. Subha Kalyaanamoorthy<br>Assistant Professor |
| Internal-external Member | Dr. Todd Holyoak<br>Associate Professor |

## Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

# Abstract:

Protein misfolding and aggregation are hallmarks of many diseases, including amyotrophic lateral sclerosis (ALS). In familial ALS, aberrant self-association of mutant Cu,Zn-superoxide dismutase (SOD1) is implicated as a key contributor to disease. Mutations have the largest impacts on the stability of the most immature form of SOD1, the unmetallated, disulfide-reduced monomer (apoSH SOD1). Here we demonstrate that, despite the marginal stability of apoSH SOD1, aggregation is little correlated with the degree of protein unfolding, and multiple modes of aggregation occur, depending on the mutation and solution conditions. Additionally, we explore the local differences across apoSH SOD1 folded and unfolded structure and dynamics through chemical shift perturbations, both amide proton and nitrogen temperature coefficients, and peak shape analysis. These results provide new evidence that ALS-associated mutations promote the aggregation of apoSH SOD1 through multiple pathways and potentially through subtle differences in local structure and dynamics with broad implications for understanding mechanisms of protein self-association in disease and biotechnology.

# Acknowledgements:

With great honesty, I must admit this was the hardest section to write. To list all the people that helped me get here and to then express how much they have changed my life would make this section longer than the entire thesis, if you can believe it. I have opted for a more personal approach by thanking a lot of these people individually, so please understand these words do not do justice to how impactful every person was on my life.

I would like to express my deepest gratitude to Dr. Elizabeth M. Meiering for offering me this opportunity, under which I have grown immensely. Without this, I would not be the person I am today.

Thank you to my advisory committee members, both past and present: Dr. Thorsten Dieckmann, Dr. Subha Kalyaanamoorthy, Dr. Todd Holyoak, and Dr. Michael Palmer for all their support along the way. To my external, Dr. Natalie Goto, I would like to express my appreciation to you for serving on my examining committee.

To my amazing lab members, past and present, this was a phenomenal ride, filled with so many ups and downs. You all have offered so much thoughtful discussion and endless support, even on days when the science Gods were against us. A heartful thanks goes to Bruna Siebeneichler, Michael Tarasca, Shanise Reddekopp, and Jacob Afelskie. To Colleen Doyle and Kyle Trainor, one of you believed in endless love and the other believed in tough love, you know who you are. Regardless, without you two, I could never have evolved to the scientist I proudly call myself today.

Also a thank you to members in the department that made this possible. Specifically, a special thanks to Catherine Van Esch, who always put out all my fires with a smile. You take on so much for us, and I hope one day you'll know how much you mean to us. To Chem Stores, which includes but is not limited to Sarah Mark, Vladimir Miskovic, and Slavica Bogdanovic, thank you for always supporting me.

Now to my family and friends, who showed me unconditional love and support. On my hardest days, at my weakest moments, there you all stood -- proud of little old me. I wish you all could read my mind and understand how much love I have for each of you. To the friends that helped me prioritize my sanity, specifically Travis and Matthew, thank you for always showing me grace. To the ugly laughs on late nights with Volition La, the tough conversations about intrusive thoughts with Kyle Piccolo, and the self-worth coaching with Dalia Naser, I am a better person for knowing you all.

To Laura Marrone, where do I begin. I am so happy our paths crossed, and I am forever changed thanks to you. You have time and time again chose kindness and understanding in the moments that most would not. All thanks to you, I will enter every tough conversation mentally reciting, "everyone's just trying to survive." You will never believe your impact, but my goodness Laura, you have made all the difference.

Lastly, to the two people I call home, Gurmeet Jouhal and James Sorensen, thank you for always being a safe place to rest my head. Although you are gone now, my dearest *jaan*, you got me so far and look at us now! Thank you for believing in me, especially on days where you had to do it for the both of us. We did not have the easiest life, but I am so grateful for every memory with you. Lastly, to my sweet James Sorensen, you will

never know the depth I feel your love and support. Your love is so patient and kind, and always unconditional. Thank you for holding my shaking hands and never letting go.

Without any of the people listed above and many more, this journey would have not been possible. Thank you endlessly.

# Dedication:

To my *jaan*, Gurmeet Jouhal and sweet James Sorensen, for always believing I was enough.

# Table of Contents

# List of Figures:

# List of Tables:

# List of Abbreviations

| | |
|---|---|
| $^{10}$Fn3D80E | Immunoglobulin-like tenth human fibronectin type III domain D80E |
| AFM | Atomic Force Microscopy |
| Agg | A4V in the presence of aggregates |
| ALS | Amyotrophic Lateral Sclerosis |
| ApoSH SOD1 | Metal-free with disulfide reduced SOD1 |
| ApoSS SOD1 | Metal-free with disulfide oxidized SOD1 |
| Aß | Amyloid-ß |
| ATR-FTIR | Attenuated total reflectance-Fourier transform infrared spectroscopy |
| CCS | Copper chaperone for SOD1 |
| CSP | Chemical shift perturbation |
| Cu | Copper ion |
| Curvature | Curved temperature dependence |
| $D_h$ | Hydrodynamic diameter |
| $D_h$ | Number distribution of hydrodynamic diameter |
| $D_{h,I}$ | Intensity distribution of hydrodynamic diameter |
| DLS | Dynamic light scattering |
| DSC | Differential scanning calorimetry |
| dsRBD-1 | RNA-binding domain |
| DSS | Sodium trimethylsilylpropanesulfonate |
| EDTA | Ethylenediaminetetraacetic acid |
| EZnSH SOD1 | Zinc bound with disulfide reduced SOD1 |
| fALS | Familial amyotrophic lateral sclerosis |
| Folded intensity slope | The intensity slope at the folded temperature range |
| Folded Temperature Range | Temperatures at which the protein remains folded |
| GB1 | B1 domain of protein G |
| HC | V148G at high concentration |
| HEPES | Hydroxyethyl piperazine N- 2-ethanesulfonic acid |
| HoloSS SOD1 | Copper and Zinc bound with disulfide oxidized SOD1 |
| HSQC | Heteronuclear single quantum correlation |
| ITC | Isothermal titration calorimetry |
| K | Boltzmann constant |
| LC | V148G at low concentration |
| Max temperature | Temperature corresponding to the maximum peak intensity observed throughout the temperature range |
| NMR | Nuclear magnetic resonance |

| | |
|---|---|
| PAR | 4-(2-pyridylazo)rescorcinol |
| PDB | Protein data bank |
| PDI | Polydispersity index |
| pWT | Pseudo-wildtype |
| sALS | Sporadic amyotrophic lateral sclerosis |
| SLS | Static light scattering |
| SOD | Superoxide dismutase |
| SOD1 | Cu,Zn-superoxide dismutase 1 |
| $t_0$ | Initial time point |
| $T_{0.5}$ | Melting temperature |
| TCEP-HCl | Tris(2-carboxyethyl)phosphine hydrochloride |
| TDP43 | TAR DNA-binding protein 43 |
| ThT | Thioflavin-T |
| $t_{late}$ | Last time point collected |
| $t_{mid}$ | Time point between initial and the last time point |
| Trp-32 | Tryptophan-32 |
| Trp-32F | Trp-32 sidechain corresponding to the folded state |
| Trp-32U | Trp-32 sidechain corresponding to the unfolded state |
| $T_{unfolding}$ | Temperature at which protein starts to unfold |
| Unagg | A4V in the absence of aggregates |
| Unfolded intensity slope | The intensity slope at the unfolded temperature range |
| Unfolded Temperature Range | Temperatures at which the protein experiences protein unfolding |
| WT | Wildtype |
| Zn | Zinc ion |
| $\delta_{Obs}$ | Measured chemical shift |
| $\delta_{RC}$ | random coil chemical shift |
| $\delta_{Sec}$ | Secondary chemical shift |

# Chapter 1: Introduction

## 1.1 Protein folding, misfolding and aggregation

The intricate balance of driving forces governing protein folding, misfolding and aggregation remain obscure in their intricate details. However, much research has been done on understanding each of these processes and their underlying mechanisms. A common model for protein folding, the energy landscape model, which offers a plausible explanation for how protein folding, misfolding and aggregation can be linked.[1] The model highlights that along the path to find its native structure, a protein samples multiple conformations, which can include both native and non-native conformations. The number of possible conformations and the landscape are ultimately determined by the amino acid sequence.

Even though the energy landscape explains the drive to a more stable structure, with so many possible conformations available, it is unclear how a protein arrives to a completely folded, functional state. Some key models have been proposed to explain the driving forces behind protein folding, and they include: hydrophobic collapse, initial nucleation, and diffusion-collision.[2,3] All these models share similar elements, but show subtle distinctions in the critical residues involved. However, an attractive feature of the energy landscape model is that it allows all three models to exist within itself and creates a unified general mechanism of protein folding.

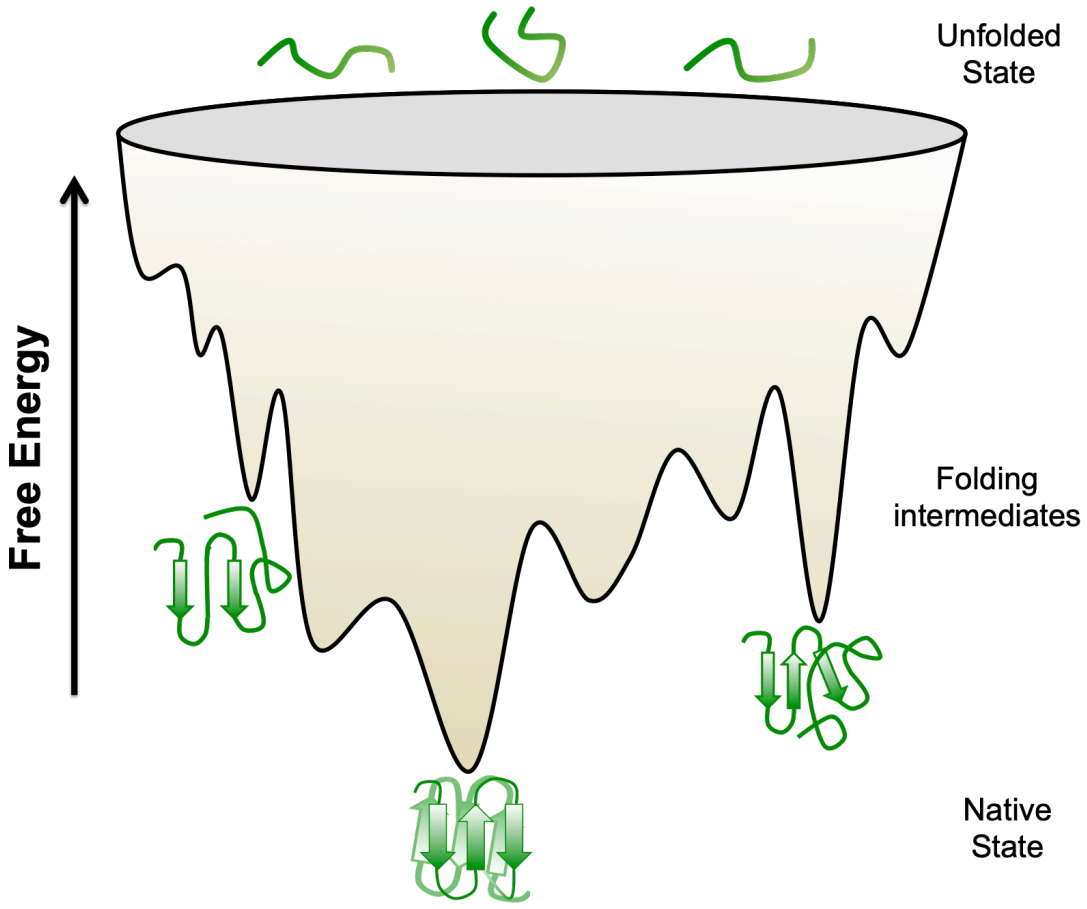



Figure 1.1: Protein energy landscape model. The funnel shape represents the decrease in available conformations the protein can sample as it folds. The dips before native (native state) folding are due to metastable non-native or native intermediates (folding intermediates) that require a certain amount of energy to change conformation to continue on the folding pathway. Adapted and modified from Toal and Schweitzer-Stenner, 2014 and Dill and MacCallum, 2012.[4]

Although the native structure is typically the more stable and therefore the thermodynamically preferred conformation, the protein can also access additional conformations, with some being more or less native-like along the folding pathway.[2,3,5,6] For the non-native conformations, they can represent a local stability minimum, allowing the non-native conformation to populate for a longer duration and sometimes trap the

protein in that conformation. Therefore, along the folding pathways, multiple native-like and non-native-like partially folded states, unfolded protein and natively folded protein are populated.[2] With this large range of differently folded states, it is not surprising that the details of protein folding and aggregation mechanisms have remained elusive.

If the non-native interactions persist and form a non-natively folded protein, it can lead to aberrant intermolecular interactions. However, eukaryotic systems have developed multiple mechanisms to deal with improperly folded proteins: the cells can either attempt to refold them with the aid of chaperones[7], or remove them through the ubiquitin-proteasome pathway.[8] Both of these quality control processes have also been observed to be upregulated during stress. Irrespective of these measures to control misfolded proteins, some misfolded proteins still evade regulatory mechanisms and aberrantly interact with other proteins or the same protein to form protein aggregates.

Protein aggregation is a widely observed phenomenon in all organisms.[9–11] Frequently, protein aggregation is a hallmark of disease but, in some cases, protein aggregation can serve a protective role. To minimize cellular disruption, the cell may concentrate misfolded or aggregated proteins together.[12] Moreover, with so many proteins showing tendency to aggregate, it is not surprising the multitude of possible aggregate structures that have been observed. Aggregate structures can range from soluble globular aggregates to insoluble fibres; however, it is unclear how different aggregate structures are related. Some have proposed that soluble oligomers represent early aggregates, which grow and undergo conformational rearrangement to form insoluble fibres, while others propose they are independent of each other.[13–16] With all

these aggregation conformers observed in diseases, it is imperative to understand their mechanism of formation.

With so many different aggregate structures, and some on linked pathways, while others are on independent pathways, it is difficult to elucidate a mechanism of formation. More importantly, it is unclear what role different protein conformations play in aggregate formation. Many studies find the misfolded structure important for initiation[17–20], but additional studies have found native-like structures in cellular aggregates.[21–23] Additionally, aggregation propensity has been linked to the proportion of globally unfolded species[13,14,24–27], and interestingly the unfolded species may be vital for protein templating in prion-like aggregation.[28,29] Understanding the role each species serves in protein aggregation will be critical to deciphering the role of protein aggregation in disease.

There are numerous protein aggregation diseases, many being neurodegenerative disorders with different key proteins involved in aggregation. For Alzheimer's disease, amyloid fibrils consist of predominantly amyloid-ß (Aß) peptides[30], whereas Parkinson's disease involves Lewy bodies formed from $\alpha$-synuclein.[31] Additionally, many diseases like amyotrophic lateral sclerosis (ALS) have both sporadic (sALS) and familial subtypes (fALS); fALS has multiple proteins associated with its cellular aggregates.[32–35] TAR DNA-binding protein 43 (TDP43), RNA-binding protein FUS and Cu,Zn-superoxide dismutase 1 (SOD1).[35–38] More specifically, there are over 180 different mutations in SOD1 associated with ALS patients, and to convolute things further, SOD1 mutants are associated with different disease characteristics, notably duration. Collectively, understanding of the aggregation-relevant species becomes more convoluted with the

key aggregating proteins varying between diseases, and introducing point mutations in the key proteins adds to the complexity.

## 1.2 Cu,Zn Superoxide Dismutase 1 (SOD1)

In humans, there are three types of superoxide dismutases (SOD): SOD1 is located mostly in the cytoplasm, SOD2 is in the mitochondria and SOD3 is extracellular.[39] SOD1 is the most abundant form, and is responsible for catalyzing the dismutation of the superoxide radical. SOD1 is a 32 kDa homodimeric metalloenzyme (Figure 1.2); the dimer is held together by predominantly hydrophobic interactions.[40] Each monomer has a ß-barrel arranged in a Greek key motif.[41] Moreover, each monomer binds a copper (Cu) and zinc ion (Zn), with the former serving a catalytic role, while the latter is structural.[42] Additionally, the two long loops, loops 4 and 7, are involved in forming the active site and parts of the dimer interface. Loop 4 contains residues responsible for metal-binding and is thus appropriately named the metal-binding loop. Loop 7, also named the electrostatic loop, contains a short helix and one of the cysteines responsible for forming the intrasubunit disulfide bond. Collectively, the metal ions, disulfide bond and dimerization are important post-translational maturation events critical to forming a stable and functional SOD1.

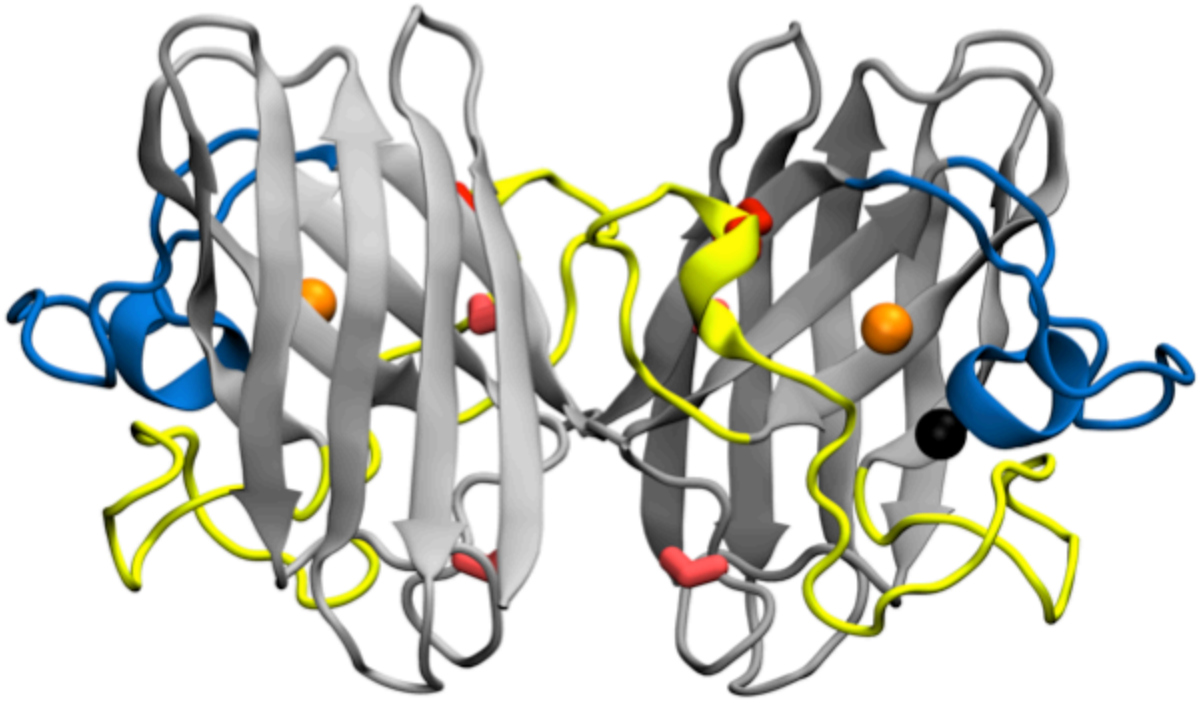


Figure 1.2: Ribbon structure of holoSS superoxide dismutase 1 (PDB: 1HL5, rendered with PyMol). Holo SOD1 consists of two 8-stranded Greek key beta-barrel monomers, with one zinc ion (black sphere) and one copper ion (orange sphere) per monomer. The red represents the Cys57-Cys146 disulfide bond and two free cysteines at positions 6 and 111, and loop IV (yellow) contributes to the dimer interface and zinc binding and loop VII (blue) is the electrostatic loop. Adapted from Broom *et al*., 2014.

The CCS dimerizing to immature SOD1 on the same interface SOD1 forms a homodimer. For the CCS to be able to bind to immature SOD1, it was originally thought the zinc ion must be bound prior to the CCS:SOD1 dimerization. More recently, there have been findings suggesting the CCS might also help in zinc transfer in addition to the copper transfer and disulfide formation.[43] The dimer interface affinity is weak in the most immature form (metal-free and disulfide bond reduced) which results in a predominately monomeric state. However, upon zinc addition, the dimer becomes more favourable.[44,45] Considering the CCS binds at the homodimer interface to form a heterodimer, a SOD1 state that is more permissive to dimerization is also a likely candidate to form a heterodimer with the CCS. With that in mind, the conflicting data centered around the

SOD1 state responsible for CCS binding prove more research is necessary to understand SOD1 maturation. Upon the addition of the metal cofactors and bond formation, SOD1 can then dimerize to form the full mature Cu,Zn SOD1 (holoSS SOD1).

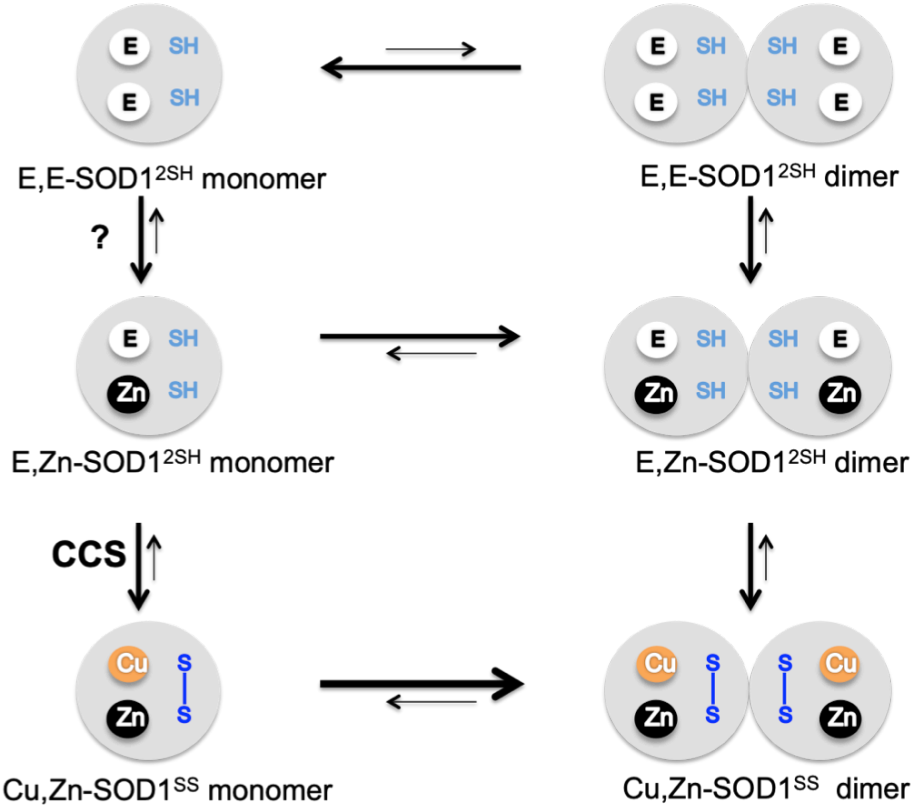



Figure 1.3: Schematic of Cu,Zn SOD1 folding. The most immature state, reduced apo (apoSH) is the initial folding state. The mechanism of zinc incorporation is unknown but is thought to be the initial maturation step. The copper chaperone for SOD1 (CCS) binds to E,Zn SOD1 (E,ZnSH) to deliver the copper and oxidize the conserved disulfide bond between Cys57 and Cys146. Adapted and modified picture from Banci *et al*., 2012.[46]

In addition to the mature enzyme catalyzing the dismutation of superoxide radicals in the cytosol, SOD1 has also been found in the nucleus. Under oxidative stress, SOD1 can relocate to the nucleus and act as a transcription factor that regulates oxidative resistance and repair genes.[47] Our research could not support the binding of any SOD1s to RNA or DNA.[48] However, with so much conflicting literature, and with multiple

maturation states and mutations, it is clear studying SOD1 folding must be done systematically to address all the vital questions. Why are different types of studies, such as SOD1 aggregation studies[49–51], seeing different results? Why do comparable mutations, in terms of location and side chain, in SOD1 cause such different biophysical effects? And most importantly, why do these mutations cause such diverse disease characteristics?

## 1.3 ApoSH SOD1: interplay between mutations and disease is complex

SOD1 requires multiple post-translational modifications to attain its mature native state. Without metal ions and the disulfide bond, the apoSH SOD1 exists predominantly as a monomer. Upon zinc addition, the apoSH SOD1 becomes E,ZnSH SOD1 (where E stands for empty, Zn for zinc-bound, and SH for disulfide-reduced) which is more favourable for dimer formation, whereas with the formation of the disulfide bond, the apoSS SOD1 exists predominantly as a dimer, and with the addition of both metals and disulfide bond, the holoSS SOD1 dimerization is more favourable than in any other state. As more structurally-stabilizing post-translational modifications are added, the thermal stability increases as well. For instance, holoSS SOD1 has a melting temperature ($T_{0.5}$) at 92 °C whereas the apoSH monomer is closer to 48 °C. Therefore, the decreased stability of the more immature forms of SOD1 might explain its implication with disease.

In addition to the lack of post-translational modifications causing marginal stability in apoSH SOD1, the mutations have an added destabilizing impact as well. Mutations in

apoSH SOD1 show a larger range of thermal stabilities,[45] whereas in the case of the mature enzyme, holoSS SOD1, mutants have more comparable thermal stabilities.[52] Therefore, these data suggest a mutation is more impactful in the less stable forms of the protein. Similar to thermal stability being more perturbed in the less mature forms, the dimer interface is also more perturbed by mutations in the more immature states.[53] Collectively, the biophysical properties of immature forms of SOD1 are more sensitive to mutations and are a logical culprit in contributing to ALS.

Due to its sensitivity to ALS-associated mutations and being a sensible culprit in disease, the role of immature forms of SOD1 and ALS-associated mutations in ALS is vital to understand. Previously, Sekhar *et al*. reported differences in the formation of transient conformations across apoSH mutants.[54,55] Their data suggest that apoSH mutants show perturbations in maturation processes and/or the ability to access additional non-native processes the control could not. Similarly, Culik *et al*. also show that the perturbed maturation in apoSH mutants can have consequences on CCS binding.[56] Thus, apoSH SOD1 mutant studies highlight that point mutations can change the protein dynamics such that they affect both maturation and non-native processes.

Furthermore, the more immature forms of SOD1 are also more aggregation prone. In 2022, Naser *et al.* studied in vivo apoSH aggregation and showed significant rescue from aggregation upon the addition of zinc.[57] Additionally, they reported the aggregation propensity roughly correlated with thermal stability, suggesting the unfolded species may be a vital species in aggregation. However, the aggregate structure they observed consists of mostly native-like structures, which hints that the unfolded state may be pulled into the aggregate, but undergoes restructuring upon association. Although these data

suggest the aggregate mechanism is similar for most mutants and the aggregate structures are mostly native-like, other studies propose alternative aggregate structures, which call attention to factors that may modulate aggregation.

Moreover, aggregate structures observed for SOD1 range from native-like to amyloid[21,22] with a wide range of structures among both fibres[58] and oligomers.[49–51] For example, near-native monomers can self-associate to form small oligomers[23] or linear assemblies.[59] Alternatively, different SOD1 peptides adopt a toxic corkscrew oligomer[60] or various steric zipper fibres[60], and unfolded full-length SOD1 forms amyloid structures upon agitation.[24] The partial unfolding of full length SOD1 was also proposed to give rise to the corkscrew oligomer.[60] Other studies have found that the more malleable the protein, the more aggregation-prone conformations are generated[16], giving rise to a higher diversity of aggregate structures.[61–63] These considerations in conjunction with previous findings that apoSH SOD1 has a rugged free energy landscape where the population of mature and non-native conformations is altered by mutation[54,55] strongly support the formation of different aggregate structures by SOD1 variants.

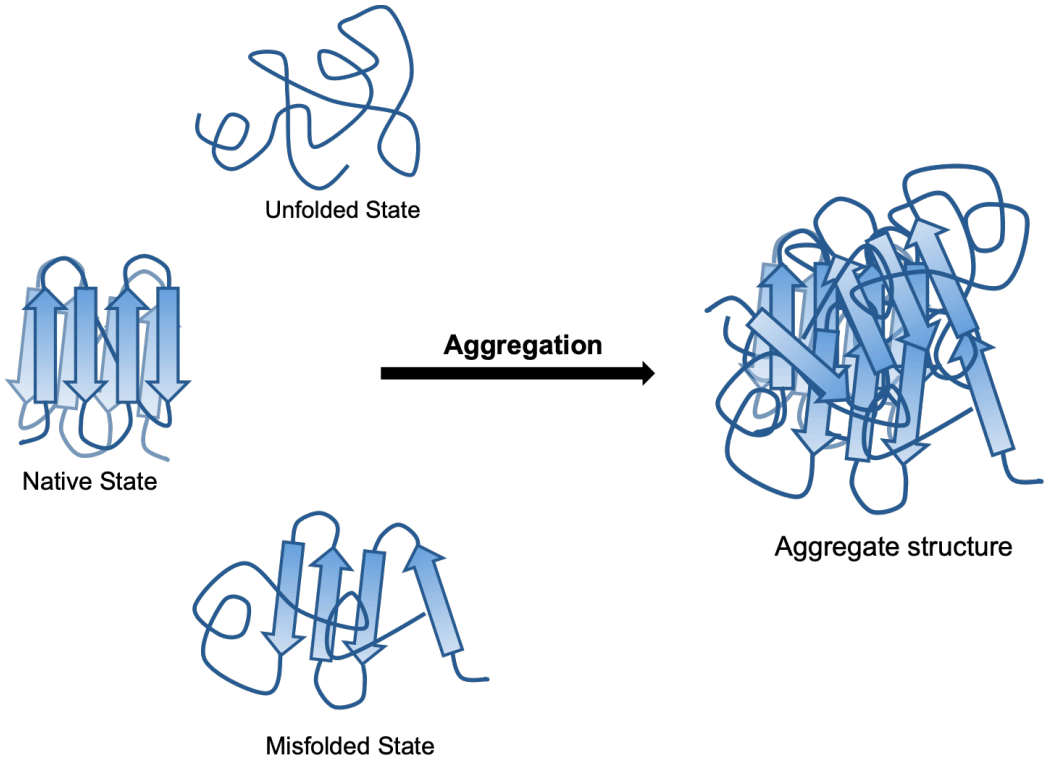



Figure 1.4: Aggregation schematic. The aggregate structure seems to contain a mixture of native-like, partially folded and unfolded states of the protein. Many contrasting studies support either the native structure, misfolded or the unfolded state being the initial aggregating species involved, however the possibility of protein templating further complicates the identification of involved species.

With all these structural polymorphisms, the impact solution conditions have on the aggregation mechanism is important to decipher considering the extent and type of apoSH aggregation in vitro varies greatly with experimental conditions.[21,22,24,64] For example, at physiological temperature and pH, agitation of the metal-free SOD1 promotes the formation of amyloid aggregates[24], whereas non-amyloid aggregates of variable sizes form at different rates for apoSH SOD1 mutants under quiescent conditions.[28,29,65] These and other studies have demonstrated that mutant apoSHs may aggregate in distinct ways, which are sensitive to varying conditions. Intriguingly, structural polymorphisms of apoSH aggregates may be analogous to different conformational “strains” observed in prion-like aggregation.[28,29] Studies of mutant SOD1 ALS mice models support that such

strains are influenced by mutation and contribute to different disease characteristics. Further studies are needed to unravel the mechanisms that may underlie mutant SOD1 aggregate polymorphisms.

## 1.4 Multi-faceted approach to study differences across apoSH mutants

To decipher how mutant SOD1s lead to aggregate polymorphisms, the focus of this thesis is to investigate how ALS-associated mutations in apoSH SOD1 alter the aggregation propensity and structure under different solution conditions and how mutations impact SOD1 structure, dynamics and maturation. For this systematic analysis, we employ multiple biophysical techniques to effectively analyze the impact of different mutations on apoSH SOD1. To allow for appropriate comparability to literature findings and to minimize aberrant self-association, the pseudo wildtype SOD1 (pWT) model is used which involves mutating the two free cysteines (C6 and C111) in the wildtype SOD1 to an alanine and serine respectively. This model has been extensively characterized and shows remarkably comparable biophysical characteristics to the wildtype SOD1, with an additional advantage of minimal observed aggregation throughout protein purification or biophysical characterization.[45,52,65,66]

The pWT SOD1 and its disease-relative mutations are investigated using the following techniques: dynamic light scattering (DLS), atomic force microscopy (AFM) and attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) for

assessing protein aggregation and structural insight, nuclear magnetic resonance (NMR) for both structural and dynamics information, and isothermal titration calorimetry (ITC) with differential scanning calorimetry (DSC) for exploring binding affinity and thermal stability.

Due to the lack of a reliable structure for the apoSH SOD1, most of the analyses in this work are imaged on the holoSS SOD1 structure (PDB: 1HL5). Without the metals and the intrasubunit disulfide bond, the metal-binding loop and the electrostatic loop are largely disordered for the predominately monomeric apoSH SOD1.[54,55,67] Although there is a crystal structure available for a metal-free and cysteine-free SOD1 (PDB: 2GBU)[67] which could be considered similar to apoSH SOD1, this structure shows unexpected differences to the solution-based NMR data.[54,55] The 2GBU structure shows structuring similar to the holoSS SOD1, which consists of ordered loops in a dimeric state. Therefore, it is likely this structure is not an accurate representation of the apoSH SOD1 state, and further highlights the necessity for structural information on apoSH SOD1.

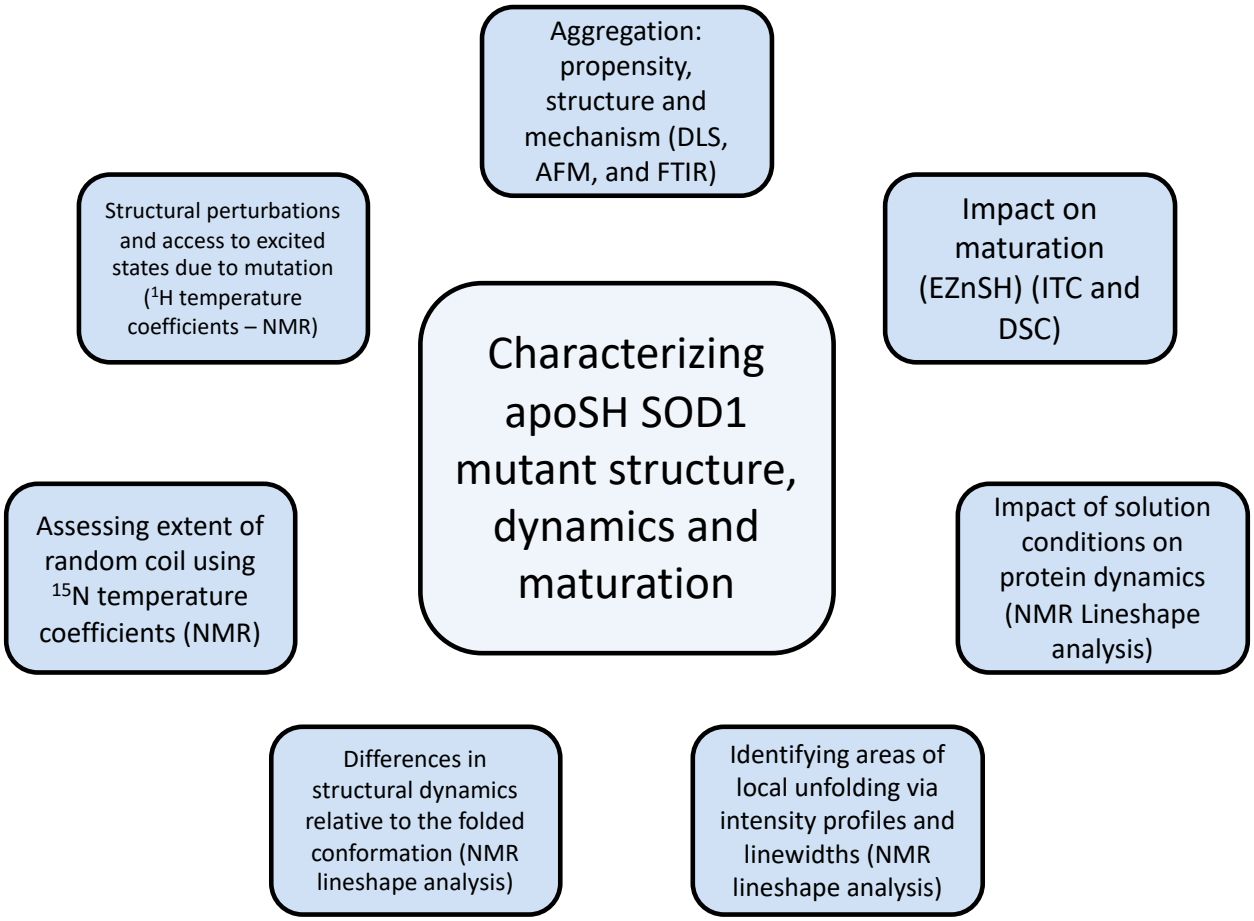



Figure 1.5: Approach to characterizing apoSH SOD1 presented in this thesis. The plan to characterize different aspects of the apoSH SOD1 state and its mutants involve many biophysical techniques (indicated in brackets).

### 1.4.1 Dynamic Light Scattering (DLS)

We used DLS to obtain a sensitive analysis of the presence of oligomers and aggregates different sized species in samples of variant SOD1. Light scattering is a particularly sensitive method for detecting aggregated species, as the light scattering intensity is proportional to the $D_h^6$, where $D_h$ is the hydrodynamic diameter.[68] Typically, DLS results are reported as distributions of particle hydrodynamic diameters, $D_h$, based on the associated intensity of scattered light ($D_{h,I}$). In addition to the $D_{h,I}$, the particle number distribution ($D_{h,N}$) is also useful because it shows the number or relative

concentration of different size species.[68] Intensity distributions from light scattering report the amount of light scattered as a function of particle diameter, hence will always emphasize larger particles, and such distributions can be misleading when trying to assess extent of aggregation. Number distributions, however, show particle abundance according to their concentration in solution, and thus provide a useful alternative measure of aggregation propensity. Number distributions are underutilized and here by considering both distributions offers insight into apoSH aggregate sizes as well as their abundance, and ultimately a deeper insight into self-association of SOD1 variants.

A diverse set of mutations that vary in their structural features, thermal stability, and extent of aggregation are examined in Chapter 2. Mutants with $T_{0.5}$ close to pWT and above physiological temperature are considered stabilized, while destabilized mutants have a $T_{0.5}$ below pWT and physiological temperature.[65] The destabilized mutants have an increased fraction of unfolded protein compared to pWT, which will increase the average measured $D_h$. Based on empirical calculations using the polypeptide chain length[69], the $D_h$ of folded and unfolded apoSH SOD1 is ~4.2 and 7.8 nm, respectively. Thus, in addition to the extent of aggregation, both $D_{h,I}$ and $D_{h,N}$ can reveal shifts in the fraction of unfolded protein.

### 1.4.2 Atomic Force Microscopy (AFM)

Atomic force microscopy (AFM) involves the use of a sharp probing tip to measure the surface of a sample.[70] Interactions between the tip and surface are analyzed to generate high-resolution images of the sample, which are useful for assessing aggregate structural information. The images are sensitive enough to decipher between aggregate

structures and sizes. In Chapter 2, we discuss extending typical AFM aggregate size analysis to quantify the number of monomers in aggregates.[71] Additionally, we demonstrate violin-box plots as a useful tool that augments standard box plot summaries of AFM particle statistics with additional information on data distribution. The data distributions, in addition to the other size estimates, are valuable for discerning subtle aggregate differences between variants and their aggregation behaviour.

### 1.4.3 Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR)

ATR-FTIR is a technique that produces an infrared spectrum of either the emission or absorption of a sample, and can be useful for assessing the relative secondary structure of proteins.[72] In Chapter 2, the secondary structure of apoSH SOD1s during aggregation is assessed by ATR-FTIR spectroscopy. Protein structural components show absorbance bands in different spectral ranges, such as: β-sheet (~1623-1641 $cm^{-1}$); unstructured/helix (1642-1657 $cm^{-1}$); and turns (1662-1686 $cm^{-1}$). We also obtain control spectra of folded and heat denatured apoSH SOD1, as well as sonicated apoSH SOD1 as a measure for amyloid.[73] These controls are used for gauging the extent of native-like structure in the aggregates. Additionally, the ATR-FTIR spectra are collected for the folded samples under different solution conditions, and the spectral change between solution conditions use to investigate the impact on protein structure prior to aggregation. Therefore, ATR-FTIR is sensitive to changes in secondary structure both upon aggregation and due to different solution conditions.

### 1.4.4 Nuclear Magnetic Resonance (NMR)

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool that takes advantage of the magnetic properties of atoms, and allow for determining the structure, dynamics and interactions of molecules.[74–77] A 1D NMR spectrum has peaks along one horizontal axis, which represent the chemical shift for each nucleus. However, this spectrum is limited to smaller proteins since there are more and broader peaks for larger proteins which may overlap.[76] Multidimensional NMR resolves the overlapping peaks by separating them by another dimension. 2D NMR experiments are commonly used to study protein structure and dynamics. These experiments can either be a homonuclear spectrum, where $^{1}H$ frequencies correlated with other $^{1}H$ frequencies or a heteronuclear spectrum with $^{1}H$ typically correlated with either $^{15}N$, $^{13}C$ or less commonly, $^{31}P$. A $^{1}H$-$^{15}N$ heteronuclear single quantum correlation (HSQC) spectrum provides a $^{1}H$-$^{15}N$ cross peak in the protein, which can arise from backbone amides or sidechain amines. As each amino acid residue (except proline) contains an amide group, this results in a spectrum with a single peak for each amino acid in the protein, the location of which is determined by the chemical environment of each amide.

Each NMR peak consists of a chemical shift, peak intensity and linewidth, all reflecting the local chemical environment, population and local dynamics, respectively, of the corresponding residue.[78] Chemical shift represents the resonant frequency of a nucleus relative to a standard in the same magnetic field. The value is independent of the magnetic field, and instead reports on small local magnetic fields generated by the

electrons around the nuclei. However, since chemical shifts are sensitive to many things such as conformational exchange, environmental changes, and presence or absence of local structure, chemicals can be sensitive reporters of subtle changes in chemical environment. Moreover, if a residue is able to access two conformations, depending on their rate of exchange, the spectrum can show two different chemical shifts corresponding to the two different conformations in slow exchange. However, for two conformations in fast exchange, the spectrum will show one peak at a chemical shift value that is a weighted averaged of the two chemical shifts without exchange.

In addition to the chemical shifts, the peak lineshape can also be a useful reporter of local dynamics such as conformational exchange. Peak intensity reports on the number of nuclei resonating at a given frequency, whereas linewidths change based on relaxation rates. However, both peak intensity and linewidth are influenced by additional dynamic processes, such as a residue exchanging between two distinct chemical environments. NMR lineshape analysis is a useful tool commonly used for assessing binding of ligands to a biomolecule.[79–81] It can help assess thermodynamics, kinetics and structural information of binding events. In addition to these processes, lineshape analysis has been explored for comparing protein variants for differences in dynamics. [82] Unfortunately, lineshape analysis can be convoluted due to the numerous factors that impact peak shape and position; two of these key factors include exchange rate between processes and relaxation rate.[78,83–85] Processes such as ligand binding that may be simplified to exclusively a bound and unbound state are easier to interpret. However, assessing

lineshape changes throughout multiple spectra where processes such as dimerization, aggregation and/or population of transient conformations are occurring convolute the analysis greatly.

For apoSH SOD1 mutants, population of both non-native and native dimers, aggregates, and transient conformations are possible along the temperature ramp.[12,54,56,65] Aside from these processes convoluting the sample, each of these species impact peak shape in different ways. Dimers, aggregates and transient conformations will have their own respective exchange rates with the monomer, whether that is the folded or unfolded monomer is an additional complication, as well as different relaxation rates. Impacts to exchange rates and relaxation rates will subsequently influence peak position, intensity and linewidths. Thus, with these potential complications, our aim is to attempt to develop a method to analyze peak shape relative to a reference where conclusions are only drawn with multiple residues behaving comparably as opposed to isolated residues.

By collecting an NMR spectrum at a series of temperatures, the temperature coefficients of both proton and nitrogen chemical shift can be measured (both amide proton and nitrogen chemical shifts) (Chapters 3, 5, and 6), and the chemical shift perturbations relative to pWT throughout the temperature range (Chapters 3 and 6). These data were obtained for multiple SOD1 mutants and under different solution conditions to assess the changes in local dynamics and structure (Chapters 3-6). By collecting spectra in the folded and unfolded temperature ranges, the dynamics between

the folded and unfolded temperature ranges are assessed using peak shape changes. Collectively, these data detailed insights on the subtle variability each SOD1 mutant has which in turn can explain the aggregate polymorphisms observed for SOD1.

### 1.4.5 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is traditionally used to assess binding affinities between a ligand and a protein by titrating the ligand into a cell containing the protein.[86] The heats of binding are fit to a binding model to determine key association parameters such as binding affinity. However, this technique can be extended to apply to any binding partners if the binding affinity is within the measurable range of the respective instrument used. In addition to the traditional titration of a binding party into a cell that contains the binding partner, the method of titration can be manipulated to assess dissociation instead of association. In Chapter 7, we take a highly concentrated sample of EZnSH SOD1 and measure heats of its dimer dissociation to assess the strength of the dimer for multiple mutants. Since this state is considered to be an important initial maturation state and the dimer interface is used by the CCS to finish maturation, a perturbed dimer interface is believed to have a detrimental impact on maturation.

### 1.4.6 Differential Scanning Calorimetry (DSC)

DSC is often employed to characterize the thermal stability of the protein and its respective mutants. It involves heating a cell containing protein against a reference cell. The DSC instrument detects the difference in heating between the cells throughout the

temperature ramp; these changes in heating are represented as an unfolding endotherm which can be fit to unfolding models to characterize the thermodynamics associated with protein unfolding. Previously in our lab, DSC was used to assess the thermal stability of holoSS-SOD1, apoSS-SOD1, and apoSH-SOD1.[45,52,65,87] In Chapter 7, the thermal stability of EZnSH-SOD1 is evaluated. However, the SOD1 thermograms in its many states highlight the complexities that arise upon protein unfolding. In most cases, protein unfolding is considered 2-state unfolding between the folded state and unfolded state, however, certain states and mutants show the ability to undergo 3-state unfolding. For apoSS-SOD1 unfolding, a monomeric intermediate is populated for proteins that form weak dimers, which then unfolds. However, for those that form a strong dimer, the dimer unfolds into the unfolded state without a detectable intermediate. Interestingly, for EZnSH-SOD1, the DSC thermograms suggest very different behaviour than observed for the other states.

## 1.5 Thesis Summary

With apoSH SOD1 mutants showing such diversity in both protein characteristics and aggregate structures, it is critical to understand how this state and its mutants impact protein folding and dynamics. In this thesis, I present characterization of the apoSH SOD1 state and ALS-associated SOD1 mutations using complementary biophysical methods. The objective of these studies was to explore aggregate polymorphisms under different solution conditions, apoSH SOD1 and its mutant structures and dynamics, as well as maturation of apoSH SOD1 and SOD1 mutants. The overarching goal is to understand how apoSH SOD1 and ALS mutations impact structure, dynamics, aggregation and

maturation, to better understand their role in disease. In addition to this research contributing to understanding the roles of SOD1 in aggregation and disease, it is also useful for helping to understand the role point mutations and solution conditions have on protein folding in general.

# Chapter 2: Immature ALS-associated mutant superoxide dismutases form variable aggregate structures through distinct oligomerization processes

Statement of contributions:

This chapter consists of a manuscript which has been accepted for publication at the time of the submission of this thesis. Exceptions to sole authorship of material are stated below. The manuscript in this chapter is as follows:

Harmeen K. Deol, Helen R. Broom, Bruna Siebeneichler, Brenda Lee, Zoya Leonenko, Elizabeth M. Meiering. Immature ALS-associated mutant superoxide dismutases form variable aggregate structures through distinct oligomerization processes. Biophysical Chemistry, Volume 288, 2022, 106844, ISSN 0301-4622, https://doi.org/10.1016/j.bpc.2022.106844.

Author contributions:

HRAB: DLS and AFM conceptualization, methodology, validation, formal analysis, investigation, data curation and visualization, resources, overall conceptualization, writing-original draft; BL: AFM methodology, validation, investigation, formal analysis, writing – review & editing; ZL: AFM methodology, validation, formal analysis, resources, writing – review & editing; BS: ATR-FTIR validation, formal analysis, investigation, visualization, writing-review & editing; HKD: ATR-FTIR conceptualization, methodology, validation, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft, writing- review & editing; EMM:

conceptualization, writing- review & editing, supervision, project administration, funding acquisition.

## 2.0 Context

Protein misfolding and aggregation are hallmarks of many diseases, including amyotrophic lateral sclerosis (ALS). In familial ALS, aberrant self-association of mutant Cu,Zn-superoxide dismutase (SOD1) is implicated as a key contributor to disease. Mutations have the largest impacts on the stability of the most immature form of SOD1, the unmetallated, disulfide-reduced monomer (apoSH SOD1). Here we demonstrate that, despite the marginal stability of apoSH SOD1, aggregation is little correlated with the degree of protein unfolding, and multiple modes of aggregation occur, depending on the mutation and solution conditions. Light scattering and atomic force microscopy reveal two distinct mutant SOD1 behaviours: high aggregator mutants form abundant small assemblies, while low aggregator mutants form fewer, more fibre-like aggregates. Attenuated total reflectance-Fourier transform infrared spectroscopy and Thioflavin T binding show the aggregates maintain native-like anti-parallel beta structure. These results provide new evidence that ALS-associated mutations promote the aggregation of apoSH SOD1 through multiple pathways, with broad implications for understanding mechanisms of protein self-association in disease and biotechnology.

## 2.1 Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative syndrome characterized by the progressive loss of nerve cells in the brain and spinal cord.[88] A hallmark feature of ALS is the presence of protein aggregates that accumulate in motor neurons.[89] A prominent disease hypothesis is that ALS and many other neurodegenerative diseases arise from the misfolding and self-association of disease-associated proteins which have cytotoxic effects.[90] One of the major contributors to familial ALS is mutations in superoxide dismutase 1 (SOD1), which generally promote protein misfolding and aggregation.[91–93] The widely varying biochemical characteristics and energy landscapes of mutant SOD1s have been implicated in giving rise to different disease characteristics, yet the underlying molecular mechanisms remain obscure.[94,95] Thus, elucidating the mechanisms of SOD1 aggregation is of broad interest for understanding neurodegeneration as well as for controlling aggregation in the development of proteins for therapeutics and biotechnology.[96,97]

The most immature form of SOD1 lacks both metal cofactors and the disulfide bond (apoSH) and is of particular interest owing to its marginal stability and high aggregation propensity.[94,98–102] Moreover, apoSH SOD1 is generally more destabilized by mutations compared to other forms of SOD1.[94] The thermal unfolding transition of wild-type (WT) and pseudo wild-type (pWT) apoSH begins slightly above the physiological temperature of 37 °C and is typically decreased by disease mutations. Consequently, the relative population of partially folded and unfolded monomers generally significantly increase upon mutation. With this global destabilization and the accompanying increase

in local conformational dynamics, apoSH aggregation is a logical culprit in contributing to ALS and so is of much interest to investigate.

The extent and type of apoSH aggregation in vitro has been reported to vary greatly with experimental conditions.[103–106] For example, at physiological temperature and pH, agitation of the metal-free SOD1 promotes the formation of amyloid aggregates[104], whereas non-amyloid aggregates of variable sizes form at different rates for apoSH mutant SOD1s under quiescent conditions.[94] These and other studies have demonstrated that mutant SOD1s may aggregate in distinct ways, which are sensitive to varying conditions. Notably, structural polymorphisms of apoSH aggregates may be analogous to different conformational "strains" observed in prion-like aggregation.[107,108] Studies of mutant SOD1 ALS mice models support that such "strains" are influenced by mutation and contribute to different disease characteristics. Further studies are yet needed to unravel the mechanisms that may underlie mutant SOD1 aggregate polymorphisms.

Here, we characterize apoSH SOD1 aggregation propensity and structure by light scattering, atomic force microscopy (AFM), attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) and Thioflavin-T (ThT) binding. A diverse set of familial ALS-associated mutant SOD1s are analyzed under different protein and salt concentrations. We find the mutants can be classified as high aggregators, which mainly show rapid abundant formation of oligomers and native-like fibrils, or low aggregators, which mainly show slow formation of low concentrations of fibres. The results illuminate how different point mutations and solution conditions can impact multi-modal aggregation of SOD1, which has significant implications for understanding and controlling such aggregation.

## 2.2 Results

### 2.2.1 DLS in different solution conditions reveals distinct aggregation propensities of mutant apoSHs

Dynamic light scattering (DLS) measurements were performed at physiologically relevant temperature (37°C) and pH (7.4) to obtain a sensitive analysis of the extent of formation of self-associated species by mutant SOD1s at 63 μM protein concentration. Typically, DLS results are reported as distributions of particle hydrodynamic diameters, $D_h$, based on the associated intensity of scattered light ($D_{h,I}$). Here we assess also the particle number distributions ($D_{h,N}$), which show the number or relative concentration of different size species.[109] While intensity distributions are useful for high sensitivity detection of the formation of larger species, owing to the intensity being proportional to $D_h^6$, this measure can overlook more abundant smaller species which are better captured by the number distribution. We find considering both distributions offers deeper insight into the self-association of SOD1 variants.

We examine a diverse set of mutations that vary in their structural features, thermal stability, and extent of aggregation. Mutants with a thermal melting temperature ($T_{0.5}$) close to pWT and above physiological temperature are referred to here as stabilized, while destabilized mutants have a $T_{0.5}$ below pWT and physiological temperature[94] (Table 2.1). The destabilized mutants have an increased fraction of unfolded protein compared to pWT, which will increase the average measured $D_h$. Based on empirical calculations using the polypeptide chain length[110], the $D_h$ of folded and unfolded apoSH SOD1 is ~4.2 and 7.8 nm, respectively. Consistent with the calculated $D_h$ values, at 0 hours ($t_0$) the

destabilized mutants exhibit higher values of $D_{h,I}$ (Figure 2.1A-D, Table 2.2) and $D_{h,N}$ (Figure 2.1B) relative to pWT and the stabilized mutants. Thus, both $D_{h,I}$ and $D_{h,N}$ reveal shifts in the fraction of unfolded protein.

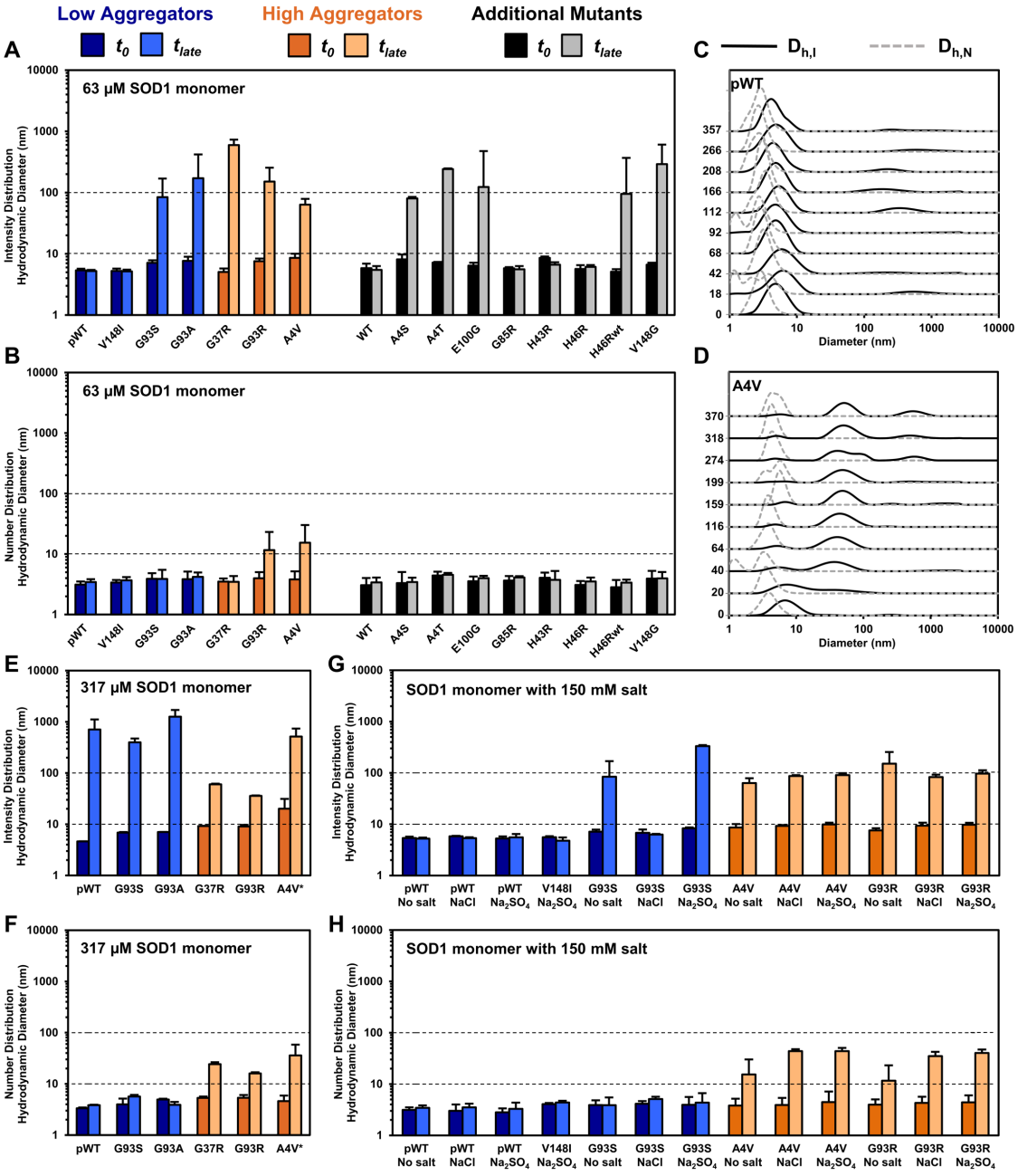



Figure 2.1: Aggregation of SOD1 variants measured by DLS hydrodynamic diameter intensity and number distributions at different protein and salt concentrations. Samples at pH 7.4, 37 °C contained (A-D) 63 μM SOD1 monomer, (E,F) 317 μM SOD1 monomer, and (G,H) 63 μM SOD1 monomer (except V148I, at

317uM) with 150 mM salt. The hydrodynamic diameters (A,B,E-H) correspond to the main peak in the distributions (C,D and Figures 2.2-2.4). (C,D) DLS measurements were taken approximately every 24 hours, with the first time point ($t_0$) at the bottom of each panel and the last time point ($t_{late}$), at the top. (A,B,E-H) Hydrodynamic diameters are averages of 3 measurements for a sample, with standard deviations for the measurements shown by error bars. For each variant, pairs of values correspond to the initial measurement ($t_0$ represented by dark blue, orange, or black bar) and after incubation of the sample for ~300 hours ($t_{late}$ represented by light blue, orange, or gray bar). Dashed lines at 10 nm and 100 nm are to guide the eye. Mutants were grouped based on the aggregation results for the mutants at 63 μM; the low aggregators are shown in blue, while the high aggregators are shown in orange; additional mutants that are not analyzed in further detail are in grey and black. Intensity distribution hydrodynamic diameter values (black solid line) are highly sensitive to large particles, while number distribution diameter values (gray dashed line) represent the dominant population of species in solution.
*A4V showed pronounced aggregation at high protein concentration, leading to the formation of aggregates visible by the eye. Such aggregates diffuse too slowly to be measurable by DLS, thus the sample could only be measured for ~120 hours.

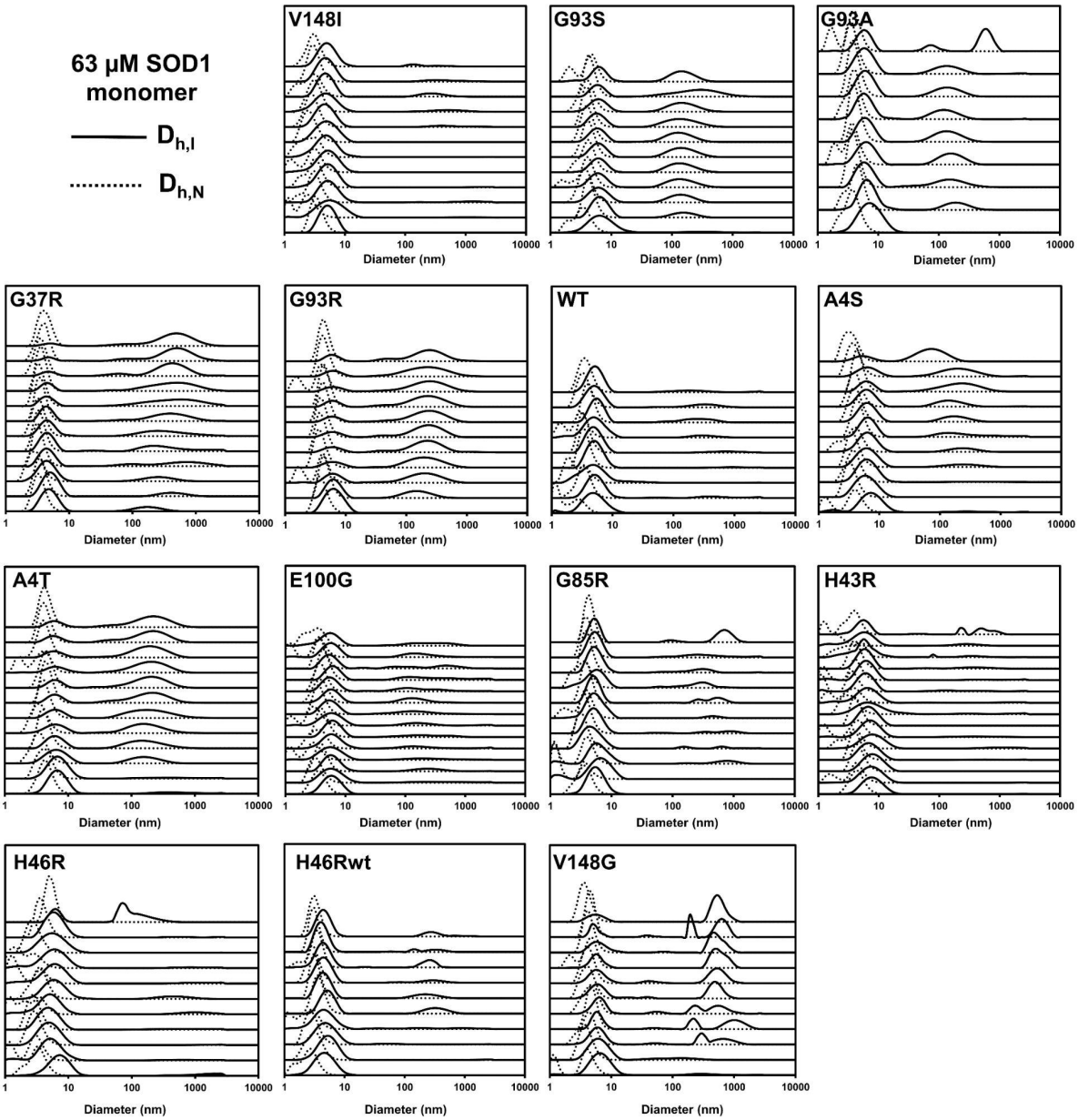



Figure 2.2. DLS shows that apoSH variants aggregate little at 63 μM SOD1 monomer. All samples were incubated in 1 mM TCEP, 20 mM HEPES. Intensity distributions are indicated by the solid lines and number distributions by the dotted lines. DLS measurements were taken approximately every 24 hours, with $t_0$ at the bottom of each panel and the last time point, greater than 300 hours, at the top. Intensity distribution hydrodynamic diameter values (solid line) are sensitive to large

particles, while number distribution diameter values (dotted line) represent the dominant population of species in solution.

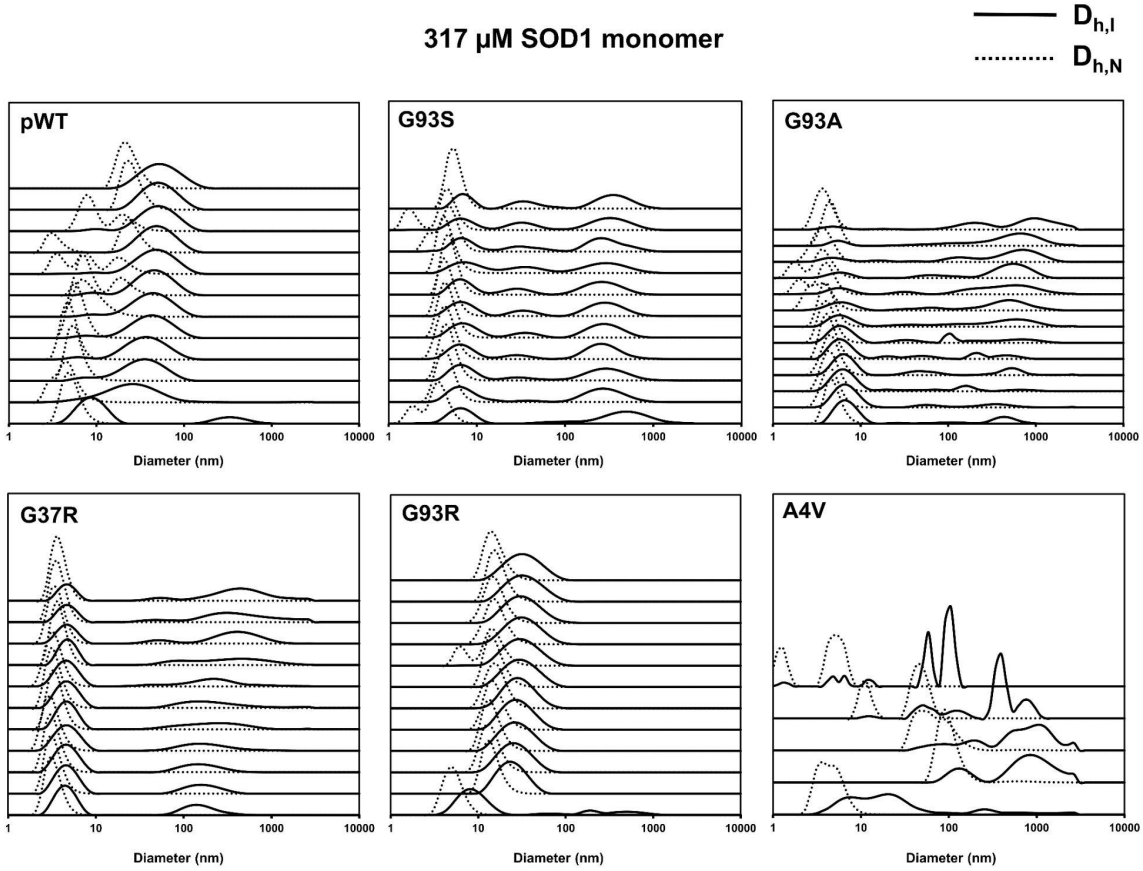



Figure 2.3. DLS also reveals different levels of aggregation for apoSH variants incubated at high concentration (317 μM). All samples were incubated in 1 mM TCEP, 20 mM HEPES. At increased protein concentration A4V shows the highest aggregation propensity. At $t_0$ there is a large discrepancy in the size reported by the intensity and number distribution, suggesting that aggregates form immediately. After ~24 hours most of the sample formed aggregates of similar size to those observed at 63 μM; however, larger aggregates are also apparent. Subsequent measurements revealed that a significant amount of protein had formed aggregates too large to be detected by DLS, resulting in poor estimates of particle size. After ~120 hours the sample was not measurable by DLS. G37R and G93R showed much high levels of aggregation at high concentration, forming aggregates similar in size to those formed by A4V and 63 μM as indicated by the shift in intensity and number distributions to larger values.

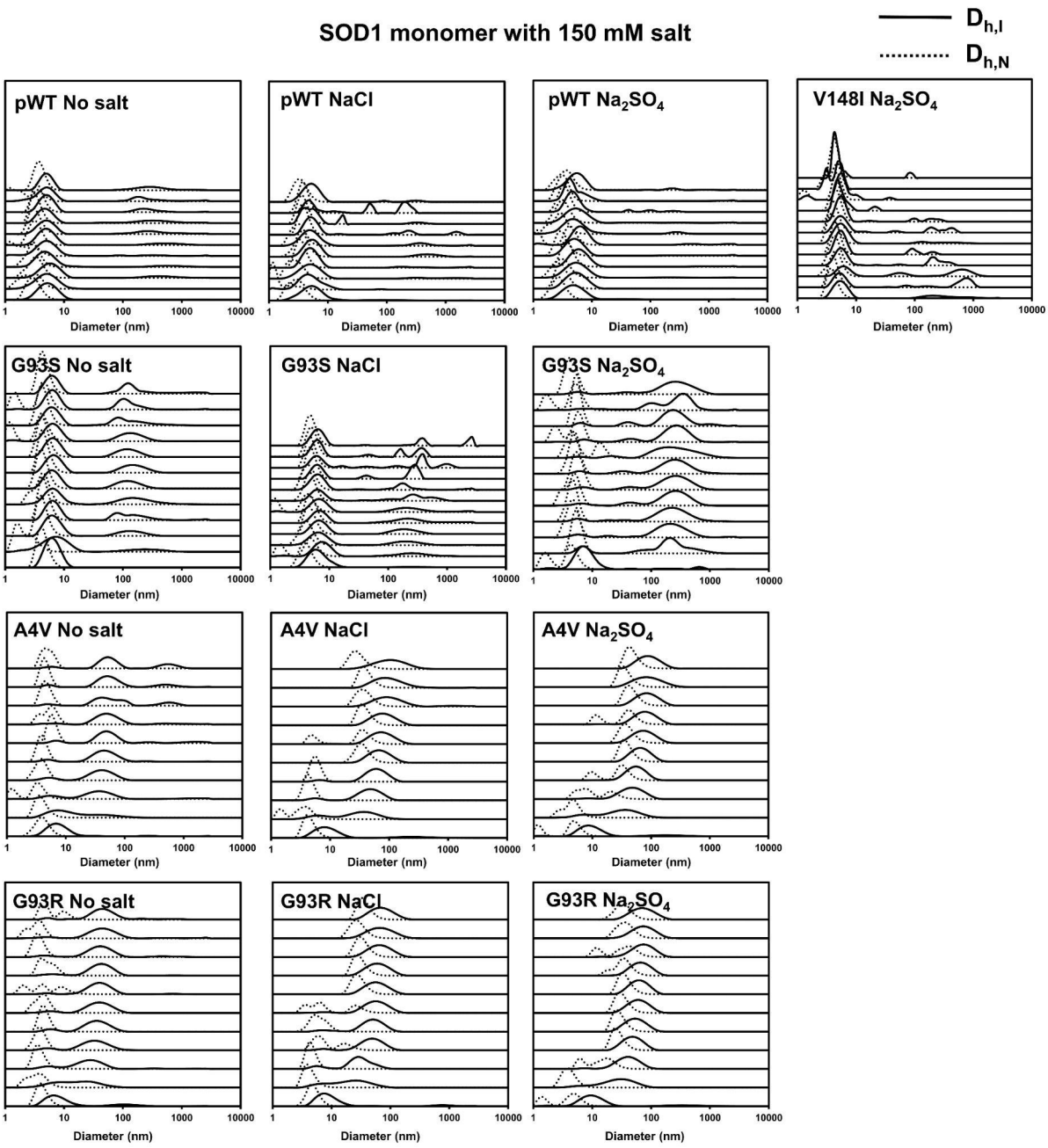



Figure 2.4. DLS experiments indicate that $Na_2SO_4$ and NaCl promote the aggregate formation, most notably for A4V and G93R mutants. The effect of salt on the size of particles in solution is shown for mutants that displayed little increase in light scattering intensity with incubation time. For these apoSH variants, DLS indicates that very low levels of aggregation are observed under all conditions. The effect of salt on the size of particles in solution is shown for mutants that displayed an increase in light scattering intensity with incubation time. For A4V and G93R, aggregation is enhanced in 150 mM $Na_2SO_4$ and NaCl as indicated by the shift in intensity (solid lines) and number distributions (dotted lines) to larger values. While G93S showed an increase in light scattering intensity with incubation time in 150 mM $Na_2SO_4$, DLS data indicate that the aggregates formed never become the most abundant species. At t0 there is a difference in the size reported by the intensity and

number distribution for all samples, which suggests that the samples exhibit some degree of heterogeneity.

Table 2.1. Apparent melting temperatures of mutant apoSH SOD1s.

| Stabilized mutants | | | Destabilized mutants | | |
|---|---|---|---|---|---|
| **ApoSH SOD1** | **$T_{0.5}$[a] (°C)** | **Disease Duration (years)[c]** | **ApoSH SOD1** | **$T_{0.5}$[a] (°C)** | **Disease Duration (years)[c]** |
| pWT | 48.5 ± 0.3 | n/a | A4V | 36.3 ± 0.2 | 1.5 |
| V148I | 50.0 ± 0.1 | 1.7 | H43R | 35.4 ± 0.4 | 1.8 |
| H46R | 52.8 ± 0.4 | 17.6 | G93R | 35.4 ± 1.8 | 5.3 |
| G85R | 41.2 ± 0.3 | 6.0 | E100G | 35.3 ± 0.7 | 4.7 |
| | | | G93A | 34.6[b] | 3.1 |
| | | | G37R | 33.5 ± 1.2 | 17.0 |
| | | | G93S | 33.4 ± 1.2 | 8.0 |
| | | | A4S | 32.5[b] | n/a |
| | | | A4T | 30.9 ± 0.3 | 1.5 |

[a]Values are the apparent melting temperature ($T_{0.5}$) where the observed $C_p$ in the differential scanning calorimetry (DSC) scan is a maximum. Values are from Vassall *et al*., "Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS," Proc. Natl. Acad. Sci. U. S. A., vol. 108, no. 6, pp. 2210–2215, 2011, doi: 10.1073/pnas.0913021108.

[b] $T_{0.5}$ based on one measurement

[c]Values are from Q. Wang *et al*., "Protein aggregation and protein instability govern familial amyotrophic lateral sclerosis patient survival," PLoS Biol. Vol. 6, no.7, e170, doi: 10.1371/journal.pbio.0060170

There is minimal change in both $D_h$ distributions upon sample incubation for ~300 hours ($t_{late}$), indicating that under these conditions most SOD1 variants aggregate only to a low level and the size of the species in solution changes little with time (Figure 2.1A,B). The notable exception is A4V, where $D_{h,I}$ and $D_{h,N}$ values increase with time (Figure 2.1A,B). While the average $D_{h,N}$ after ~300 hours of incubation has high uncertainty because the number of aggregates is too low to be well defined, the average $D_{h,I}$ clearly shows larger species form with time (Figure 2.1B).

To promote aggregation, we also performed light scattering experiments at 5-fold higher protein concentration (317 μM). In these experiments, we further define the behavior of the pWT control and destabilized mutants that exhibited some aggregation at 63 μM (i.e.

A4V, G37R, G93A, G93R, and G93S). After ~300 hours of incubation ($t_{late}$), the high concentration samples show varying increases in $D_{h,I}$ and $D_{h,N}$ relative to $t_0$, suggesting differences between SOD1 variants in the size of species formed (Figure 2.1E,F). The mutants can be classified according to their aggregation characteristics. $D_{h,I}$ and $D_{h,N}$ differ the most for pWT, G93A, and G93S, henceforth referred to as the low aggregators, with differences in particle diameter ranging from ~400-1200 nm evidenced by increases in $D_{h,I}$ but not $D_{h,N}$ (blue bars in Figure 2.1E,F, Table 2.3). In addition, the minimal aggregation for the low aggregator mutants is further shown by the limited change in polydispersity index (PDI) values (Table 2.2-2.3), indicating no significant change in sample heterogeneity (Table 2.3).

Table 2.2. DLS data summary for mutant apoSH SOD1s at 63 μM, pH 7.4, 37 °C.

| | **t = 0** | | | | **t > 300** | | | |
|---|---|---|---|---|---|---|---|---|
| ApoSH SOD1 variant | PDI[a] | PDI width[b] (nm) | Intensity Distribution Hydrodynamic[c] Diameter (nm) | Number Distribution Hydrodynamic[c] Diameter (nm) | PDI[a] | PDI width[b] (nm) | Intensity Distribution Hydrodynamic[c] Diameter (nm) | Number Distribution Hydrodynamic[c] Diameter (nm) |
| pWT | **0.205** | **2.146** | **5.383** | **3.131** | **0.267** | **5.231** | **5.228** | **3.449** |
| | *0.025* | *0.250* | *0.332* | *0.387* | *0.083* | *5.657* | *0.253* | *0.398* |
| V148I | **0.178** | **1.868** | **5.262** | **3.380** | **0.471** | **94.69** | **5.179** | **3.681** |
| | *0.029* | *0.130* | *0.472* | *0.348* | *0.210* | *110.2* | *0.333* | *0.471* |
| G93S | **0.239** | **3.226** | **7.136** | **3.899** | **0.502** | **30.34** | **84.11** | **3.880** |
| | *0.024* | *0.548* | *0.708* | *0.930* | *0.143* | *46.08* | *85.18* | *1.598* |
| G93A | **0.224** | **3.315** | **7.683** | **3.841** | **0.621** | **122.9** | **171.8** | **4.190** |
| | *0.049* | *0.851* | *1.300* | *1.286* | *0.245* | *142.5* | *248.9* | *0.766* |
| G37R | **0.300** | **48.10** | **5.058** | **3.499** | **0.909** | **149.0** | **596.6** | **3.467** |
| | *0.078* | *109.5* | *0.711* | *0.442* | *0.141* | *34.49* | *134.2* | *0.864* |
| G93R | **0.240** | **3.732** | **7.563** | **3.971** | **0.598** | **26.87** | **151.1** | **11.60** |
| | *0.056* | *1.155* | *0.806* | *1.053* | *0.334* | *9.509* | *104.2* | *11.58* |
| A4V | **0.224** | **5.566** | **8.593** | **3.807** | **0.312** | **29.55** | **63.46** | **15.40** |
| | *0.043* | *9.779* | *1.452* | *1.376* | *0.119* | *10.88* | *15.38* | *14.70* |
| WT | **0.245** | **2.672** | **5.848** | **3.061** | **0.348** | **85.11** | **5.462** | **3.397** |
| | *0.038* | *0.874* | *1.059* | *0.959* | *0.160* | *251.6* | *0.853* | *0.694* |
| A4S | **0.251** | **3.532** | **8.103** | **3.302** | **0.750** | **24.35** | **80.46** | **3.449** |
| | *0.014* | *0.134* | *1.622* | *1.744* | *0.041* | *2.141* | *4.640* | *0.626* |
| A4T | **0.250** | **3.991** | **7.190** | **4.419** | **1.0** | **45.24** | **241.8** | **4.544** |
| | *0.020* | *0.327* | *0.202* | *0.686* | *0.0* | *5.109* | *6.424* | *0.334* |
| E100G | **0.230** | **2.595** | **6.432** | **3.548** | **0.461** | **86.19** | **123.5** | **4.003** |
| | *0.014* | *0.151* | *0.739* | *0.706* | *0.163* | *52.04* | *353.4* | *0.357* |
| G85R | **0.179** | **2.516** | **5.812** | **3.644** | **0.271** | **43.79** | **5.584** | **4.095** |
| | *0.045* | *0.373* | *0.259* | *0.673* | *0.060* | *67.39* | *0.715* | *0.206* |
| H43R | **0.215** | **3.547** | **8.603** | **4.046** | **0.344** | **51.18** | **6.640** | **3.735** |
| | *0.049* | *0.650* | *0.457* | *0.885* | *0.110* | *150.6* | *0.651* | *1.503* |
| H46R | **0.186** | **2.149** | **5.671** | **3.086** | **0.192** | **5.428** | **6.113** | **3.525** |
| | *0.019* | *0.433* | *0.863* | *0.498* | *0.035* | *8.757* | *0.461* | *0.558* |
| H46R$^{wt}$ | **0.227** | **20.13** | **5.109** | **2.825** | **0.496** | **1063** | **95.32** | **3.371** |
| | *0.089* | *37.77* | *0.501* | *0.905* | *0.307* | *2061* | *273.3* | *0.403* |
| V148G | **0.291** | **10.37** | **6.646** | **3.928** | **0.567** | **191.1** | **291.5** | **3.955** |
| | *0.117* | *13.20* | *0.522* | *1.343* | *0.138* | *197.1* | *316.3* | *1.083* |

Bold and italicized values are the averages and standard deviations, respectively, resulting from multiple experiments. [a]PDI is obtained using the Cumulants method of fitting the autocorrelation function (Eq. 4). PDI is an assessment of sample heterogeneity. Values larger than 0.1 are indicative of polydispersity (i.e., the sample contains more than one species in solution). High PDI values result in larger peak widths, and thus greater distribution of sizes reported. [b]Peak width reflects the range of sizes obtained from the intensity distribution and is related to the PDI according to Eq. 4. [c]The sizes reported correspond to the main peak in the distributions.

Table 2.3. DLS data summary for mutant apoSH SOD1s at 317 μM, pH 7.4, 37 °C.

| | t = 0 | | | | t > 300 | | | |
|---|---|---|---|---|---|---|---|---|
| ApoSH SOD1 variant | PDI[a] | PDI width[b] (nm) | Intensity Distribution Hydrodynamic[c] Diameter (nm) | Number Distribution Hydrodynamic[c] Diameter (nm) | PDI[a] | PDI width[b] (nm) | Intensity Distribution Hydrodynamic[c] Diameter (nm) | Number Distribution Hydrodynamic[c] Diameter (nm) |
| pWT | **0.363** | **7.785** | **4.620** | **3.340** | **0.926** | **65.05** | **707.1** | **3.845** |
| | *0.053* | *2.993* | *0.041* | *0.115* | *0.104* | *11.39* | *404.5* | *0.080* |
| G93S | **0.806** | **38.76** | **6.852** | **3.977** | **0.833** | **32.16** | **398.6** | **5.664** |
| | *0.045* | *11.59* | *0.199* | *1.184* | *0.068* | *5.263* | *72.75* | *0.431* |
| G93A | **0.321** | **62.11** | **7.030** | **4.961** | **0.931** | **285.1** | **1256** | **3.910** |
| | *0.036* | *54.47* | *0.050* | *0.178* | *0.120* | *49.43* | *445.1* | *0.529* |
| G37R | **0.404** | **7.876** | **9.252** | **5.317** | **0.242** | **22.26** | **60.06** | **24.23** |
| | *0.017* | *0.453* | *0.072* | *0.365* | *0.014* | *0.676* | *1.920* | *2.190* |
| G93R | **0.327** | **37.70** | **9.025** | **5.352** | **0.215** | **14.33** | **35.81** | **15.98** |
| | *0.078* | *48.06* | *0.684* | *0.716* | *0.014* | *0.435* | *0.655* | *0.917* |
| A4V[d] | **0.374** | **27.00** | **20.15** | **4.616** | **1.0** | **1679** | **514.9** | **36.00** |
| | *0.150* | *25.92* | *11.08* | *1.288* | *0.0* | *322.4* | *221.7* | *22.36* |

Bold and italicized values are the averages and standard deviations respectively resulting from three measurements of a single sample. [a]PDI is obtained from the Cumulants method of fitting the autocorrelation function (Eq. 4). Values larger than 0.1 are indicative of polydispersity (i.e., the sample contains more than one species in solution). High PDI values result in larger peak widths and greater uncertainty in the sizes reported. [b]Peak width reflects the uncertainty in the hydrodynamic diameter obtained from the intensity distribution and is related to the PDI according to Eq. 4. [c]The sizes reported correspond to the main peak in the distributions. [d]A4V aggregated readily at high protein concentration, leading to the formation of aggregates visible by the eye. Such aggregates diffuse too slowly to be measurable by DLS, thus the sample could only be measured for ~120 hours.

In contrast, the high aggregator mutants A4V, G37R and G93R show extensive aggregation with increases in both $D_{h,I}$ and $D_{h,N}$ values. Although all aggregate considerably, G37R and G93R have smaller differences between $D_{h,I}$ and $D_{h,N}$ at $t_{late}$ unlike A4V (Figure 2.1E,F, Table 2.3). The smaller differences in $D_h$ values and the decrease in PDI values for G37R and G93R suggest these mutants are aggregating more homogeneously. Overall, the high aggregator variants show a large increase in the $D_{h,N}$ values at $t_{late}$, indicating extensive aggregation with increased protein concentration; on the other hand, the large aggregates formed by the low aggregator mutants scatter a large amount of light but. due to their low concentration, they do not influence $D_{h,N}$ values.

Static light scattering (SLS) measurements provided additional support for the mutant grouping based on DLS aggregate size measurements. In the SLS, the extent of aggregation with time in each sample was measured by the amount of scattered light; the low aggregator variants show limited scattering compared to the significant scattering of the high aggregator mutants (Figure 2.5). There is no obvious lag time for the high aggregators, suggesting a low/no energy barrier between the initial state(s) and aggregation-prone state(s) (Figure 2.5). In contrast, the low aggregator variants that give rise to large aggregates show some evidence for a lag phase (Figure 2.5). Similar to DLS, due to the low concentration of the large aggregates, they may not be observed in all SLS measurements, which therefore have large associated standard deviations. Taken together, DLS and SLS results suggest the SOD1 variants have either a low or high aggregation propensity, with or without a lag time, respectively.

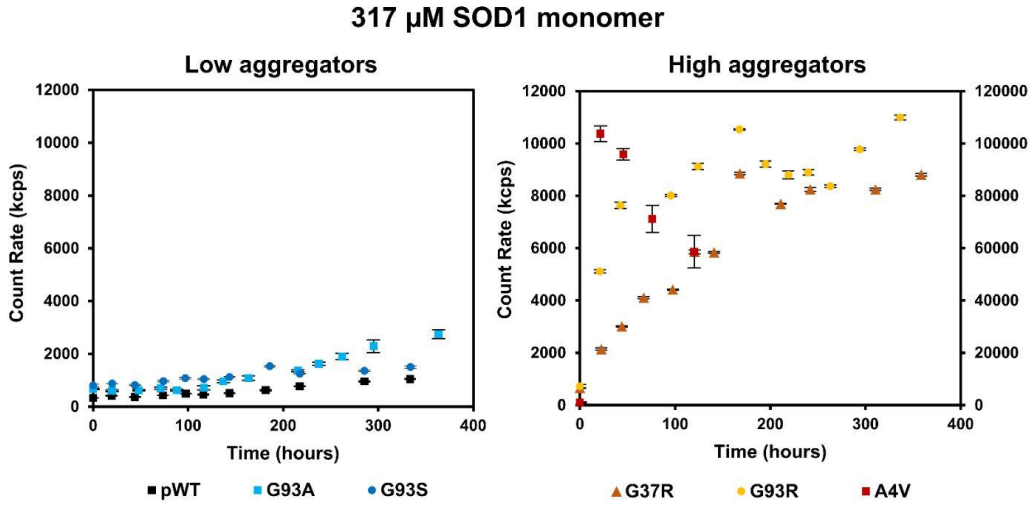



Figure 2.5. SLS reveals different levels of aggregation through low aggregators and high aggregators incubated at high concentration (317 μM). A4V, the only mutant plotted on the right-hand axis (red squares), aggregated the most and was only measurable by light scattering for ~120 hours. After this point the aggregates were too large to be detected. G37R (orange triangles) and G93R (yellow circles) also aggregated with very little detectable lag time. Incubation at high protein concentration, however, does not lead to much change in light scatter intensity over time for pWT, G93A, and G93S (black squares, light blue squares and dark blue circles, respectively). In this

case, the light scattering intensity fluctuated greatly between the measurements giving rise to large standard deviations in the measurements. This finding indicates that large particles are present in solution, yet these species are not abundant. Error bars represent the standard deviation of three measurements at each time point.

As another means to promote or modulate aggregation, we examined the effects of adding salt (NaCl and $Na_2SO_4$) to apoSH SOD1. Generally, salts can act to promote aggregation by screening electrostatic repulsion (apoSH has a net negative charge at pH 7.4 [111,112]). In addition, salts may affect protein stability and self-association according to the Hofmeister series, with $Na_2SO_4$ generally having larger effects than the relatively small effects of NaCl [113–116]. Along with the pWT control, we characterized destabilized mutants from the low aggregators (G93S) and high aggregators (A4V and G93R). Intracellular ionic strength is ~100-200 mM [117,118], thus we added 150 mM $Na_2SO_4$ and NaCl to samples with 63 μM protein. We also examined a stabilized mutant, V148I, with low aggregation propensity (Figure 2.1G,H), with both increased protein concentration (317 μM) and 150 mM $Na_2SO_4$. At $t_0$, $D_{h,I}$ and $D_{h,N}$ values reflect subtle changes between variants across all conditions (Figure 2.1G,H, Table 2.4). Overall, the differences in $D_{h,N}$ and $D_{h,I}$ between pWT and mutant samples suggest minimal initial increase in self-association under salt conditions.

In contrast, at $t_{late}$ the groupings of mutant aggregation propensity observed at high protein concentration are also prevalent under salt conditions. For the high aggregators, A4V and G93R, both the $D_{h,I}$ and $D_{h,N}$ increase for all samples, but the increase is larger for samples with salt which also show some presence of a lag time (orange bars in Figure 2.1G,H and 2.6, Table 2.4). For the low aggregator G93S, only samples with no salt and $Na_2SO_4$ showed $D_{h,I}$ increases with time (blue bars in Figure 2.1G,H, Table 2.4). The

aggregation for G93S in $Na_2SO_4$ also occurred with a comparable lag time to the high aggregators under the same conditions, but the lower light scattering intensity suggests G93S forms less aggregates than the high aggregators (Figure 2.6). For the other low aggregator, V148I, in $Na_2SO_4$ the $D_{h,I}$ and $D_{h,N}$ values are similar and relatively unchanged with time. The PDI values tends to be increased for both low aggregators which makes discerning smaller differences difficult due to increased error (Table 2.4). To summarize, the high aggregator mutants, A4V and G93R, show increased aggregation propensity and extensive aggregation under all conditions, whereas the low aggregators, G93S and V148I, undergo minimal aggregation.

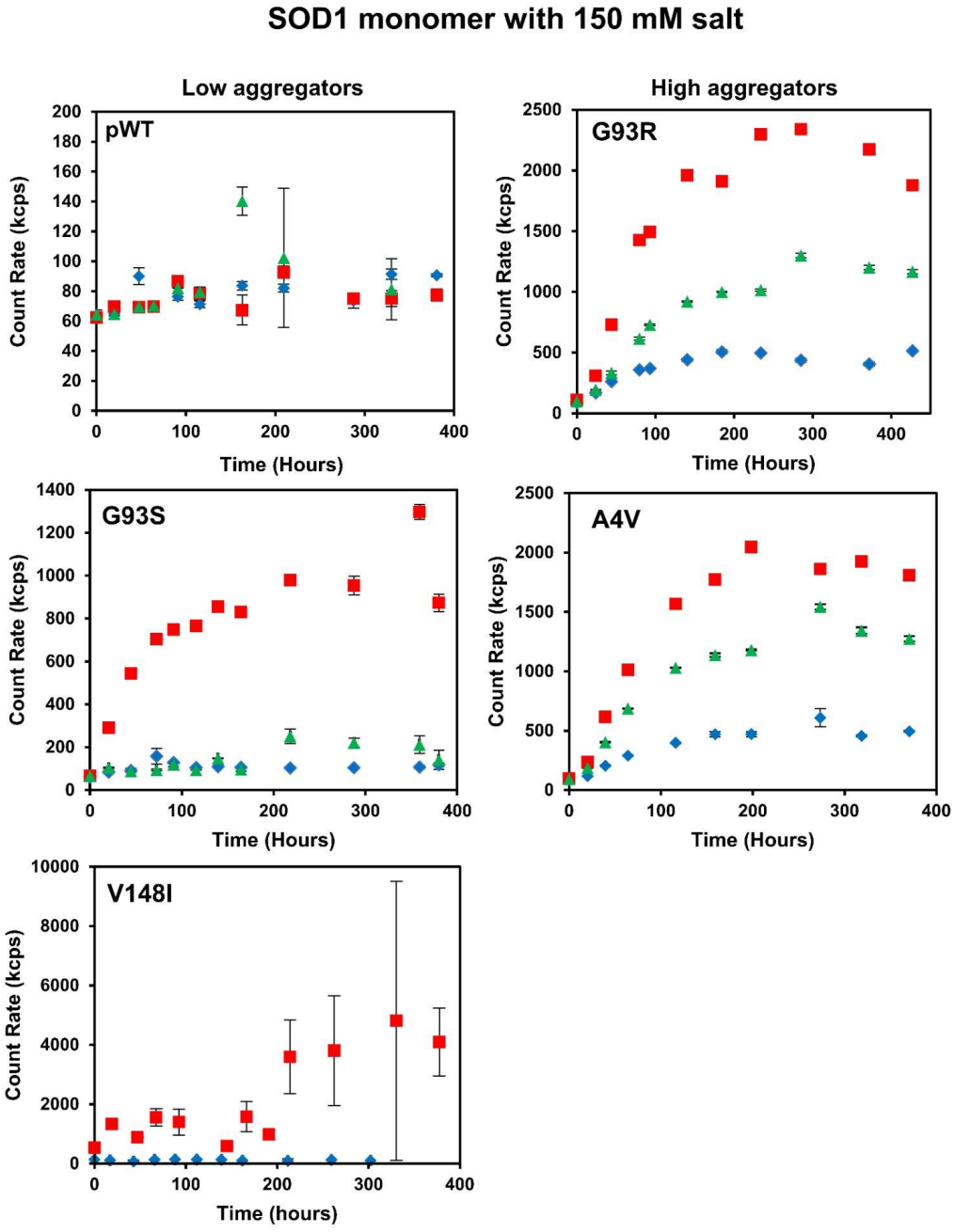



Figure 2.6. SLS reveals that mutants exhibit increased aggregation in 150 mM salt. For all panels, apoSH variants were incubated in buffer, 1 mM TCEP, 20mM HEPES, pH 7.4, with either no salt (blue diamonds), 150 mM $Na_2SO_4$ (red squares), or 150 mM NaCl (green triangles). $Na_2SO_4$ promotes aggregation of all mutants, but the effect was most pronounced for A4V, G93R, and G93S. NaCl only appears to enhance aggregation of A4V and G93R, and the effect is lower compared to $Na_2SO_4$. Error bars represent the standard deviation of three measurements at each time point.

Table 2.4. DLS data summary for mutant apoSH SOD1s at 63 μM, pH 7.4, 37 °C, 150 mM salt.

| | t=0 | | | | t > 300 | | | |
|---|---|---|---|---|---|---|---|---|
| ApoSH SOD1 variant | PDI[a] | PDI width (nm)[b] | Intensity Distribution Hydrodynamic Radius (nm)[c] | Number Distribution Hydrodynamic Radius (nm)[c] | PDI[a] | PDI width (nm)[b] | Intensity Distribution Hydrodynamic Radius (nm)[c] | Number Distribution Hydrodynamic Radius (nm)[c] |
| pWT | **0.205** | **2.146** | **5.383** | **3.131** | **0.267** | **5.231** | **5.228** | **3.449** |
| No Salt | *0.025* | *0.250* | *0.332* | *0.387* | *0.083* | *5.657* | *0.253* | *0.398* |
| pWT | **0.228** | **2.298** | **5.800** | **3.013** | **0.350** | **127.9** | **5.350** | **3.522** |
| NaCl | *0.031* | *0.273* | *0.117* | *0.980* | *0.135* | *214.9* | *0.213* | *0.619* |
| pWT | **0.240** | **2.197** | **5.262** | **2.816** | **0.236** | **61.37** | **5.551** | **3.292** |
| $Na_2SO_4$ | *0.023* | *0.093* | *0.502* | *0.541* | *0.066* | *79.72* | *0.888* | *1.050* |
| V148I | **0.330** | **10.29** | **5.568** | **4.066** | **1.0** | **5551** | **4.774** | **4.378** |
| $Na_2SO_4$ | *0.026* | *4.689* | *0.258* | *0.221* | *0.0* | *3028* | *0.728* | *0.344* |
| G93S | **0.239** | **3.226** | **7.136** | **3.899** | **0.502** | **30.34** | **84.11** | **3.880** |
| No Salt | *0.024* | *0.548* | *0.708* | *0.930* | *0.143* | *46.08* | *85.18* | *1.598* |
| G93S | **0.236** | **61.46** | **6.788** | **4.136** | **0.704** | **837.6** | **6.287** | **5.099** |
| NaCl | *0.050* | *101.5* | *1.105* | *0.503* | *0.143* | *353.3* | *0.186* | *0.558* |
| G93S | **0.239** | **4.570** | **8.312** | **3.944** | **0.629** | **214.8** | **331.8** | **4.354** |
| $Na_2SO_4$ | *0.023* | *1.851* | *0.346* | *1.660* | *0.264* | *19.22* | *17.33* | *2.296* |
| A4V | **0.224** | **5.566** | **8.593** | **3.807** | **0.312** | **29.55** | **63.46** | **15.40** |
| No Salt | *0.043* | *9.779* | *1.452* | *1.376* | *0.119* | *10.88* | *15.38* | *14.70* |
| A4V | **0.269** | **4.418** | **9.239** | **3.907** | **0.211** | **31.79** | **86.95** | **43.86** |
| NaCl | *0.015* | *0.268* | *0.409* | *1.466* | *0.013* | *1.610* | *3.608* | *3.869* |
| A4V | **0.273** | **4.665** | **9.907** | **4.446** | **0.212** | **33.47** | **90.67** | **43.77** |
| $Na_2SO_4$ | *0.024* | *0.540* | *0.804* | *2.712* | *0.022* | *4.221* | *8.127* | *6.920* |
| G93R | **0.240** | **3.732** | **7.563** | **3.971** | **0.598** | **26.87** | **151.1** | **11.60** |
| No Salt | *0.056* | *1.155* | *0.806* | *1.053* | *0.334* | *9.509* | *104.2* | *11.58* |
| G93R | **0.254** | **4.889** | **9.355** | **4.297** | **0.245** | **30.78** | **83.08** | **35.06** |
| NaCl | *0.037* | *1.813* | *1.326* | *1.378* | *0.018* | *3.004* | *9.362* | *7.290* |
| G93R | **0.257** | **14.51** | **9.754** | **4.390** | **0.233** | **36.20** | **96.74** | **40.40** |
| $Na_2SO_4$ | *0.035* | *29.85* | *0.920* | *1.633* | *0.030* | *6.532* | *16.03* | *6.737* |

Bold and italicized values are the averages and standard deviations, respectively resulting from multiple experiments. [a]PDI is obtained from the Cumulants method of fitting the autocorrelation function (Eq. 4). Values larger than 0.1 are indicative of polydispersity (i.e., the sample contains more than one species in solution). High PDI values result in larger peak widths, and thus greater uncertainty in the sizes reported. [b]Peak width reflects the uncertainty in the hydrodynamic diameter obtained from the intensity distribution and is related to the PDI according to Eq. 4. [c]The sizes reported correspond to the main peak in the distributions.

### 2.2.2 Atomic force microscopy (AFM) reveals distinct aggregation modes for mutant apoSH SOD1s

Atomic force microscopy (AFM) imaging was performed to obtain complementary structural information for the aggregates. The aggregates were formed in solution and deposited on the mica surface for imaging in air in intermittent contact mode. The AFM

images of aggregates formed at high protein concentration show that aggregate size and morphology differ markedly between the high and low aggregator mutant SOD1s (Figures 2.7 and 2.8). Analysis of AFM images was performed using the SPIP software to compare aggregate dimensions and abundance between variants (summarized in Table 2.5, Figure 2.9). We also extend typical AFM aggregate size analysis to quantify the number of monomers in aggregates.[119] Finally, we demonstrate violin-box plots as a useful tool that augments standard boxplot summaries of AFM particle statistics with additional information on data distribution (Figures 2.10-2.11). The data distributions, in addition to the other size estimates, are valuable for discerning subtle aggregate differences between variants and their aggregation grouping.

When comparing the AFM data to the DLS results, certain considerations must be acknowledged. Due to the asymmetry of the particle data distribution, median rather than average values were used to compare different SOD1 variants [120]. The median length and breadth estimates of the pWT particles at high concentration were ~20 nm and ~12 nm (Figures 2.10A and 2.11, Table 2.5), respectively, larger than the apoSH monomer reported by DLS (Figure 2.1E,F). Volume discrepancies between AFM measurements and expected size measurements based on the PDB can be due to variation in the aggregates interacting with the mica slide as well as volume changes from association [121,122]. Because the entire mica surface is not imaged exhaustively, it can be challenging to interpret differences in the absolute number of particles detected between two samples but comparing the shapes of the distributions is meaningful. Given these caveats, there is good agreement between DLS and AFM results, as considered in detail below.

Table 2.5. AFM image statistics of mutant apoSH SOD1s incubated at 317 μM, pH 7.4, 37 °C.

| Red Apo SOD1[a] | Median Length (nm)[b] | Median Breadth (nm)[b] | Average Height (nm)[b] | Number of Large Particles (>200 nm)[c] | Median Aspect Ratio[b,d] | Median Aspect Ratio of Large Particles (>200nm) [b,d,e] | Average Coverage (%)[f] |
|---|---|---|---|---|---|---|---|
| pWT | 19.6 (14.7-25.2) | 12.4 (10.4-15.9) | 0.26 (0.21-0.32) | 1 | 1.38 (1.33-1.63) | 3.91 (3.05-6.50) | 0.30 ± 0.55 |
| G93S | 39.1 (24.5-58.8) | 24.3 (14.7-41.5) | 0.13 (0.10-0.18) | 9 | 1.50 (1.33-1.88) | 2.29 (1.71-3.30) | 7.61 ± 4.55 |
| G93A | 30.1 (19.9-55.4) | 16.4 (10.9-39.4) | 0.15 (0.12-0.20) | 48 | 1.57 (1.33-2.04) | 3.37 (2.38-4.90) | 4.98 ± 2.65 |
| G93R | 48.1 (28.2-80.4) | 27.5 (17.9-44.1) | 0.50 (0.32-0.67) | 117 | 1.57 (1.33-1.97) | 2.02 (1.63-2.54) | 14.1 ± 7.14 |
| G37R | 24.3 (17.0-39.3) | 13.9 (10.8-21.8) | 0.24 (0.18-0.31) | 55 | 1.50 (1.33-1.89) | 2.23 (1.72-2.79) | 9.72 ± 2.03 |
| A4V | 46.9 (26.0-83.0) | 30.7 (17.0-49.8) | 0.44 (0.24-0.70) | 58 | 1.50 (1.33-1.81) | 1.71 (1.44-2.14) | 8.85 ± 4.39 |

[a]All apoSH SOD1 samples were incubated for over 300 hours at 37 °C before being deposited on the mica and imaged. The exception is G37R labelled $t_0$, which was not incubated. [b]Data represent the median and first to third quartile range in brackets of SPIP measurements obtained from all AFM images. Due to asymmetry in the values, the median rather than the average is shown. [c]The number of large particles was calculated by summing the total number of particles over 200 nm and dividing that number by the number of images obtained for that sample. Thus, values represent the average number of particles detected per image. In cases where very few large aggregates were detected and normalization according to the number of images obtained resulted in a number between 0 and 1, the value was rounded to 1. [d]The aspect ratio equals the particle length divided by breadth. [e]The aspect ratio of particles with lengths greater than 200 nm. [f]The coverage is the area taken up by aggregates divided by the total area. Values represent the average and standard deviation of all images. Notably, the lengths of fibres that are not linearly displayed on the mica will be underestimated and the breadths overestimated; therefore, the aspect ratios (length divided by breadth) will similarly be underestimated.

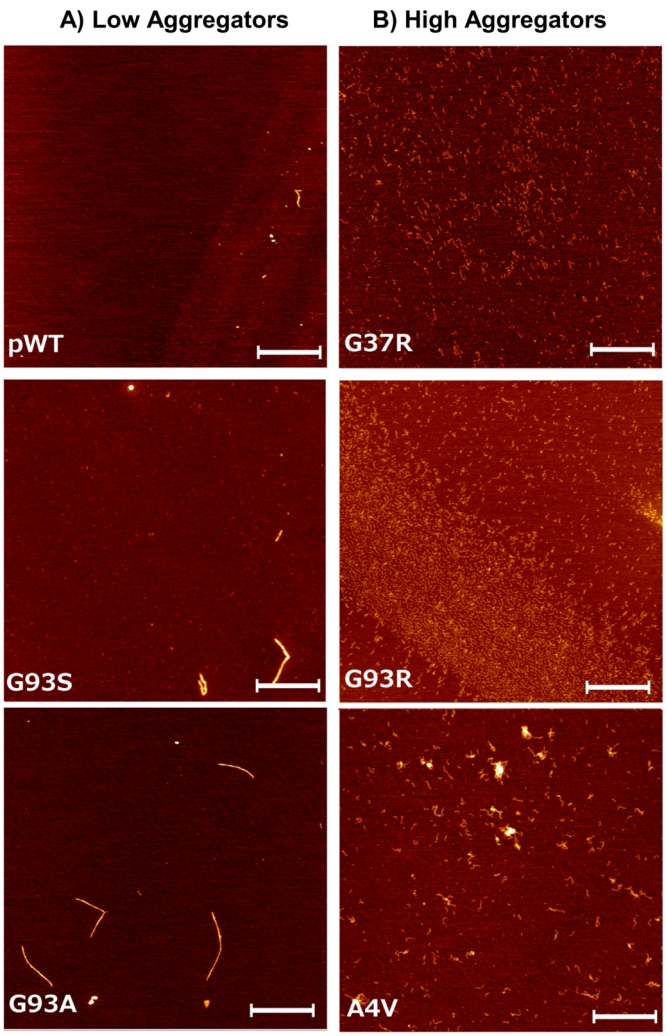



Figure 2.7: AFM reveals that different SOD1 variants form aggregates of differing sizes and morphologies when incubated at high concentration (317 μM SOD1 monomer). AFM images of SOD1 variants incubated for ~300 hours, except A4V which was only incubated for ~120 hours. Images are 5 x 5 μm and white bars indicate 1 μm. Consistent with DLS results, pWT showed the lowest levels of aggregate formation, although some fibres are apparent. Samples that displayed lower aggregation propensity (A), pWT, G93A, and G93S, formed a small number of fibres, greater than 1 μm. Samples that exhibited the highest levels of aggregation (B), A4V, G37R, and G93R formed numerous particles ~25-80 nm in length. For all images, the aggregates are so large that they mask the presence of small particles resembling monomers and oligomers.

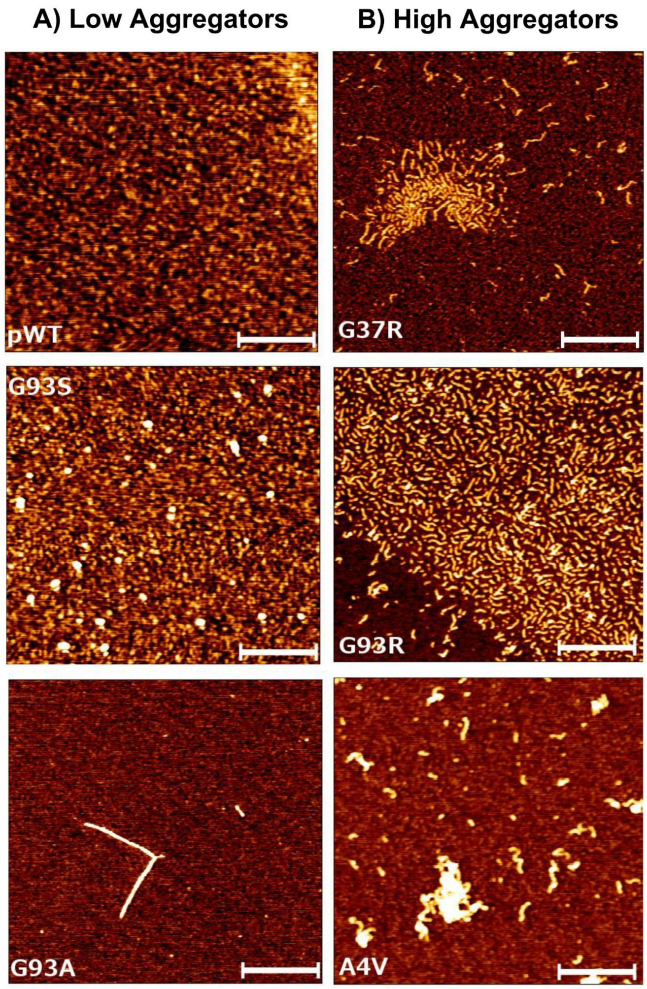



Figure 2.8. Close-up AFM images of different apoSH variants incubated at high concentration (317 μM) reveal the abundance of monomers and small oligomers. Background 2 x 2 μm images of apoSH samples incubated at 317 μM are shown. White bars indicate 500 nm.

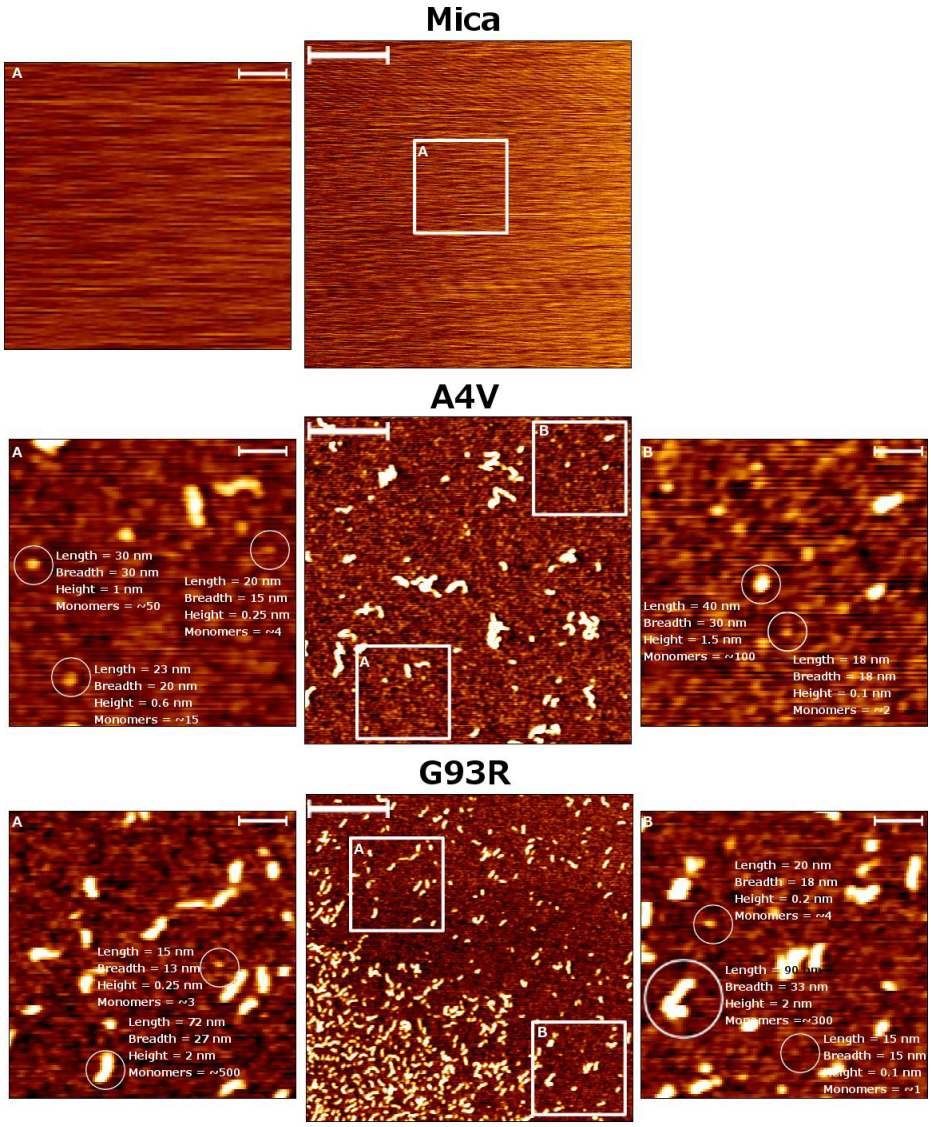



Figure 2.9. Size estimates of small oligomers that are abundant in the AFM images of apoSH variants incubated at high concentration (317 μM). Numerous small particles representing monomers and small oligomers 10-20 nm in length are evident in the background of the 5 x 5 μm images shown in Figure 2.7. The presence of these particles is clear when comparing the images to those obtained for cleaved mica (top images). In this figure, centre images are 2 x 2 μm and white bars represent 500 nm. White boxes represent the regions that are shown close-up (~580 x 580 nm) in the right and left panels, and in these images white bars represent 100 nm. Size estimates were calculated by hand using JPK Data Processing Software (JPK Instruments AG).

Consistent with the light scattering data, pWT at high concentration contained the least aggregated material. There is only a slight indication of low levels of long large aggregates (>200 nm) we refer to as fibres (Figures 2.7A and 2.10A,B), with the majority

of particles having volumes of ~1 to 4 monomer units (Figure 2.10C). The AFM slides for pWT have the lowest average coverage (area taken up by the aggregates divided by the total area, Table 2.5), in agreement with the minimal aggregation observed by DLS (Figure 2.7.1E,F, Table 2.3). For the few particles larger than 200 nm, the average aspect ratio is relatively large, suggesting fibre formation (Table 2.5). AFM images show more aggregation for all mutants at high concentration relative to pWT, even for mutants categorized as low aggregators (Figures 2.7 and 2.10A-C). The low aggregator mutants G93A and G93S exhibit larger median length and breadth than pWT (Figures 2.10A and 2.11); volume estimates indicate more particles containing ~10 monomers (Figure 2.10C), which we refer to as oligomers due to their small amorphous structure. These smaller oligomers may not have been discernible by the DLS measurements due to the large size range in species present. In addition to the increase of small oligomers in G93A and G93S, there are more particles larger than 200 nm compared to pWT (Table 2.5), which nevertheless are still considered outliers due to their low concentration (Table 2.5) and they have little impact on the shapes of the size distributions (Figure 2.10). Thus, the AFM data agree with the DLS data: the low aggregator mutants show some enhanced aggregation relative to pWT.

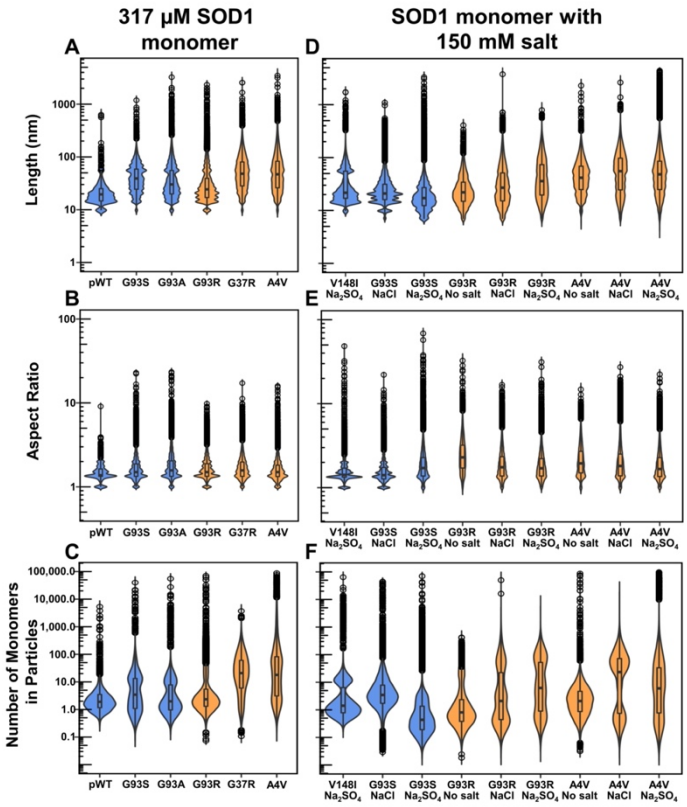



Figure 2.10: AFM measurements of aggregated SOD1 variants represented as box and whisker superimposed on violin plots. Samples at pH 7.4, 37 °C contained (A-C) 317 µM SOD1 monomer and (D-F) 61 µM SOD1 monomer (except V148I, at 317uM) with 150 mM salt. In each panel, the median of the data is shown by a horizontal line in the middle of the box and the box represents the first to third quartile of the data (i.e., the middle 50% of the data). The whiskers extend 1.5 times the interquartile range from the edges of the box and all points outside this range, which are considered outliers, are shown individually as black circles. Violin plots were superimposed on the box and whisker plots to illustrate the density of the data. The blue and orange violin plots correspond to the low and high aggregators, respectively. (A,B) Length, and aspect ratio distributions show that the median of the data is shifted to higher values for all mutants compared to pWT. The mutant distributions are broader and there are more large particles considered outliers. (C) Distributions showing the number of monomers within the particles are also broader for all mutants. Although the resolution of these distributions is lower than the others due to the higher uncertainty in the data, they reveal large populations of particles ~1-2 monomers for all samples and for pWT and G93A, these particles are the most abundant particles detected. For G37R and G93S, the median is shifted to slightly higher values, although the majority of particles observed are very small (>10 monomers). Only for A4V and G93R are particles containing more than 10 monomers the most abundant. (D,E) Length and aspect ratio distributions broaden to higher values for samples containing $Na_2SO_4$ and NaCl. Particle length and breadth estimates are larger for A4V and G93R compared to G93S and V148I, which is consistent with light scattering results that point to higher aggregation propensity of A4V and G93R. Furthermore, the A4V and G93R distributions are broader than G93S and V148I distributions. For G93S, in all distributions there are more large particles considered outliers (black dots), which can be accounted for by the presence of fibres in the images. (F) Distributions showing the number of monomers per particle indicate that the majority of particles detected have volumes consistent with ~1-5 monomers for salt free A4V and G93R samples, as well as G93S $Na_2SO_4$ and NaCl samples. For A4V and G93R $Na_2SO_4$ and NaCl samples, these distributions become broader as more particles containing 10-100 monomers are detected. For samples of V148I in $Na_2SO_4$, particles sized between ~1-2 monomers are abundant.

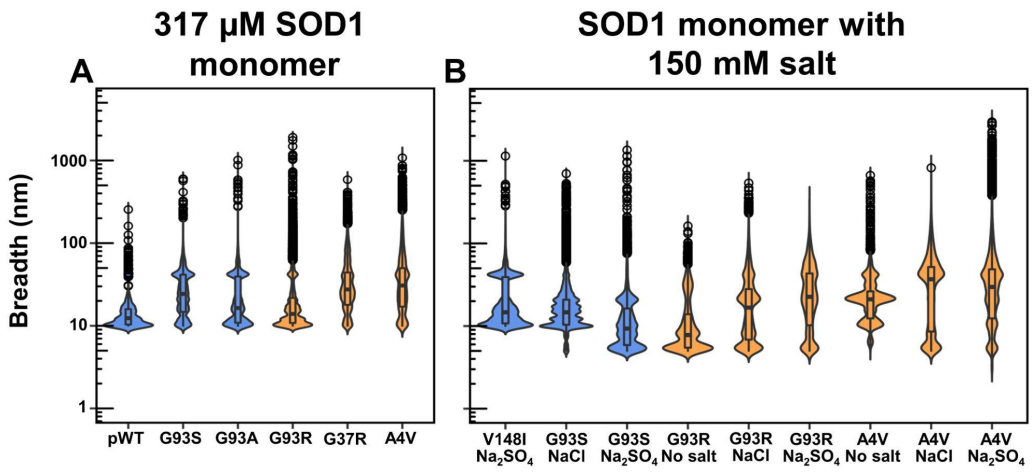



Figure 2.11. AFM measurements of aggregated apoSH SOD1 variants represented as box and whisker superimposed on violin plots. In each panel, the median of the data is shown by a horizontal line in the middle of the box and the box represents the first to third quartile of the data (i.e., the middle 50% of the data). The whiskers extend 1.5 times the interquartile range from the edges of the box and all points outside this range, which are considered outliers, are shown individually as black circles. Violin plots were superimposed on the box and whisker plots to illustrate the density of the data. The blue and orange violin plots correspond to the low and high aggregators, respectively. Breadth distributions corresponding to mutants at high concentration (A) and mutants with salt conditions (B).

While DLS data show the high aggregators variants have considerable differences in aggregation propensity and mechanism relative to the low aggregators, AFM imaging offered more insight on structural differences (Figure 2.7). For these variants, they show higher values for average coverage (Table 2.5) and the aggregates are composed predominantly of oligomers containing ~3-21 monomers (orange plots in Figure 2.10C). Interestingly, despite their increased aggregation, the high aggregators do not appear to form fibres observed for the low aggregators. The aspect ratios of aggregates formed by the high aggregator mutants are only slightly higher for particles greater than 200 nm compared to particles smaller than 200 nm (Table 2.5), which suggests that fibre formation is rare. The observations by AFM and DLS that the high aggregator variants

form numerous smaller aggregates, without a lag time, contrasts with the low aggregator variants and suggests the mutants undergo distinct modes of aggregation.

AFM images were also collected to assess the aggregates formed under the different salt conditions (Figure 2.12). Since the $D_{h,I}$ and $D_{h,N}$ distributions are quite broad and may not reflect subtle differences in the size of the particles formed under the different conditions, AFM provides a more powerful technique for assessing these differences. The imaging confirms that salt generally promotes fibrillization of A4V, G93R, G93S, and V148I (Figures 2.12-2.14). For the high aggregators, A4V and G93R, average length and breadth estimates indicate that longer and thicker fibrils form in $Na_2SO_4$ and NaCl than with no salt or at high protein concentration (orange plots in Figure 2.10D, Table 2.6). For both A4V and G93R, particles larger than 200 nm are more numerous in samples containing salt, and the average coverage for these samples are also generally higher than without salt (Table 2.6). Thus, salt promotes aggregation and increased length of fibrils (Figure 2.10D, Table 2.6).

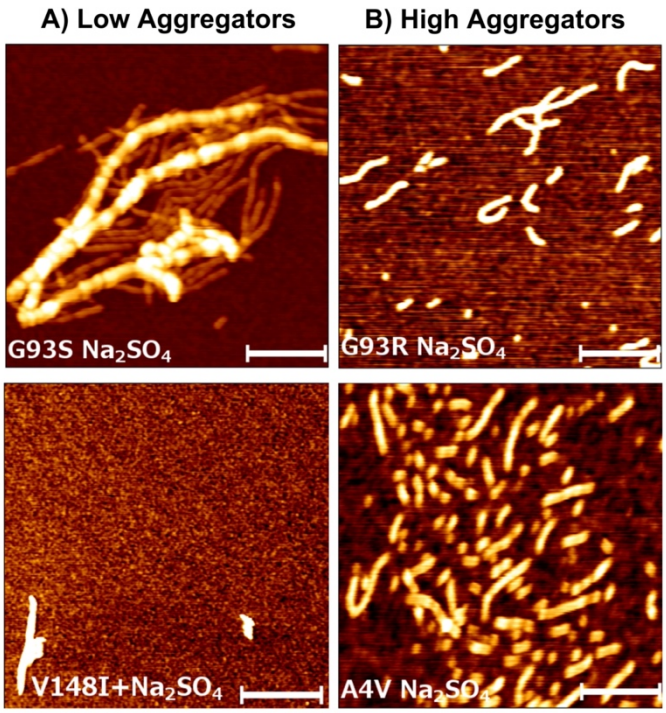



Figure 2.12: AFM reveals that SOD1 variants form aggregates of differing sizes and morphologies in $Na_2SO_4$. AFM images of SOD1 variants incubated in $Na_2SO_4$, for ~300 hours. All images are 2 x 2 µm and white bars indicate 500 nm. Only samples exhibiting high levels of aggregation by light scattering were imaged. A4V, G93R and G93S are at 63 µM and V148I in $Na_2SO_4$ is at 317 µM. AFM images for NaCl and salt free samples were consistent with DLS results (Figure 2.13). SPIP software indicates that the fibrils and fibres formed in salt tend to be longer than those formed in salt free buffer.

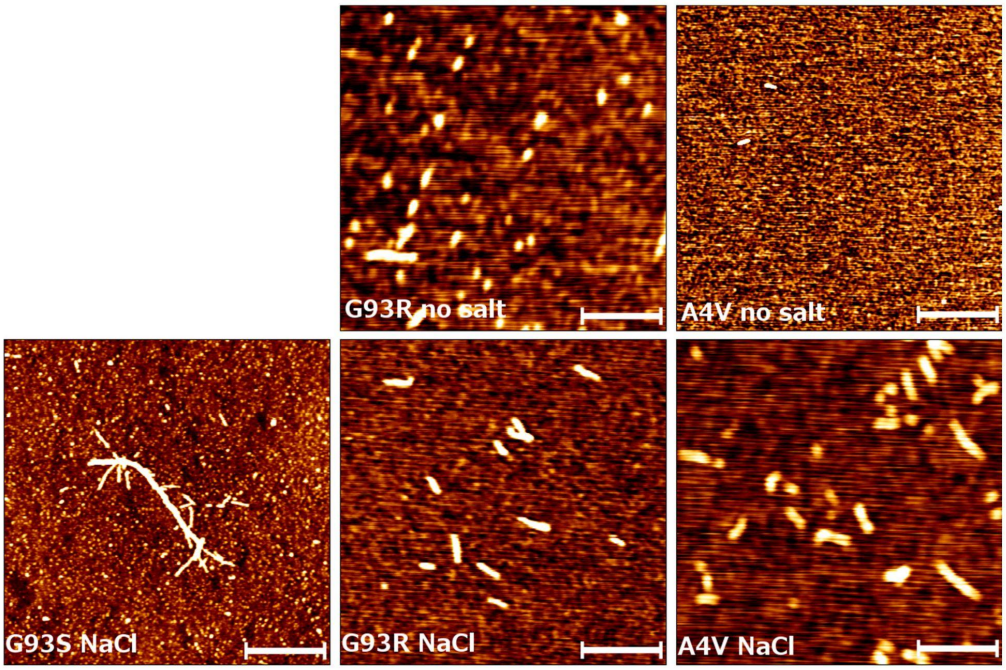



Figure 2.13. AFM reveals that SOD1 variants form aggregates of differing sizes and morphologies in NaCl. AFM images of apoSH variants incubated in NaCl for ~300 hours. All images are 2 x 2 μm and white bars indicate 500 nm. Only samples exhibiting high levels of aggregation by light scattering were imaged. A4V, G93R and G93S are at 63 μM in NaCl. Consistent with DLS results, less aggregation was observed for salt free samples. SPIP software indicates that fibrils formed in salt tend to be longer than those formed in salt free buffers.

Interestingly, the low aggregator mutants, G93S and V148I, form less aggregates but they still resemble fibres that are longer than the fibrils formed by A4V and G93R. The low aggregator mutant G93S forms fibres in $Na_2SO_4$ and, to a lesser extent, in NaCl (Figures 2.12 and 2.14, Table 2.6) compared to A4V and G93R. For both G93S and V148I, most of the particles are small (<5 monomers), but there are considerable aggregates larger than 200 nm in the presence of $Na_2SO_4$ (Figures 2.10D,E and 2.11). These large aggregates have the highest aspect ratios (Figure 2.10E, Table 2.6), even larger than aggregates formed at high protein concentration. Taken together, the increased length and breadth estimates for G93S and V148I in combination with the higher average coverage point to enhanced aggregation propensity relative to pWT, again

consistent with the light scattering results. This trend is familiar: at 317 μM, SOD1 mutants that displayed lower aggregation propensity formed, although minimally, longer aggregates. However, high aggregating mutants tend to form smaller, more numerous aggregates. Again, we see distinct aggregate characteristics between the high and low aggregators, suggesting different modes of aggregation.

Table 2.6. AFM image statistics of aggregated mutant apoSH SOD1s, pH 7.4, 37 °C, 150 mM salt.

| Red Apo SOD1[a] | Median Length (nm)[b] | Median Breadth (nm)[b] | Average Height (nm)[b] | Number of Large Particles (>200 nm)[c] | Median Aspect Ratio[b,d] | Median Aspect Ratio of Large Particles (>200nm) *b,d,e* | Average Coverage (%)[f] |
|---|---|---|---|---|---|---|---|
| V148I $Na_2SO_4$ | 21.6 (17.0-55.4) | 14.7 (10.8-39.1) | 0.14 (0.12-017) | 14 | 1.42 (1.33-1.71) | 2.86 (2.04-3.99) | 2.61 ± 1.49 |
| G93S NaCl | 20.9 (16.1-31.3) | 14.7 (10.4-20.8) | 0.32 (0.21-0.46) | 17 | 1.41 (1.27-1.74) | 2.19 (1.62-2.79) | 6.51 ± 2.34 |
| G93S $Na_2SO_4$ | 17.2 (12.4-27.1) | 9.3 (5.9-16.3) | 0.10 (0.07-0.12) | 30 | 1.71 (1.38-2.28) | 5.09 (3.32-8.19) | 2.68 ± 0.80 |
| G93R no salt | 22.0 (14.9-34.2) | 7.8 (5.5-13.9) | 0.16 (0.11-0.54) | 7 | 2.29 (1.71-3.20) | 3.69 (2.74-4.97) | 4.39 ± 1.04 |
| G93R NaCl | 27.0 (15.2-51.3) | 16.7 (6.8-28.0) | 0.16 (0.11-0.54) | 52 | 1.75 (1.38-2.34) | 2.25 (1.70-3.32) | 6.11 ± 6.62 |
| G93R $Na_2SO_4$ | 35.9 (19.6-73.5) | 22.6 (10.2-43.3) | 0.25 (0.16-0.53) | 78 | 1.70 (1.38-2.22) | 2.87 (2.12-3.75) | 8.31 ± 6.06 |
| A4V no salt | 21.9 (14.9-55.0) | 12.4 (9.8-33.9) | 0.12 (0.10-0.39) | 5 | 1.56 (1.33-2.22) | 2.01 (2.47-3.07) | 1.32 ± 1.96 |
| A4V NaCl | 55.4 (24.5-97.6) | 36.8 (8.6-51.9) | 0.33 (0.12-0.45) | 167 | 1.81 (1.38-2.51) | 2.11 (1.63-2.80) | 13.4 ± 3.50 |
| A4V $Na_2SO$ | 48.2 (24.9-85.7) | 29.7 (12.4-48.6) | 0.12 (0.10-0.28) | 109 | 1.67 (1.35-2.28) | 2.03 (1.57-2.76) | 16.4 ± 7.48 |

[a]All apoSH SOD1 samples were incubated for over 300 hours at 37 °C before being deposited on the mica and imaged.[b]Data represent the median and first to third quartile range in brackets of SPIP measurements obtained from all AFM images. Due to asymmetry in the values, the median rather than the average is shown. [c]The number of large particles was calculated by summing the total number of particles over 200 nm and dividing that number by the number of images obtained for that sample. Thus, values represent the average number of particles detected per image. In cases where very few large aggregates were detected and normalization according to the number of images obtained resulted in a number between 0 and 1, the value was rounded to 1. [d]The aspect ratio equals the particle length divided by breadth. Notably, the lengths of fibres that are not linearly displayed on the mica will be underestimated and the breadths overestimated; therefore, the aspect ratios (length divided by breadth) will similarly be underestimated.[e]The aspect ratio of particles with lengths greater than 200 nm. [f]The coverage is the area taken up by aggregates divided by the total area. Values represent the average and standard deviation of all images. Notably, the lengths of fibres that are not linearly displayed on the mica will be underestimated and the breadths overestimated; therefore, the aspect ratios (length divided by breadth) will similarly be underestimated.

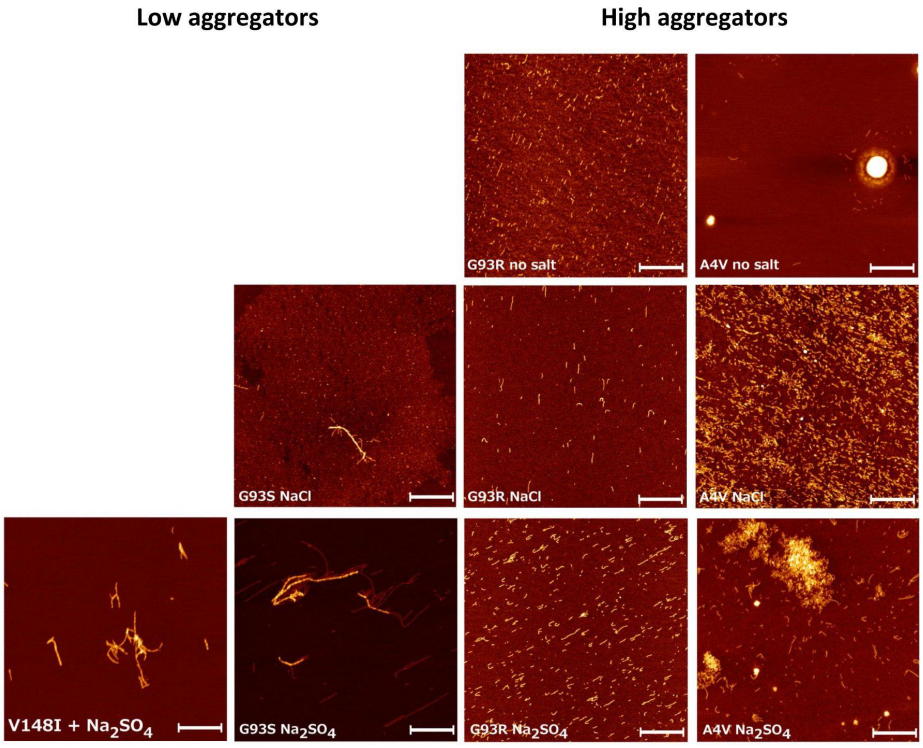



Figure 2.14. AFM reveals that A4V, G93R, G93S and V148I form aggregates of differing sizes and morphologies in $Na_2SO_4$ and NaCl. AFM images of apoSH variants incubated in salt free, $Na_2SO_4$, and NaCl for ~300 hours. All images are 5 x 5 μm and white bars indicate 1 μm. Only samples exhibiting high levels of aggregation by light scattering were imaged.

### 2.2.3 ATR-FTIR and ThT-binding data indicate different aggregate structures for SOD1 variants

The secondary structure of apoSH SOD1s during aggregation was assessed by ATR-FTIR spectroscopy.[72] Protein structural components show absorbance bands in different spectral ranges, such as: β-sheet (~1623-1641 $cm^{-1}$); unstructured/helix (1642-1657 $cm^{-1}$); and turns (1662-1686 $cm^{-1}$). We also obtained reference spectra of folded and heat denatured apoSH SOD1, as well as sonicated apoSH SOD1 as a measure for amyloid.[73] For the low aggregators at high protein concentration, the ATR-FTIR spectra at $t_0$ correspond to a combination of the folded and unfolded SOD1 references (black lines

on the left of Figure 2.15A). The ATR-FTIR spectra show little change with time (Figure 2.15A, blue lines in Figure 2.15B), in agreement with the minimal aggregation observed by DLS and AFM. The low aggregator variants, G93A and G93S, show a small shift toward the unfolded spectrum that pWT does not (Figure 2.15A). This shift may be reporting on the enhanced aggregation of G93A and G93S relative to pWT reported by the DLS and AFM data. The shift to unfolded signal is difficult to decipher; it could represent a shift in the protein population from the folded to unfolded state and/or local protein unfolding.

For the high aggregator mutants, the $t_0$ spectra are intermediate between the folded and unfolded apoSH SOD1 references except for A4V (black lines on the right side of Figure 2.15A). The spectrum for A4V at $t_0$ shifted to lower wavenumbers relative to the other mutants, which may be explained by the presence of initial aggregates, observed by DLS. With time, differences in spectral changes are observed between the high aggregator mutants (Figure 2.15A, orange lines in Figure 2.15B,C). A4V is the only mutant in this group that has a signal increase with time in the region corresponding to the unstructured protein (1642-1657 $cm^{-1}$) (Figure 2.15C). This distinct behavior may be due to its rapid aggregation and perhaps indicate more developed or different aggregate structure compared to G37R and G93R. Interestingly, while G37R and G93R display comparable aggregation propensity by DLS and similar aggregate morphology by AFM, their ATR-FTIR spectra differ. G37R shows very little spectral change, whereas G93R shows an increased signal in the range of 1620-1625 $cm^{-1}$ with time. G93R has a signal decrease in the unstructured region and an increase in the regions that correspond to increased intermolecular β-sheets (Figure 2.15A,C), similar to those observed for other

protein oligomers [123–126]. These spectral differences at high protein concentration point to changes in secondary structure that provide additional support for different modes of aggregation between and within the two apoSH SOD1 variant groups.

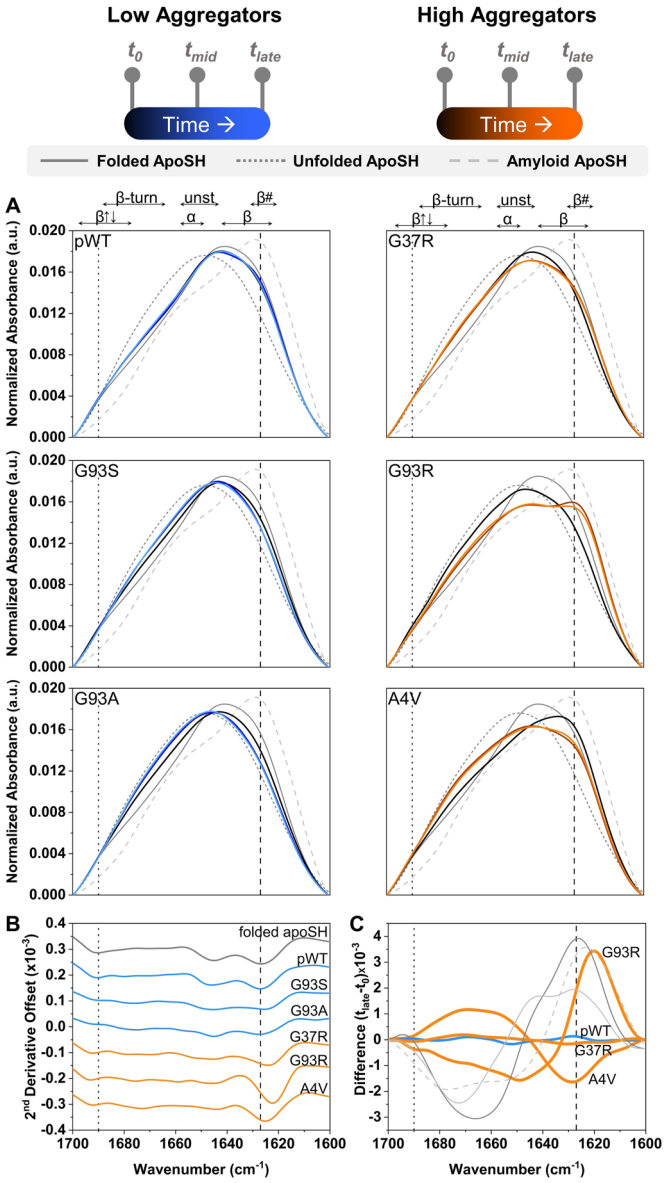



Figure 2.15: ATR-FTIR spectra reveal variable underlying aggregate structures for 317 μM samples. ATR-FTIR spectra were collected at an initial time point, $t_0$ (black), a time point in between the initial and the last time point ($t_{mid}$, light blue or light orange), and the last time point ($t_{late}$, dark blue or dark orange). The $t_{mid}$ spectra validate the spectral changes observed with time. All spectra are normalized by area and are overlaid with reference spectra of folded apoSH SOD1 (solid grey line), unfolded apoSH SOD1 (dotted grey), and sonicated apoSH SOD1 (amyloid reference; dashed light grey). Approximate spectral regions for antiparallel β-sheet (β⇓↑), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and β-sheet in aggregates

with strong intermolecular hydrogen bonds (β#) are indicated [127–129]. (A) The left, blue plots represent spectra for the low aggregators, which show minimal changes with incubation time, consistent with their minimal aggregation observed by DLS data and AFM images. The right, orange plots represent spectra for the high aggregators, which show distinct changes with time, although DLS data and AFM images show all variants to aggregate considerably. A4V and G93R show relatively large shifts, however in opposite directions. While G37R aggregates comparably to G93R based on DLS, it shows a smaller shift to lower wavenumbers by ATR-FTIR, comparable to shifts observed for the low aggregators. (B) The second derivative was calculated from the absorbance plots for $t_0$ (grey) and $t_{late}$ (blue or orange) collected in (A) for G93A and A4V (Figure 2.16 for remaining mutants). These representative plots highlight the minimal changes in the low aggregators (blue), and significant changes for the high aggregators (orange) specifically at the lower wavenumbers. To aid comparison, the spectra for variants at $t_{late}$ have been offset vertically. (C) Difference spectra ($t_{late}$ - $t_0$) show the low aggregators (illustrated by pWT, blue) show minimal change with time, in accordance with minimal extent of aggregation, whereas the high aggregators (orange) show variable changes and relative signals between each other and the folded relative to unfolded reference (solid grey line). The other difference reference spectra also are relative to the unfolded spectrum, which includes folded apoSS-unfolded apoSH (solid light grey) and sonicated apoSH-unfolded apoSH (amyloid reference; dashed light grey). The vertical dashed and dotted lines at 1627 nm and 1690 nm, respectively, are visual guides.

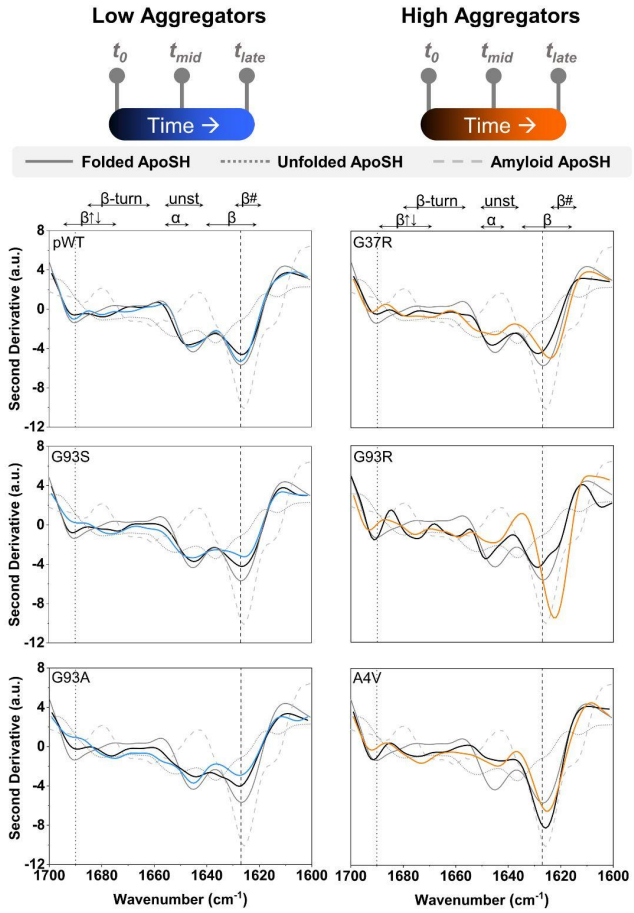



Figure 2.16. FTIR second derivative spectra reveal variable underlying aggregate structures for 317 μM samples. FTIR spectra were collected at an initial time point, $t_0$ (black), a time point roughly in

between the initial and the last time point, ($t_{mid}$, light blue or light orange) and the last time point ( $t_{late}$, dark blue or dark orange). All spectra are normalized by area, the second derivative was taken using Savitzky-Golay with 7 points of order and are overlaid with controls for folded apoSH (solid grey line), unfolded apoSH (dotted grey), and sonicated apoSH SOD1 (amyloid control; dashed light grey). Approximate spectral regions for antiparallel β-sheet (β⇓⇑), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and parallel or antiparallel β-sheet in aggregates (β#) are indicated." The vertical dashed and dotted lines are visual guides present at 1627 nm and 1690 nm, respectively.

In order to further investigate extended β-sheet structure, we conducted Thioflavin-T (ThT) binding experiments.[130] Typically, ThT-binding is taken as a measure of amyloid formation, but studies have also found ThT binding to the hydrophobic pockets within non-fibrillar, globular proteins.[130–132] The apoSH SOD1 samples show minimal ThT-binding (Figure 2.17) which agrees with the ATR-FTIR spectra not corresponding to the sonicated amyloid reference (Figure 2.15A). The ATR-FTIR spectra of the apoSH SOD1 samples show two distinct bands corresponding to native antiparallel β-sheet structure, at 1630 and 1690 $cm^{-1}$ (dotted line Figure 2.15); in contrast, the sonicated reference exhibits a decrease in signal at 1690 $cm^{-1}$, a hallmark of aggregates with parallel β-sheets generally present in amyloid structure.[128,133–136] Therefore, the aggregates studied here likely contain native antiparallel structure, not amyloid, suggesting native-like aggregation.

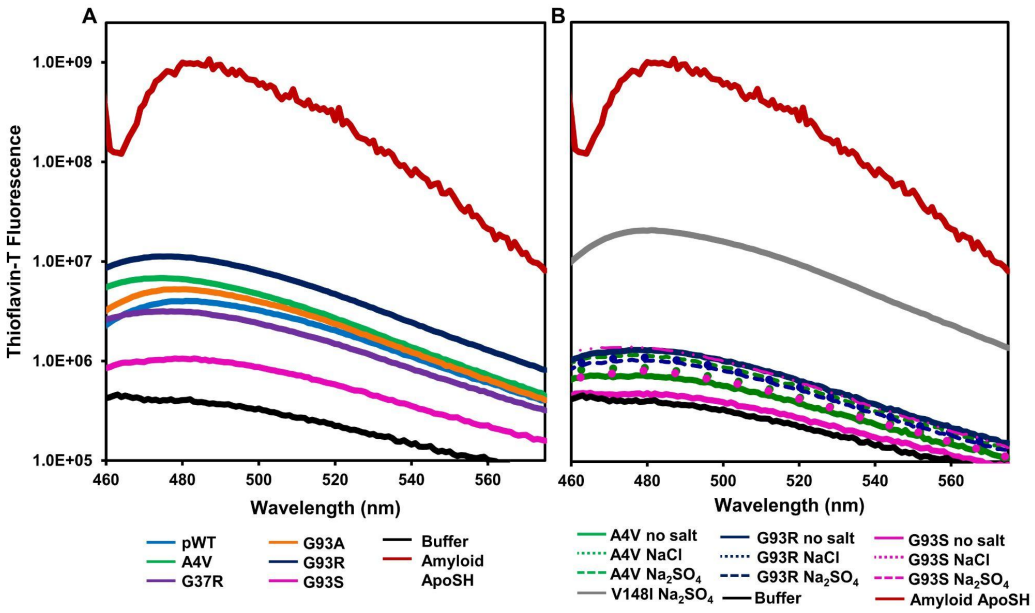



Figure 2.17. Differences in ThT fluorescence upon binding apoSH aggregates formed at 317 μM (A) and at 63 μM with 150 mM of salt (B) point to structural differences in the aggregates. Aggregates formed by apoSH incubated at high concentration (317 μM) and low concentration (63 μM) with 150 mM display differences in ThT fluorescence, which suggest that the β-sheet content of these aggregates differs.

We also use ATR-FTIR and ThT-binding to assess the structure of aggregates formed under different salt conditions. For the ATR-FTIR spectra at $t_0$, all variants in NaCl have increased signal in the unstructured region (1642-1657cm$^{-1}$) relative to no salt (Figure 2.18B, 2.20), and the effects of $Na_2SO_4$ are variable and complex. Both the A4V and G93R high aggregators exhibit spectral shifts to wavenumbers associated with increased intermolecular β-sheet structure, while the 1690 cm$^{-1}$ signal remains, corresponding to oligomerization regardless of initial spectral differences (Figures 2.18-2.21). In contrast to the high aggregators, we expected the low aggregators to show minimal changes. Interestingly, although G93S has similar initial ATR-FTIR spectra for both salt conditions (Figure 2.18B), G93S in $Na_2SO_4$ also develops increased signal at lower wavenumbers corresponding to oligomerization whereas G93S in NaCl shows

minimal changes (Figure 2.19). The stabilized mutant V148I in $Na_2SO_4$ behaves as expected for a low aggregator and shows minimal signal changes with time. It may be that V148I is dimerizing since its ATR-FTIR spectra resemble the apoSS (unmetallated oxidized SOD1) reference (Figure 2.19), which is a dimer[137], and AFM particle size estimates show V148I mostly forms dimers (Figure 2.10F and Table 2.6). This enhanced formation of native dimer by V148I may explain its lack of aggregation. Overall, the ATR-FTIR spectroscopy, DLS, and AFM results agree: $Na_2SO_4$ and to a lesser extent NaCl impact aggregation to variable extents for different mutant SOD1s.

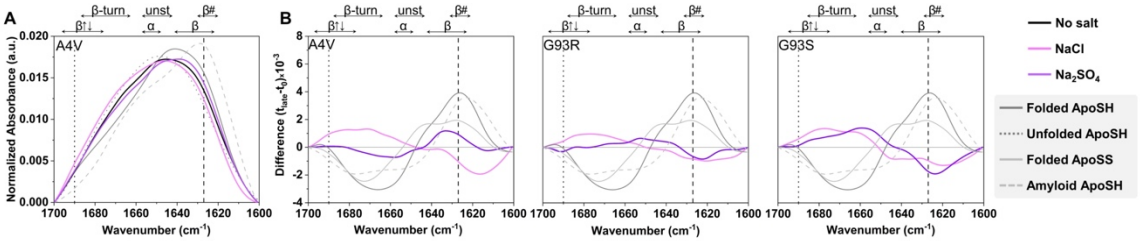



Figure 2.18 ATR-FTIR spectra reveal varying aggregate structures in 150 mM salt. Spectral regions for antiparallel β-sheet (β ⇓⇑), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and parallel or antiparallel β-sheet in aggregates (β#) are indicated." (A) ATR-FTIR spectra for 63 μM A4V samples for initial time points ($t_0$) in no salt (black), NaCl (pink) and $Na_2SO_4$ (purple). Spectra are normalized by area. Reference spectra include folded apoSH SOD1 (solid, grey line), unfolded apoSH SOD1 (unmetallated, disulfide-reduced SOD1 monomer, dotted, grey line), folded apoSS (unmetallated, oxidized-disulfide SOD1 dimer, dashed, light grey line), and sonicated apoSH SOD1 (amyloid reference; dashed, light grey line). (B) Difference spectra ($t_{late}$ - $t_0$) of A4V (left), G93R (middle) and G93S (right) in no salt, NaCl and $Na_2SO_4$. In salt, all variants, except A4V in $Na_2SO_4$, show more pronounced shifts with time toward unstructured signal compared to the absence of salt. The vertical dashed and dotted lines at 1627 nm and 1690 nm, respectively, are to guide the eye. For comparison, the difference spectra are overlaid with the differences between a specific form of SOD1 relative to unfolded protein: folded apoSH-unfolded apoSH, (solid grey line), folded apoSS-unfolded apoSH (solid light grey) and sonicated apoSH-unfolded apoSH (amyloid reference; dashed light grey). The differences between these reference spectra and the unfolded spectrum illustrate spectral changes associated with structure formation.

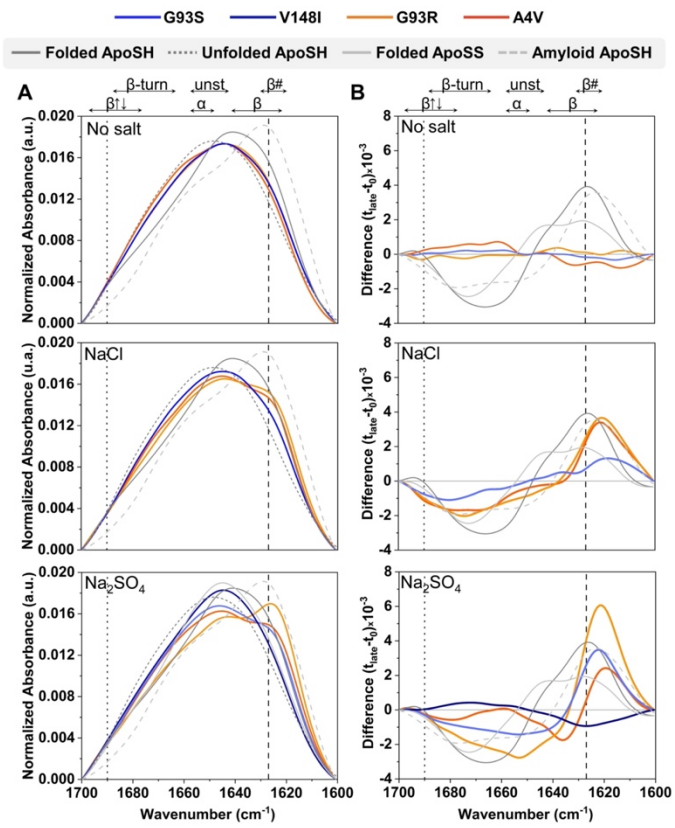



Figure 2.19: ATR-FTIR spectra reveal variable underlying aggregate structures for samples in 150 mM salt. The last time point data for all variants are compared for identical solution conditions. Spectral regions for antiparallel β-sheet (β⇓⇑), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and parallel or antiparallel β-sheet in aggregates (β#) are indicated. Variants are coloured by their extent of aggregation, blue for the low aggregators and orange for the high aggregators. (A) All spectra are normalized by area, and are overlaid with reference spectra for folded apoSH SOD1 (solid, grey line), unfolded apoSH SOD1 (dotted, grey line), folded apoSS (dashed, light grey line), and sonicated apoSH SOD1 (amyloid reference; dashed, light grey line). Without salt (top), absorbance spectra (A) and difference spectra ($t_{late}$ - $t_0$) (B) are similar for all variants. In NaCl (middle) the high aggregators, A4V and G93R (orange), show shifts to lower wavenumbers with NaCl relative to the low aggregator, G93S (blue). In $Na_2SO_4$ (bottom) however, the presence of salt has variable impact on mutants. All destabilized mutants in $Na_2SO_4$ show enhanced signal at lower wavenumbers except for the stabilized mutant, V148I, which shows little change. The vertical dashed and dotted lines are visual guides present at 1627 nm and 1690 nm, respectively. For comparison, the difference spectra are overlaid with the differences between a specific form of SOD1 relative to unfolded protein: folded apoSH-unfolded apoSH, (solid grey line), folded apoSS-unfolded apoSH (solid light grey) and sonicated apoSH-unfolded apoSH (amyloid reference; dashed light grey). The differences between these reference spectra and the unfolded spectrum illustrate spectral changes associated with structure formation.

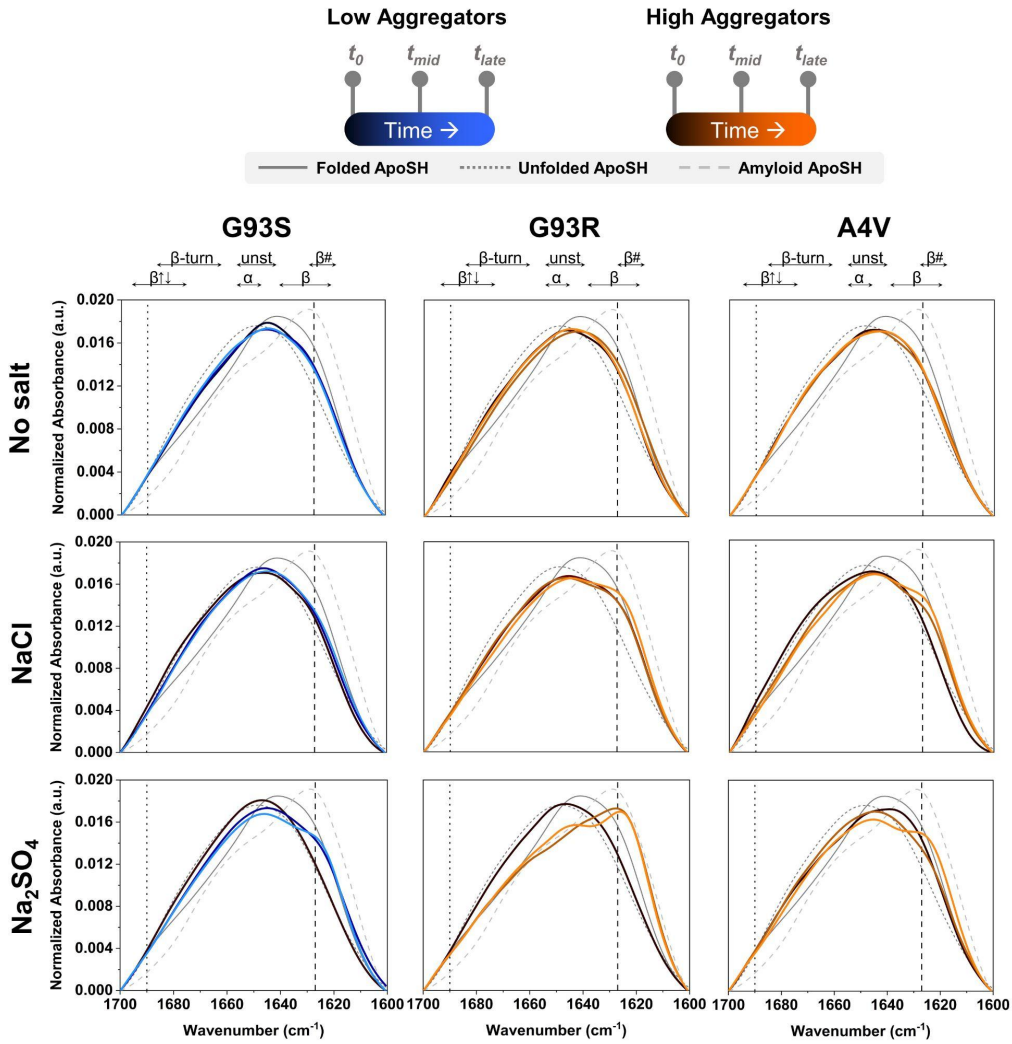



Figure 2.20. FTIR spectra reveal variable underlying aggregate structures for 63 μM with 150 mM salt. FTIR spectra were collected at an initial time point, $t_0$ (black), a time point roughly in between the initial and the last time point, ($t_{mid}$ light blue or light orange) and the last time point ( $t_{late}$, dark blue or dark orange). All spectra are normalized by area and are overlaid with controls for folded apoSH (solid grey line), unfolded apoSH (dotted grey), and sonicated apoSH SOD1 (amyloid control; dashed light grey). Approximate spectral regions for antiparallel β-sheet (βǂ), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and parallel or antiparallel β-sheet in aggregates (β#) are indicated." The vertical dashed and dotted lines are visual guides present at 1627 nm and 1690 nm, respectively.

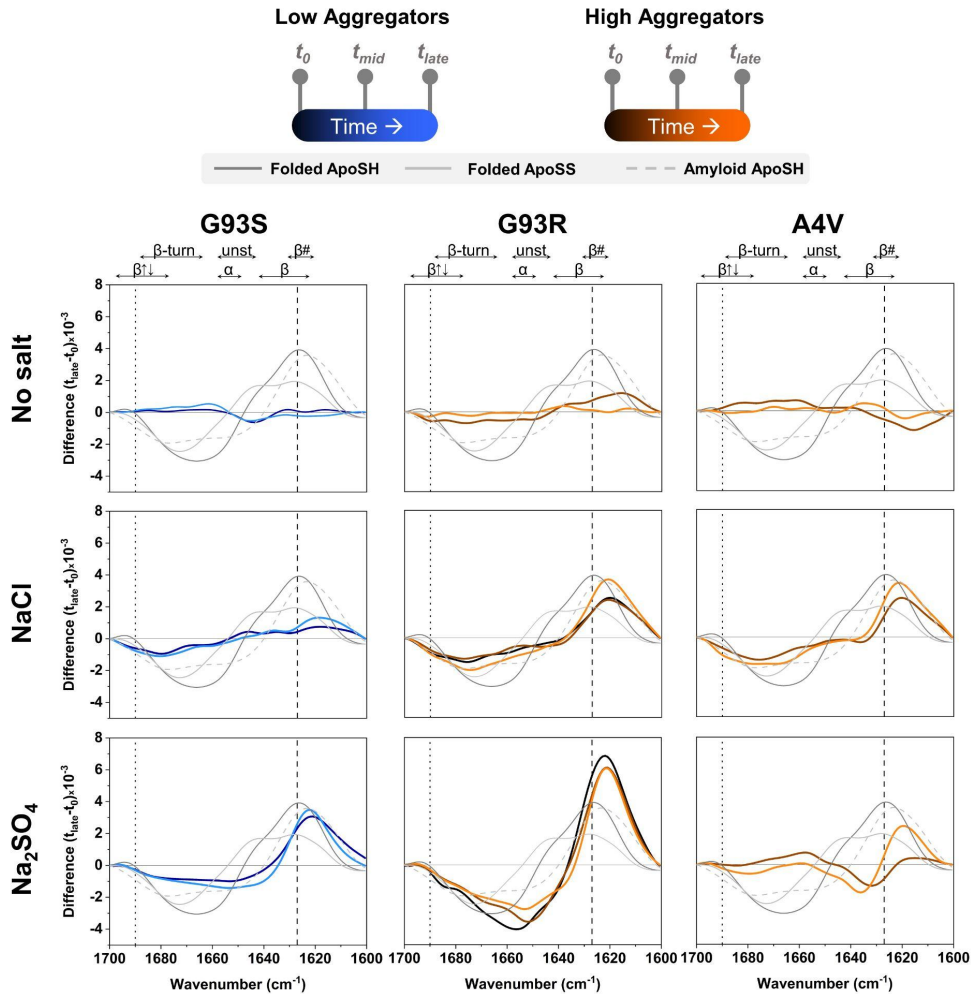



Figure 2.21. FTIR difference spectra ($t_{late}$-$t_0$) of different point mutations at 63 μM with 150 mM salt. The difference spectra are performed between a $t_{mid}$ and the $t_0$ (dark blue, black and dark orange), and between the $t_{late}$ and $t_0$ (light blue and light orange). All spectra are overlaid with the difference between the specific control relative to the unfolded control: folded apoSH-unfolded apoSH, (solid grey line), folded apoSS-unfolded apoSH (solid light grey) and sonicated apoSH-unfolded apoSH (amyloid control; dashed light grey). Approximate spectral regions for antiparallel β-sheet (β↓↑), β-turn, α-helix (α), unstructured (unst), β-sheet (β), and parallel or antiparallel β-sheet in aggregates (β#) are indicated. The vertical dashed and dotted lines are visual guides present at 1627 nm and 1690 nm, respectively.

The ATR-FTIR spectra in combination with ThT-binding also support that the fibrillar aggregates formed in salt vary in structure between mutants and solution conditions (Figure 2.17). Notably, although salt increases the length of the aggregates, there is no ATR-FTIR signal decrease in 1690 cm$^{-1}$, a hallmark of most amyloid proteins (Figure 2.19). Although ThT-binding is different between variants, the relative signal and

extent of ThT-binding vary for the same mutants under different conditions (Figure 2.17). However, it is difficult to ascertain whether these aggregate structures are truly different or at different points along the same aggregation pathway. For G93S and V148I, which aggregate minimally, the ThT-binding signals and ATR-FTIR spectra varied significantly. V148I has the highest ThT fluorescence even though it shows the least extent of aggregation of all the variants. G93R and A4V, which also share comparable aggregation propensity and aggregate morphology in salt, also show variability in their spectral changes and ThT-binding. Thus, mutants show variable structures under different solution conditions, especially in $Na_2SO_4$.

Table 2.7: Summary of SOD1 variant DLS, AFM and ATR-FTIR results.

| Method | | Variants | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | pWT | | | V148I | G93A | | G93S | | G37R[x] | | G93R[x] | | A4V[x] | | | |
| | | Sample conditions[a] | | | | | | | | | | | | | | | |
| | | 317 μM | NaCl | $Na_2SO_4$ | $Na_2SO_4$ | 317 μM | 317 μM | NaCl | $Na_2SO_4$ | 317 μM | 317 μM | NaCl | $Na_2SO_4$ | 317 μM | NaCl | $Na_2SO_4$ | |
| **DLS[b]** | Number[d] | NC | NC | NC | NC | NC | NC | NC | NC | * | * | *** | *** | *** | *** | *** | |
| | Intensity[d] | *** | *** | *** | NC | *** | *** | NC | *** | * | * | *** | *** | *** | *** | *** | |
| **AFM[c]** | Monomer[e] | *** | - | - | ***+ | *** | *** | *** | *** | R | R | R | R | R | R | R | |
| | Oligomer[e] | R | - | - | * | * | *** | * | * | *** | *** | *** | *** | *** | *** | *** | |
| | Fibres[e] | * | - | - | * | * | * | * | * | R | R | R | R | R | R | R | |
| **ATR-FTIR[b]** | Unfolded[f] | NC | - | - | NC | * | * | NC | NC | NC | NC | NC | NC | *** | NC | NC | |
| | Native[f] | NC | - | - | NC | NC | NC | NC | *** | NC | *** | *** | *** | NC | *** | *** | |

[x] high aggregator variants; low aggregator variants unless otherwise noted
[a] Sample conditions are: 317 μM: high protein concentration with no added salt; NaCl: 63 μM protein with 150 mM NaCl; $Na_2SO_4$: 63μM (and 317 μM for V148I) protein with 150 mM $Na_2SO_4$
[b] change in the DLS or ATR-FTIR measurement with time
[c] observation by AFM of species at $t_{late}$
[d] NC denotes no change in DLS measurement, *small aggregate detected, ***large aggregate detected
[e] R denotes species is rarely detected, *some species detected, ***abundant species detected
[f] An increase in signal in the spectral range 1642-1657 $cm^{-1}$ was taken to correspond to an increased proportion of unfolded protein and an increase in 1623-1641 $cm^{-1}$ to increased proportion of native-like protein; NC denotes no change observed in species' spectral features, *small signal increase observed, ***large signal increase observed
***+ denotes dimer formation by V148I based on ATR-FTIR
- no data

## 2.3 Discussion

Collectively, the DLS, AFM, and ATR-FTIR results reported here show consistent differences in aggregation characteristics of different apoSH SOD1 variants over a range of solution conditions (Table 2.7). Based on DLS and AFM data, the mutant SOD1 aggregation propensities can be grouped into low aggregators or high aggregators. The low aggregators show low net aggregation, and a lag phase associated with the slow formation of small numbers of large particles. In contrast, the high aggregators show extensive formation of intermediate size particles with little to no lag time (Figure 2.5). AFM images reveal distinct aggregate morphologies, with low aggregators forming low amounts of long fibres (>200 nm), and high aggregators predominantly forming oligomers. Further, ATR-FTIR and ThT-binding experiments indicate differences in the secondary structure and limited but varying amyloid character between aggregates for SOD1 mutants, including between different high aggregator mutants. Taken together, these results suggest distinct modes of aggregation between and among the low and high aggregators. Below, we consider further the structural and mechanistic features of SOD1 and related protein self-assembly.

### 2.3.1 Degree of protein unfolding does not explain aggregation of apoSH SOD1

Protein aggregation can arise from global and partial protein unfolding [104,138–141] as well as from native states modulated by a range of protein structural characteristics.[103,142–146] Here we define degree of unfolding to encompass both global and local unfolding. A

notable finding is that the global thermodynamic stability of mutant SOD1 does not correlate simply with the aggregation observed here. In the 63 μM DLS measurements, the extent of aggregation varies markedly between mutants and does not correlate with their thermal melting temperature, $T_{0.5}$. Mutants with the lowest $T_{0.5}$ (i.e. G37R, E100G and V148G) have the highest fraction of unfolded protein at 37°C and do not show increased aggregation relative to others (Figure 2.1, Table 2.1). These findings are consistent with a weak correlation between SOD1 aggregation propensity with stability and lack of correlation with disease characteristics observed previously (Table 2.1)[94], and implicate factors other than simply stability or the globally unfolded state in determining apoSH SOD1 aggregation.

The present results provide insights on additional structural features of aggregation under physiological-like solution conditions. Due to the marginal stability of apoSH SOD1 variants, they are susceptible to global and local unfolding. Yet, the ATR-FTIR spectra for the high concentration samples further support the degree of unfolding not correlating with the observed aggregation. The $t_0$ ATR-FTIR spectra show comparable unstructured signal between all variants but widely varying aggregation. Incubating the SOD1 samples in different salt conditions also supports unfolding not being essential to SOD1 aggregation, as all destabilized mutants in NaCl have comparable unstructured ATR-FTIR signal initially but undergo different extents of aggregation. In addition, NaCl consistently enhances the unstructured signal relative to no salt at $t_0$, but the NaCl samples do not consistently aggregate more than samples with no salt or $Na_2SO_4$. For instance, samples of A4V in NaCl show an increase in unstructured signal relative to A4V in no salt and $Na_2SO_4$. Irrespective of these differences in unstructured signal, A4V under

all conditions aggregates a comparable amount. Conversely, for G93S the most aggregation is observed in $Na_2SO_4$ which started with less unstructured signal than in NaCl. Although deconvoluting the ATR-FTIR signal in the unstructured range to fully versus partially unfolded protein is not possible, the extent of the unstructured signal clearly does not correlate with observed aggregation.

Aggregation not correlating with the degree of unfolded protein is important for deciphering aggregation mechanisms. Previous studies found that agitation induced formation of amyloid arose from unfolded SOD1.[104] It is important to note, however, that SOD1 aggregates in ALS patients do not show amyloid characteristics.[147] Considering most mutations and certain conditions increase unfolding of SOD1, a key finding here is that ATR-FTIR and ThT-binding show the aggregates are not mostly unfolded and not amyloid. Instead, the samples of destabilized mutants in salt and G93R at high concentration have increasing signal with time for intermolecular antiparallel β-structure in the oligomers and fibrils indicating native-like structure in the aggregates. In addition, the high aggregators self-associate mainly without a lag phase, consistent with non-nucleated, non-amyloid self-association from a near-native state, as also reported for proteins[141,142,148], see below. However, some extent of unfolding may also enhance aggregation, as less structured species may have exposed hydrophobics or more available conformations which aid in aggregate growth. Recruitment of SOD1 conformations with varying extents of native structure may also occur during propagation of different SOD1 aggregate strains in vivo.[107,108] Moreover, native-like oligomer formation has been widely observed and in many cases considered a distinct process from the formation of well-defined fibres with cross-β structure.[138,142,149,150] Other works have also

reported on partially folded or natively folded proteins being critical in the disease-relevant aggregation pathways.[143,151,152] Specifically for SOD1, the G93A rat model for ALS shows different amounts of antiparallel-β structure in accumulated protein in the brain.[151] Cataract-associated mutant crystallin also forms amorphous aggregates with native-like structure under physiological solution conditions in vitro whereas different amyloid-like fibrils form under destabilizing conditions.[143] Based on the above considerations, the lack of correlation observed here between degree of unfolded and aggregation under physiological conditions may be critical in understanding the disease aggregation pathways of SOD1 and other proteins.

### 2.3.2 Multimodal SOD1 aggregation shifts with mutation and changing solution conditions

It is of interest to compare the aggregation observed for SOD1 variants in different solution conditions with results observed for other proteins.[153–158] Generally, increasing protein concentration increases crowding, which may minimize unfolding and/or favour intermolecular interactions.[157,159] Such interactions in early aggregates are often weak and reversible.[158,160,161] These weak interactions may dominate aggregation for the SOD1 high aggregators, since the aggregates predominately exist as oligomers and show relatively little growth in aggregate size over time. Increasing protein concentration does not, however, strongly promote aggregation for all the apoSH SOD1 mutants in the same way (Figure 2.1E,F, 2, Table 2.3). Moreover, the effects of salts are still more complex. In addition to screening charges, Hofmeister salt effects can alter protein stability and self-association. Salts can also alter the balance of electrostatic and hydrophobic interactions between proteins.[115,162–165] The results for apoSH SOD1 are consistent with

NaCl mainly acting to decrease electrostatic repulsion between monomers while $Na_2SO_4$ also enhances folding and association.[113–116,164] The underlying molecular bases for the effects of protein concentration and salt on apoSH SOD1 aggregation differ and so affect the preferred aggregation modes.

Thus, the solution conditions influence differences in aggregate structure and mechanism. For the high aggregator mutants, A4V and G93R, aggregates formed under salt conditions consisted of mostly fibrils and some fibres, whereas the aggregates at high protein concentration were predominantly oligomers and rare fibres. On the other hand, the G93S low aggregator shows increased oligomer and fibre formation in salt relative to high protein concentration. These shifts in aggregate ensemble are also evidenced by changes in aggregation lag behaviour. The time course for the increased formation of fibrils and fibres for the high aggregators in salt has increased sigmoidal character. This lag suggests more contribution of a nucleation-dependent mechanism relative to nucleation-independent aggregation.[148] The low aggregator G93S shifts in $Na_2SO_4$ toward lag-independent aggregation compared to samples with no salt or NaCl which show more evidence of lag, similar to the high concentration sample. The preceding observations provide evidence for multiple co-existing aggregation mechanisms contributing to an aggregation ensemble, with the net behaviour modulated in complex ways by both mutations and solution conditions.

As the results here show an ensemble of aggregates consisting of a range of assemblies from oligomers to fibres, it is of interest to consider whether these species may be on separate, the same, or connected pathways. Some samples that predominantly form oligomers do so with no lag, whereas those with more fibre formation

show more lag, similar to results reported previously for SOD1 and other proteins.[148] Moreover, the lack of growth of the oligomers with time and differences in lag behaviour suggest the oligomers and fibres are on separate pathways.[138] Previous works have also found oligomers and fibres form through competing pathways[138,142,149,166]; their cellular impacts also differ, with oligomers generally having increased toxicity compared to fibres, which have been sometimes considered protective.[103,105,167–169] Since typically fibres form through nucleation-dependent mechanisms, it is possible the oligomers observed here represent the nuclei for fibre elongation; however, the narrow distribution of predominantly small species observed in the violin plots do not support this mechanism. While it is not possible to conclude definitively from the available information the relationships between oligomers and fibres observed here, we hypothesize there are multiple competing processes occurring that do not connect with each other, and so give rise to the observed ensemble of aggregates.

Likewise, multiple modes of aggregation have been reported for a variety of proteins including apoSH SOD1. Aggregate structures observed for SOD1 range from native-like to amyloid[103,105] with a wide range of structures among both fibres[99,104,170–173] and oligomers.[167,174,175] For example, near-native monomers can self-associate to form small oligomers[176] or linear assemblies.[177–179] Alternatively, different SOD1 peptides adopt a toxic corkscrew oligomer[180] or various steric zipper fibres[181], and unfolded full-length SOD1 forms amyloid structures upon agitation.[104] Partial unfolding of full length SOD1 was also proposed to give rise to the corkscrew oligomer.[180] Other studies have found that the more malleable the protein, the more aggregation-prone conformations are generated[166] giving rise to a higher diversity of aggregate structures.[149,182–184] These

considerations in conjunction with previous findings that apoSH SOD1 has a rugged free energy landscape where the population of mature and non-native conformations is altered by mutation[185,186] strongly supports the formation of different aggregate structures by SOD1 variants. Thus, the results herein along with previous findings support that SOD1 can form multiple kinds of aggregates and characterizing such multimodal aggregation is central to understanding and ultimately controlling SOD1 aggregation.

## 2.4 Conclusion

SOD1 is an attractive model for elucidating the molecular complexities of protein aggregation. Examining diverse disease mutations, we discern distinct aggregation processes by varying protein and salt concentrations. Depending on the solution conditions, the destabilized mutants studied here show shifts in the ensuing ensemble of aggregates. Low aggregator mutants form an ensemble of predominantly small particles (consisting of 1-10 monomers) and low levels of large fibres. The high aggregators assemble to larger oligomers or fibrils, with less fibre formation than the low aggregators. Both oligomer and fibre structures have been observed in other studies of SOD1 [99,104,167,170–175]; oligomers are generally thought to be more toxic compared to fibres, which have been suggested to be protective.[103,105,167–169] While it is not possible to exclude that the different aggregates may be on the same aggregation pathway but at different points, it appears instead that multiple separate modes of aggregation occur here. The SOD1 aggregate ensemble observed in this study aligns with previous work where in vitro studies revealed intermediates along the multiple pathways associated with aggregation that may not be discernible in vivo.[158] With multiple modes of aggregation available, our

work highlights essential considerations that are critical when investigating aggregation, including the substantial impacts of solution conditions and mutations, and how variations in aggregation can be discerned by combining complementary structure analysis tools. Questions remain regarding why proteins access certain modes of aggregates over others, which will require more systematic and atomic resolution studies; investigating the role of protein dynamics may be of particular interest. Overall, the work performed here extends previous research on SOD1 aggregation and may provide insight not only into the disease progression in SOD1-ALS, but also extend general understanding of the intricate mechanisms of protein aggregation.

## 2.5 Methods

### 2.5.1 Preparation of apoSH SOD1

We use the pseudo-wild type (pWT) construct (C6A, C111S) which was validated against wild-type mutants. ApoSH SOD1 was prepared from apo oxidized (apoSS) SOD1 as described previously.[94] Reduction of the disulfide bond was verified by SDS-PAGE as described previously.[94] After reduction, 63 μM apoSH SOD1 samples were filtered using Anotop 10 filters with a 0.02 μm cut-off (GE Healthcare Life Sciences) to remove aggregates formed during the reduction procedure. 317 μM samples could not be filtered using these filters, due to extensive sample loss; instead, samples were centrifuged at 16 300 rpm for 10 minutes to remove large aggregates or dust particles. Salt samples were prepared according to the normal protocol and then salt buffered in 20mM HEPES, 1mM TCEP was added to the apoSH SOD1 samples after reduction. During incubation, all samples were stored in a desiccator to avoid reoxidation. Aggregation was monitored

using a variety of techniques (vide infra), and all these experiments were performed on 45 µL aliquots removed from the apoSH SOD1 sample starting immediately after filtration ($t_0$) and approximately every 24 hours after $t_0$.

2.5.2 Light scattering experiments

Light scattering measurements were made according to the protocol described previously [94]. The intensity of light scattered (I) depends on the size and concentration of particles in solution. Fluctuations in light scattering intensity as a function of time (I(t)) depend on the translational diffusion coefficient ($D_c$) of the particles in solution [187,188]. According to the laws of Brownian motion, large particles diffuse at a slower rate than small particles. $D_c$ relates to the hydrodynamic diameter ($D_h$), the diameter of a sphere that diffuses at the same rate as the particle being measured, according to the Stokes-Einstein equation (Eq. 1),

$$D_c = \frac{KT}{3\pi\eta D_h} \quad \text{(Eq. 1)}$$

where, K is the Boltzmann constant, T is the temperature, and η is the viscosity of the sample. $D_c$ is determined from an autocorrelation function, which measures the correlation of I(t) with itself after a series of delay times (τ) [187,188]. For a monodisperse sample, the decay in autocorrelation is related to the translational diffusion of particles within the sample according to Eq. 2,

$$|g(\tau)| = exp\ (-qD_c\tau) \quad \text{(Eq. 2)}$$

where g(τ) represents the exponential decay in autocorrelation and q is the light scattering vector, defined according to Eq. 3,

$$q = \frac{4\pi\eta sin\left(\frac{\Theta}{2}\right)}{\lambda} \quad \text{(Eq. 3)}$$

where Θ is the light scattering angle and λ is the wavelength of the light [189]. For polydisperse samples, the decay in autocorrelation of the signal will be polyexponential.

The PDI is related to the width (σ), or standard deviation, of the particle size distribution (Eq. 4) [169,190].

$$PDI = \frac{\sigma^2}{Z_d^2} \quad \text{(Eq. 4)}$$

This parameter is obtained from fitting the autocorrelation function using the Cumulants method[169,191,192], which assumes the sample is monodisperse and the decay in autocorrelation is single exponential.[169] The size obtained from this approach is referred to as the Z-average ($Z_d$) and represents the average $D_h$ of all particles in solution. Therefore, the $Z_d$ is only an accurate quantity when the sample is monodisperse [169,193]. In principle, monodisperse samples show PDI values less than 0.1.[188,190,194] For polydisperse samples, the Cumulants method of fitting is not valid and a multi-exponential fitting approach is required. This method is referred to as the CONTIN method.[169,187,192] It is worth noting that because many different distributions of species can lead to the same decay in autocorrelation when the sample is polydisperse, it is not possible to get precise distributions of similarly sized species that are not resolved by DLS. Therefore, PDI values provide an assessment of polydispersity, but not composition of the sample.[169]

2.5.3 AFM experiments

To prepare the AFM samples, 45 µL of apoSH SOD1 solution was deposited on freshly cleaved mica, with a 0.15 mm thickness and 22 mm diameter (SPI Supplies, Structure Probe Ltd.). The sample was incubated on the bench at room temperature for

5 minutes to let the protein adhere to the mica surface. The excess solution was drawn off the surface of the mica by capillary action using a Kimwipe and then washed four times with 45 µL of milli-Q water using a micropipette. Between each rinse, the excess water was drawn off with a Kimwipe. Samples were then dried under a stream of air and then stored in a plastic Petri dish inside a desiccator to prevent dust or moisture accumulation.

Images were obtained using a JPK Nanowizard II atomic force microscope (JPK Instruments) in intermittent contact mode, where vertical deflection of the cantilever was recorded and used to reconstruct an image of the sample. For all experiments, monolithic silicon AFM tips with a radius of curvature less than 8 nm, force constant of 42 $Nm^{-1}$, and resonance frequency of 320 kHz were used (NCH type, Nanoworld Innovative Technologies). Multiple lower resolution 10 x 10 µm images were taken first (512 x 512 or 1024 x 1024 pixels) followed by multiple higher resolution 2 x 2 or 5 x 5 µm images (1024 x 1024 or 2048 x 2048 pixels).

The higher resolution images were initially analyzed through the automatic processing application within the Scanning Probe Image Processor (SPIP). The Particle and Pore analysis (Image Metrology) application in height mode was used to determine the length, breadth, height, volume, and aspect ratio of the particles detected as well as sample properties such as percent surface coverage, which can be used to compare the extents of aggregate formation in different samples, and sample roughness, which assess how evenly distributed the protein is on the mica. Average coverage estimates obtained by SPIP were also used to gauge the extent of aggregation of the different samples. Coverage is calculated by dividing the sample area (monomers and aggregates) by the

total image area. Samples containing many aggregates will generally exhibit higher coverage on the mica.

SPIP software returns several different measures of particle size, which have differing accuracies depending on the aggregate morphology. Since apoSH SOD1 particles lacked amyloid-like structural features, fiber length and skeleton length were not deemed as accurate parameters to describe the particles. Fiber length is defined as the longest segment within the aggregate skeleton and is found by reducing the particle to a width of one pixel and measuring the longest distance from end to end. The skeleton length, on the other hand includes all branches of the fibril (Figure 2.22). For particles that are highly spherical, the fiber length and skeleton length will be an underestimate of the true length as there will be a significant loss in size at the ends of the aggregate when it is shrunk to one pixel in width. In such cases, fiber width is then overestimated. This was found to be the case for apoSH SOD1 particles; therefore, a different measure of length was used. Length estimates are determined simply by the end to end distance of the particle and breadth is defined as half the distance in the perpendicular direction (Figure 2.22). To obtain an estimate of the number of monomers within each particle, the volume obtained by SPIP software was divided by the volume of the cysteine free apo monomer, ~18.6 nm3, calculated from the crystal structure of the cysteine free apo monomer (2GBU) using the Volume, Area, Dihedral Angle Reporter (VADAR) [121]. The particle size estimates were plotted on violin-box plots using the R program.

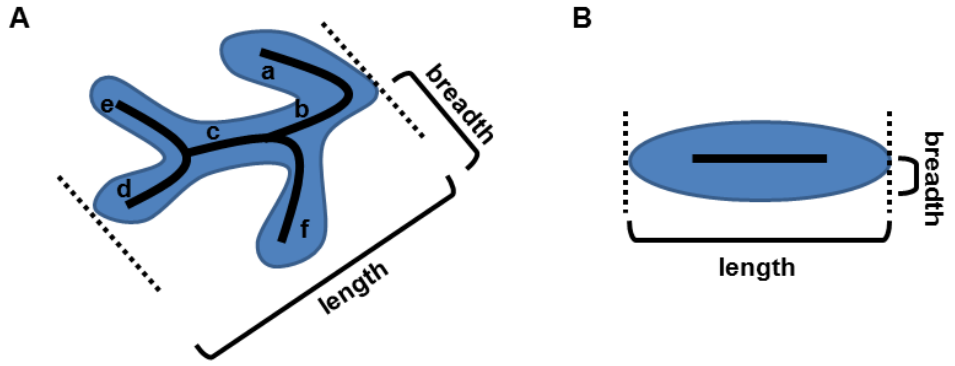



Figure 2.22. Aggregate morphology influences the accuracy of different size measurements. The measured area of the aggregates (blue) is calculated using SPIP, Particle and Pore Analysis software and is based on the number of pixels within the aggregate image (refer to SPIP manual at http://www.imagemet.com/). Fiber length is the longest segment of the aggregate reduced to a width of 1 pixel, shown as a solid black line in (A) and (B). (A) For the branched fibril, fiber length is the sum of segments a, b, c, and d. Skeleton length is the sum of all segments, a, b, c, d, e, and f. (B) For small, more spherical aggregates, reducing the aggregate width to 1 pixel causes a significant loss in aggregate length and an underestimate of the aggregate length. Since the majority of SOD1 particles resembled the type shown in (B), length as opposed to fiber length was used.

To validate the automatic particle size estimates were accurate, the cross-section profile tool in the SPIP software was used to manually measure the length and height of the particles. This was done to a representative number of variants (A4V and G37R) to both aid in visualizing the aggregates observed by DLS and to confirm the Particle and Pore analysis was providing accurate size estimates.

### 2.5.4 ATR-FTIR Spectroscopy

Attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) was performed following established recommendations [195,196] using a Tensor 37 ATR-FTIR spectrometer (Bruker Optics, Germany) continuously purged with N2 gas, equipped with a LN-MCT (mercury-cadmium-telluride) detector cooled with liquid N2. A Bio-ATR II cell with a ZnSe crystal was used to monitor changes in samples with time. For each protein sample, four series of measurements were collected: a background (air), buffer (HEPES and TCEP), protein sample, and post- sample buffer after several buffer washes.

Each measurement consisted of four spectra, each with 256 scans with a spectral range of 4000-1200 $cm^{-1}$ and 4 $cm^{-1}$. Protein samples were spun at 21000 g for 15 minutes and each supernatant and pellet were measured at 20 °C, controlled with a Thermo HAAKDC water bath (Haake).

2.5.5 ThT binding experiments

Aggregated apoSH SOD1 samples (45 µL aliquots) were diluted into ThT buffer, final concentration 25 µM ThT, 50 mM glycine, pH 9.0, as described.[197] The final concentration of total protein was 32 µM and 160 µM for samples incubated at 63 µM and 317 µM, respectively. Fluorescence spectra from 455-575 nm were obtained immediately after dilution of the apoSH SOD1 samples using a first a Photon Technology International QuantaMaster 4 (Horiba-Jobin-Yvon Spex Inc.), and then a Fluorolog3-22 spectrofluorometer (Horiba-Jobin-Yvon Spex Inc.) (all other experiments), with an excitation wavelength of 440 nm. Excitation and emission slit widths of 1 and 5 nm, respectively, were used. The spectra are the accumulated average of 5 scans of one sample.

# Chapter 3: Using amide proton temperature coefficients to probe for local structural perturbations and conformational heterogeneity in immature Superoxide Dismutase-1 variants

Author Contributions:

Harmeen Kaur Deol: conceptualization, methodology, data acquisition, validation, resources, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft; Jeffrey Alexander Palumbo: NMR data acquisition, resources; Dalia Naser: NMR data acquisition, investigation; Duncan Mackenzie: NMR data acquisition, investigation; Michael Vincenzo Tarasca: methodology, NMR data acquisition, resources, investigation; Elizabeth M. Meiering: conceptualization, supervision, project administration, funding acquisition.

## 3.0 Context

With the many aggregate polymorphisms that have been reported for SOD1, it is clear point mutations and solutions conditions influence the aggregate structure. However, it is unclear if either influence protein dynamics and in turn the aggregation pathway(s). Here, amide proton temperature coefficients for numerous ALS-associated SOD1 mutants are investigated to build on the previous framework for holoSS and apoSH pWT temperature coefficient data. The impact of ALS-associated SOD1 mutants in the apoSH state on local structural stability and extent of conformational heterogeneity is

explored. Due to its marginal stability, heating apoSH SOD1 required attention to potential protein unfolding which was explored using both chemical shift perturbations (CSP) as well as by comparing measured chemical shifts to those from urea-denatured SOD1 to assess loss of structure throughout the heat ramp. For certain destabilized mutants, HSQC spectra were acquired both in the folded and unfolding temperature range, relative to their thermal stability, to compare changes in local stability and understand potential dynamics that may be occurring at physiological temperature. The results in this chapter reveal impacts of point mutations on local structure and dynamics throughout a marginally stable protein.

## 3.1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for determining the structure, dynamics and interactions of molecules.[198,199] A 1D NMR spectrum has peaks along one axis, which represent the chemical shift for each nucleus. However, this spectrum is limiting to larger proteins since larger proteins will have more peaks which are likely broader, both factors that increase the chances of peak overlap.[200] Multidimensional NMR decreases overlap by spreading them across another axis and separate overlapping peaks. Accordingly, 2D NMR experiments are commonly used to study protein structure and dynamics. These experiments can either be a homonuclear spectrum, where it is $^{1}H$ with $^{1}H$ frequencies or a heteronuclear spectrum with $^{1}H$ with typically either $^{15}N$, $^{13}C$, or $^{31}P$. A $^{1}H$-$^{15}N$ heteronuclear single quantum correlation

(HSQC) spectrum provides a peak for each covalently-bonded $^{1}H$-$^{15}N$ nuclei pair in the protein. As each amino acid residue (except proline) contains a backbone amide group, this results in a spectrum with a single peak for each amino acid in the protein, the location of which is determined by the chemical environment of each amide.

By collecting an NMR spectrum at a series of temperatures, the temperature dependence of each chemical shift can be measured, from which the linear temperature coefficient is calculated.[201] Amide proton temperature coefficients are represented by the slope of the temperature dependence of the amide proton chemical shift, which previously was thought to report on presence of intramolecular hydrogen bonds in the protein.[199] Additionally, Tomlinson and Williamson found that in the B1 domain of protein G (GB1), structural loss with increasing temperature is a larger contributor to temperature coefficients than solely hydrogen bonds.[202] The results suggested that temperature-dependent loss of structure caused a greater change in amide proton chemical shifts and consequently is a larger determinant of amide proton temperature coefficients.[202] This method was previously applied to determine both the local structural stability of mature (holoSS) superoxide dismutase-1 (SOD1) and conformational heterogeneity due to non-linearity in the temperature dependence.[66,203]

Typically, the temperature dependence of chemical shifts for amide protons are linear; however there are some that exhibit curved temperature dependence which will be referred to as curvature.[66,198,202,204] The curvature occurs as a result of conformational

heterogeneity where the ground state is accessing an alternative low energy state with a distinct temperature dependence.[205] The relative proportion of both states vary along the temperature range, resulting in a curved temperature-dependence.[66,206] The alternative state is in fast exchange on the NMR timescale with the ground state, and within ~5 kcal/mol in energy.[204,207] Prior reports on holoSS pWT curvature suggests the hydrogen bonded residues access an alternative conformation and the alternative state is forming a different bonding pattern.[66]

Additionally, the temperature coefficients for holoSS pWT (pseudo-wildtype) SOD1 revealed that the global stability is high, and as the temperature increases, the structure disruption starts at the periphery of the protein.[66] When holoSS pWT is compared to apoSH pWT, certain dimer interface residues were more structured in apoSH pWT, while the regions that undergo post-translational modifications were less structured in the immature state. Chemical shift changes also correlated to local structural perturbations around the site of mutation, providing evidence for changes in the global protein stability through propagating effects of the mutation on nearby residues.[66] For instance, holoSS G93A in comparison to pWT showed large local perturbation around the site of mutation, but also smaller changes further from the mutation site. Destabilizing mutants also had the greatest portion of residues accessing low free-energy alternative states, whereas the slightly stabilizing mutant (i.e. V148I) had a similar proportion to pWT.

Here we explore how ALS-associated SOD1 mutants in the apoSH state affect local structural stability and extent of conformational heterogeneity. Due to its marginal stability, heating apoSH SOD1 required attention to potential protein unfolding which was explored using both chemical shift perturbations (CSP) as well as by comparing chemical shifts to urea-denatured SOD1 to assess loss of structure of the protein throughout the heat ramp. For certain destabilized mutants, HSQC spectra were acquired in temperature ranges where the protein was predominately folded and at higher temperatures where global unfolding becomes significant; by using the folded and unfolding temperature ranges, the changes in local stability as well as understanding potential dynamics that may be occurring at physiological temperature are investigated. Overall, the data presented here focuses on exploring the impact point mutations have on the local structural environment for a marginally stable protein and assessing dynamic changes that may be relevant at physiological temperatures.

## 3.2 Methods

### 3.2.1 Protein preparation for NMR

SOD1 NMR samples were prepared as previously described.[65,66] Proteins for NMR samples were expressed using M9 minimal media. Briefly, BL21 E. *coli* were used, which do not confer any antibiotic resistance. Once transformed with the pHSOD1ASlacIq, plasmid for expressing SOD1, which confers ampicillin resistance, cells were grown overnight in LB broth with a final concentration of 100 μg $mL^{-1}$ ampicillin. Next, 10 mL of

the overnight culture was used to inoculate 1 L of M9 minimal media, which contains $^{15}NH_4Cl$ as the sole nitrogen source. Cells were grown to OD600=~0.6, and then induced with IPTG (1 mM final concentration). All protein samples underwent a copper charging and heat treatment before being purified by hydrophobic-interaction column chromatography to ensure proper metalation. Copper sulfate was added (final concentration of 1.4 mM) and then the sample were heated for 20 minutes at 70 °C. The solution had ammonium sulfate (final concentration of 3 M in 20 mM Tris at pH 7.8) added and then was filtered through a nitrocellulose filter (0.45 µm cutoff) prior to application onto a hydrophobic-interaction column. The protein was eluted using gradual increasing gradient of 20 mM Tris buffer at pH 7.8 and placed in dialysis tubing (6-8 kDa) to be dialyzed against milliQ $H_2O$ for 4x8-hour exchanges. With the Amicon ultrafiltration membrane, the volume of the sample was reduced, and the protein concentration was measured with the use of the UV-vis spectrophotometer and the molar extinction coefficient of 5,300 $M^{-1}$ $cm^{-1}$ for SOD1 monomer at 280 nm. Aliquots were then syringe-filtered into Eppendorf tubes and flash frozen in liquid nitrogen. Metals were removed via dialysis of holo SOD1 against ethylenediaminetetraacetic acid (EDTA) at pH 3.8.[208] Another set of exchanges were done with 50 mM sodium acetate, 100 mM NaCl at a pH of 3.8. The last set of exchanges were against milliQ water with one final exchange against 1 mM HEPES at pH 7.8. Metal contamination was assessed by PAR assay, initially developed by Crow *et al*.[209] and further modified by Colleen Doyle and Heather Primmer as described previously.[210]

Reduction of the disulfide bond was performed as described previously.[65] The method used the reducing agent tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl). Initially, oxidized apo SOD1 was denatured using 2 M guanidine hydrochloride (GdnHCl) and 20 mM HEPES for a pH of 7.8. TCEP was then added to the sample, and it was placed in a desiccator for one hour before the sample was exchanged into buffer (1 mM TCEP and 20 mM HEPES at pH 7.4). The success of the reduction was confirmed by running an SDS-PAGE gel. Preparation of the NMR sample required concentration of the sample to ~15-20 mg/mL SOD1, with a final sample volume of ~450 mL. 50 µL of $D_2O$ and sodium trimethylsilylpropanesulfonate (DSS) (final concentration of 1 mM) was added and the sample was transferred to an NMR tube (final sample volume of 500 mL). Upon final sample preparation, we also employed an initial $^{1}H$ 1D NMR experiment (p3919gp, Appendix H) to confirm both the TCEP reduction state and the metalation status (Appendix D and E).

3.2.2 NMR experiments

$^{1}H$-$^{15}N$ HSQC spectra were acquired at varying temperatures from 1.2 °C to 38.6 °C using a Bruker Avance 600 MHz spectrometer, using the acquisition parameters described in Appendix J. The sample was equilibrated for 20 minutes at each temperature prior to spectra acquisition, as described by Doyle *et al*.[66] All spectra were referenced to DSS. Spectra were processed in TopSpin 4.0.7, and analyzed in CCPN 2.4.2. Sequence specific resonances were previously determined for pWT, A4V, G85R, E100G, and V148I apoSH SOD1 at pH 7.4 at 298 K.[54,55]

All chemical shift perturbation values were calculated using the following equation, as previously done[55]:

$$\Delta\boldsymbol{\delta}_{\boldsymbol{H,N}} = \sqrt{\left[(\Delta\boldsymbol{\delta}_{\boldsymbol{H}})^{\mathbf{2}} + \frac{\mathbf{1}}{\mathbf{5}}(\Delta\boldsymbol{\delta}_{\boldsymbol{N}})^{\mathbf{2}}\right]}$$

For each residue, the slope of the chemical shifts for amide protons against temperature represents the temperature coefficient and were determined using CCPN 2.4.2. The difference between the observed chemical shift and the slope (the residual chemical shift) were calculated, plotted as a function of temperature and fit to a quadratic equation to determine the extent of curvature using R scripts previously developed in Meiering lab.[211] All protons for which the quadratic fit had a $R^2 > 0.8$ and a p-value mean $< 0.05$ are considered curved temperature-dependences. Chemical shift perturbations were analyzed as previously done.[66]

## 3.3 Results and Discussion

Unlike the highly stable holoSS SOD1, apoSH SOD1 is marginally stable and cannot be heated past physiological temperatures without protein unfolding occurring. The mutants chosen varied in location, disease duration, and most importantly, stability (Table 3.1 and Figure 3.1).[65] The variation in thermal stability means the folded temperature range varied significantly; stabilized mutants like V148I had a relatively large, folded range similar to pWT, but destabilized mutants like A4V had a smaller range. However, to ensure enough data points were acquired, apoSH mutants were heated to a

temperature just below their onset of global thermal denaturation with more intermediate temperature points collected along the folded temperature range. To make valid comparisons between mutants, the same temperature range (9-18 °C) was used to compare across mutants unless otherwise stated.

Table 3.1: Summary of SOD1 mutant characteristics.

| Mutants | Location | Disease Duration (years)[a] | ApoSH $T_{0.5}$ (°C)[b] | HoloSS $T_{0.5}$ (°C)[c] |
|---|---|---|---|---|
| **pWT** | - | - | 48.5 | 92.7 |
| **A4V** | ß1 | 1.2 | 36.3 | 86.7 |
| **G85R** | ß4 | 6.0 | 40.7 | 77.5 |
| **E100G** | ß6 | 4.7 | 33.2 | 86.2 |
| **V148G** | ß8 | 2.1 | 34.0[d] | 86.9 |
| **V148I** | ß8 | 1.7 | 51.0 | 92.7 |

[a]Average disease durations are from Wang *et al*.[212]

[b]Melting temperatures are from Vassall *et al*.[65]

[c]Melting temperatures are from Stathopolous *et al*[52] and Doyle *et al*.[66]

[d]Based on visual inspection.

For the destabilized mutants where spectra are collected at their unfolding temperatures, the differences between local and global unfolding are vital to define. At temperatures corresponding to global unfolding based on their DSC endotherms[65], the protein experiences both global and local unfolding. The global unfolding refers to the native, folded state transitioning to its completely unfolded state. In addition to global unfolding, the protein will also undergo structural changes due to local unfolding of the native, folded state throughout the unfolding temperatures. Here, any small, local

changes will be considered a consequence of local unfolding, whereas global unfolding will only be considered if the entire structure and/or if the residues behaviour correlates to their respective DSC endotherm (e.g. the peak shape begins to change as the protein enters its protein unfolding temperatures).[65]

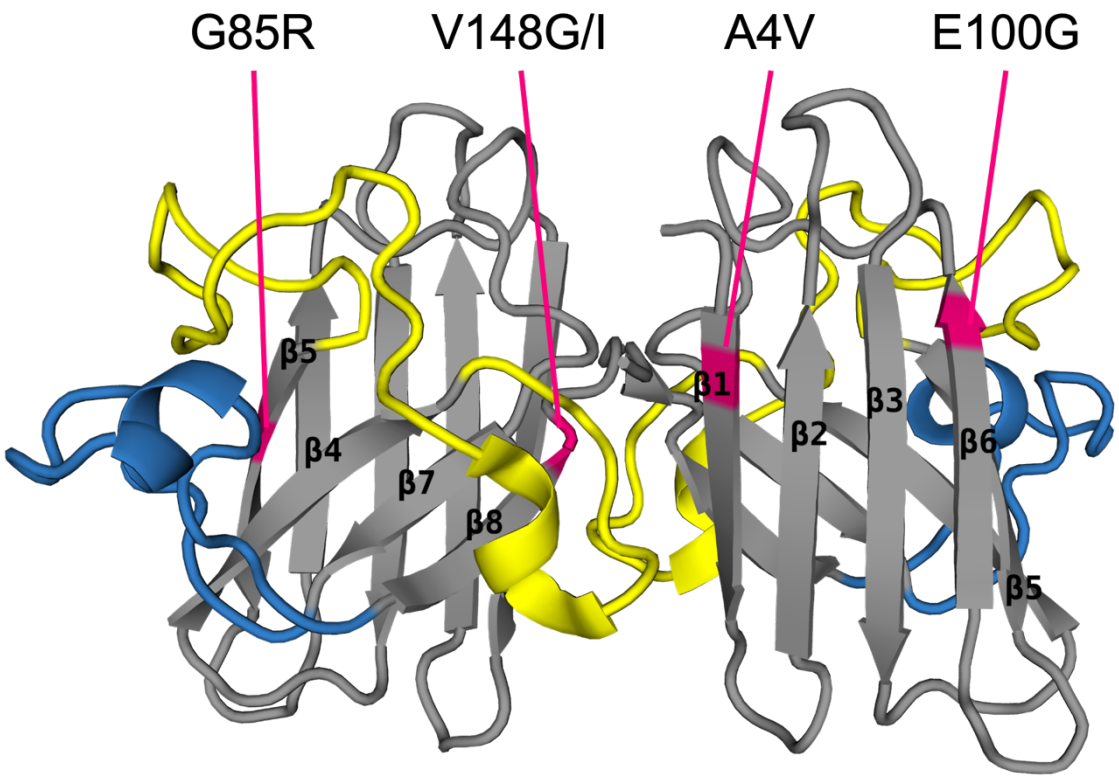



Figure 3.1: SOD1 mutations mapped onto holoSS SOD1. Mutations studied here are mapped onto the 3D structure of holoSS SOD1 in PyMOL (PDB 1HL5, rendered with PyMol) in magenta. A4V, G85R, E100G and V148G/I are in ß1, ß4, ß6 and ß8, respectively. Loops 4 and 7 become structured with maturation events, and are coloured yellow and blue, respectively.

### 3.3.1 Chemical Shift Perturbations (CSP)

Chemical shift perturbations (CSP) can be used to assess the effects of mutations, ligand titration, changes in solution conditions, along with many others on the structure.[213] CSPs can report on changes in the local environment as a result of the changed condition, which highlight critically impacted regions. Here, backbone chemical shifts are used to assess the effect mutations have on the apoSH SOD1 structure under identical solution conditions. Backbone chemical shifts are calculated from the weighted average difference in amide nitrogen chemical shift ($\delta_{NH}$) and amide proton chemical shift ($\delta_{H}$) between mutants and pWT (see Methods), as previously done.[55,66] The closer (or further) the backbone chemical shift is relative to 0, the more structurally similar (or different) the mutant is to pWT.

Due to the range of stability in the 5 mutants assessed (Table 3.1), the temperature at which the CSP is calculated is important. For more stable proteins like holoSS SOD1, CSPs can be calculated at ~298 K (~25 °C); however, because some of the destabilized mutants start unfolding at temperatures as low as 20 °C[65], areas of local protein unfolding for these mutants may show higher CSPs. For this reason, CSP values are calculated at both 19 °C and 25 °C. The CSP values at 19 °C should not reflect any changes due to thermal protein unfolding for any of the mutants; instead, significant CSP values at 19 °C are dominated by changes in the local environment due to the point mutation.

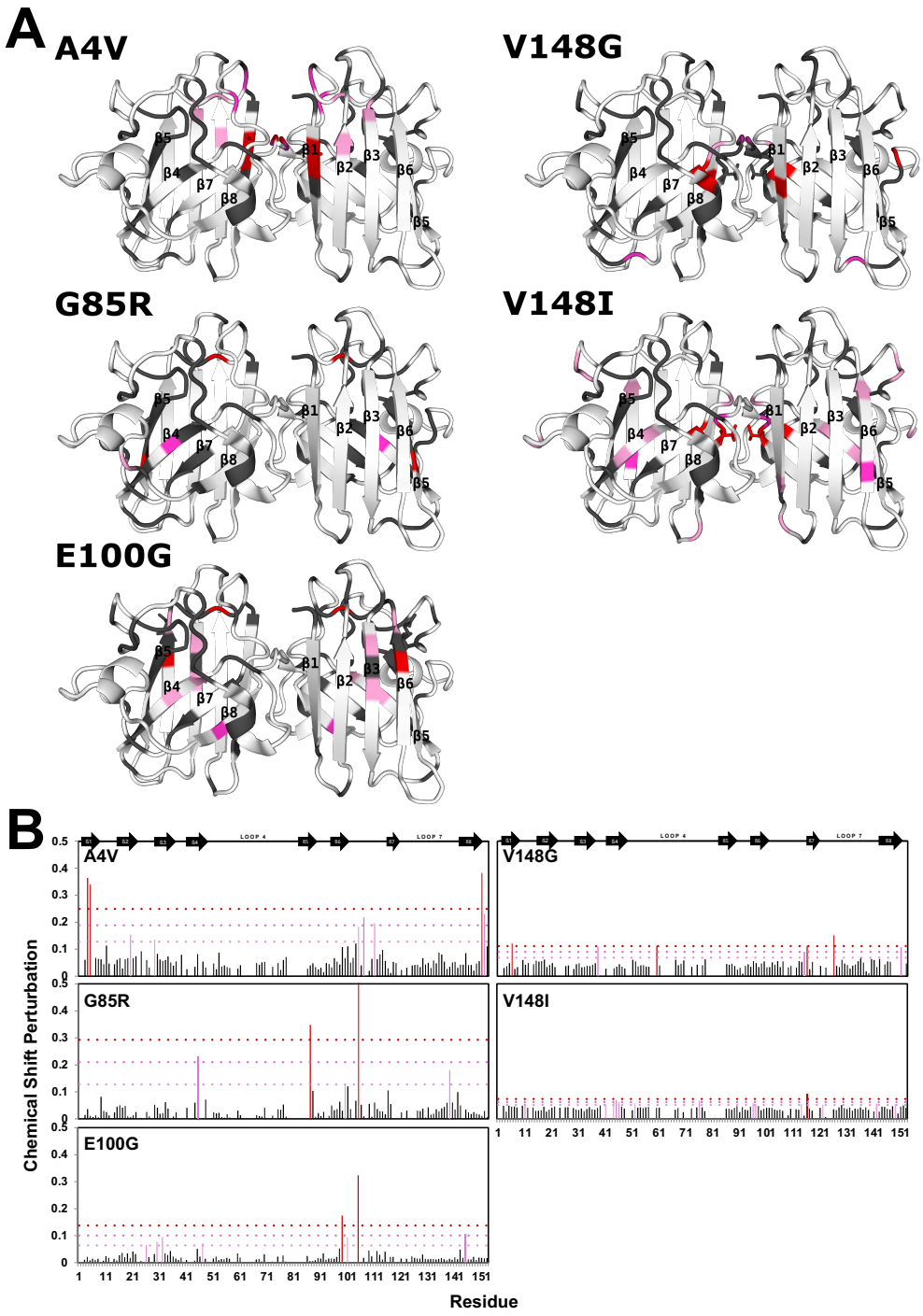



Figure 3.2: CSP for folded apoSH mutants relative to apoSH pWT at 19 °C. CSP values were coloured based on the standard deviations for each respective mutant. Residues are coloured to aid in visualization according to the relative magnitude of CSP values. The pink, dark pink and red colours/lines represent average + 1 standard deviation, average + 2 standard deviations, and average + 3 standard deviations, respectively. (A) CSP values are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). White denotes CSP values smaller than average + 1 standard deviation, and grey indicates residues where no data are available. (B) CSP values are shown as vertical bars, where black bars are

values that are smaller than average + 1 standard deviation. The beta-strands based on 1HL5 are shown above the plots.

3.3.1.2 ApoSH mutants have high CSP values near the site of mutation, at 19 °C

Along with CSP values, the standard deviation of CSP values across mutants is another useful measure for assessing the extent of structural impact of point mutations. CSP changes considered per mutant and between mutants are mapped onto the protein structure in Figure 3.2A and 3.3A, respectively. From CSP plots at 19 °C (Figure 3.2A), V148I behaves most pWT-like with very minimal perturbations and has a small range of standard deviation for CSP values. However, the four destabilized mutants show much larger standard deviations, which can be explained by the large CSP values for residues near the site of mutation. Some mutants consistently show similar values for all residues for the entire sequence. This is likely due to small differences in temperature between the mutant and pWT sample, which cause shifts in chemical shift values for all residues. Interestingly, A4V also consistently shows CSP values greater than 0 throughout the sequence, but unlike for V148I and V148G which have similar CSP values for most residues aside for near the site of mutation, A4V CSP values vary throughout. This suggests that A4V could be forming a different native structure, an observation that was previously reported[214] and that is further supported below.

Furthermore, mapping the CSP values at 19 °C on the protein structure using the standard deviation cutoffs for each mutant helps identify areas of significant perturbation (Figure 3.2 B). Although the apoSH SOD1 state exists predominantly as a monomer and is not as structured as the structure of dimeric SOD1 with intrasubunit disulfide bond

formed (PDB 1HL5, file used to illustrates the changes in the figures here). Due to the limited reliability of the current possible apoSH SOD1 structure (PDB 2GBU) available (discussed in Chapter 1), this structure remains useful for assessing changes in structure in areas relevant to SOD1 maturation. For all mutants, residues structurally near the site of mutation are perturbed consistently with smaller changes propagating throughout the structure (Figure 3.2B). V148I also shows differences in ß4 relative to pWT, which is interesting as those residues in ß4 are not near the site of mutation. For an otherwise pWT-like mutant, this subtle change could be suggesting local structural differences which may have consequences on the protein dynamics.

To allow for relative comparison of CSP values between mutants, all values were binned to a global cutoff for all mutants (Figure 3.3). It is important to note, by comparing mutants with a global cutoff, subtle differences for mutants that show smaller changes will no longer be considered significantly different. However, by now using a global cutoff, the extent of differences between mutants to pWT can be compared. Unsurprisingly, V148I shows the fewest changes relative to pWT, whereas A4V has most of its structure coloured as significantly perturbed. This finding can support many theories. Most substantially, it supports that A4V may not be accessing a native-like fold whereas the rest of the mutants are forming native-like structures with different extents of perturbations around the site of mutation. The alternative theory could be A4V is actively self-associating, however, the complementary DLS data suggests this is not the case.

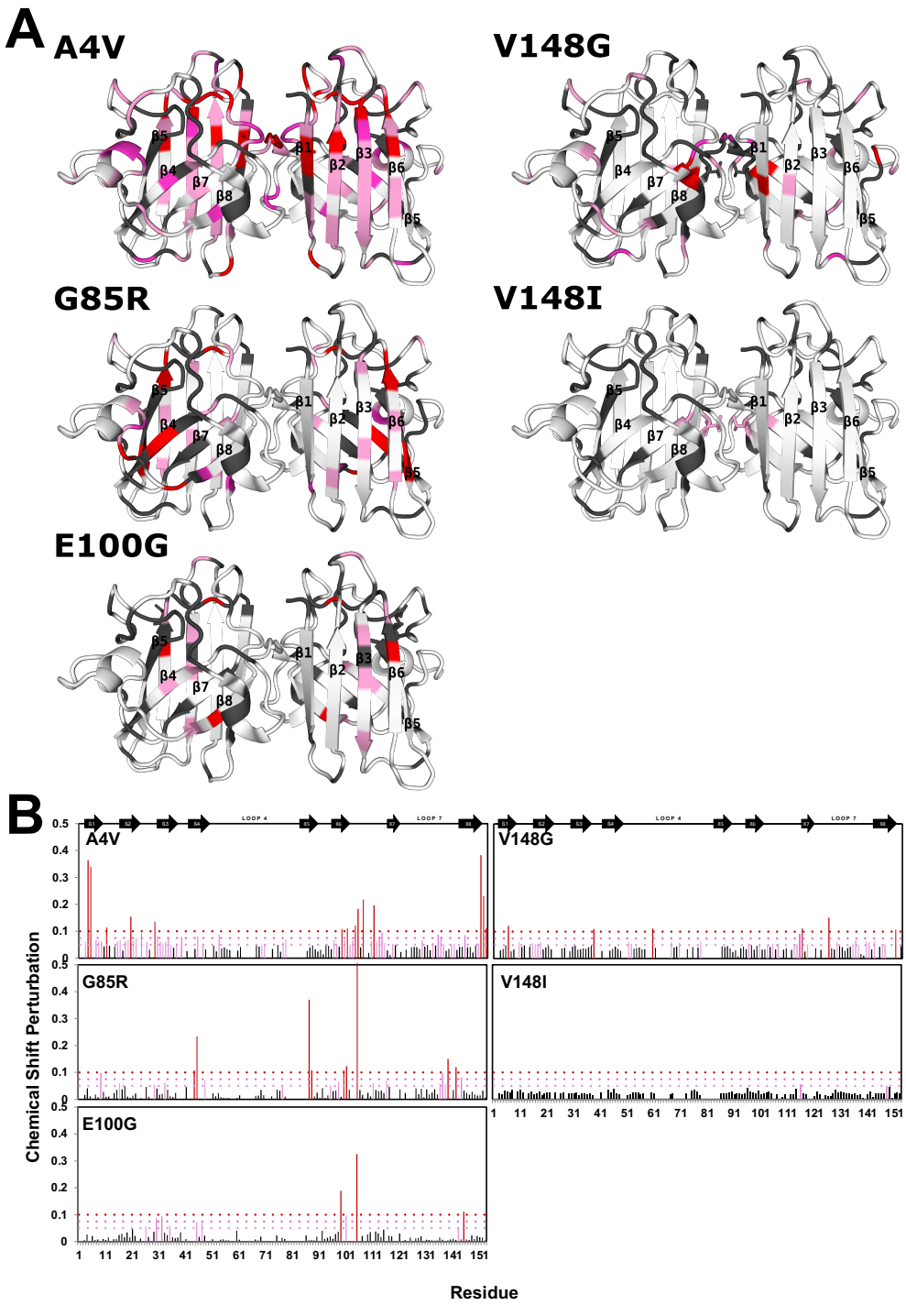



Figure 3.3: CSP for folded apoSH mutants relative to apoSH pWT at 19 °C (global cutoffs). CSP values were coloured based on the same cutoff for all mutants. Residues are coloured to aid in visualization according to the relative magnitude of CSP values. The pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1, respectively. (A) CSP values are mapped onto the 3D ribbon structure for

holoSS SOD1 (1HL5, rendered with PyMol). White denotes CSP values smaller than 0.05, and grey indicates residues where no data are available. (B) CSP values are shown as vertical bars, where black bars are values that are smaller than 0.05. The beta-strands based on 1HL5 are shown above the plots.

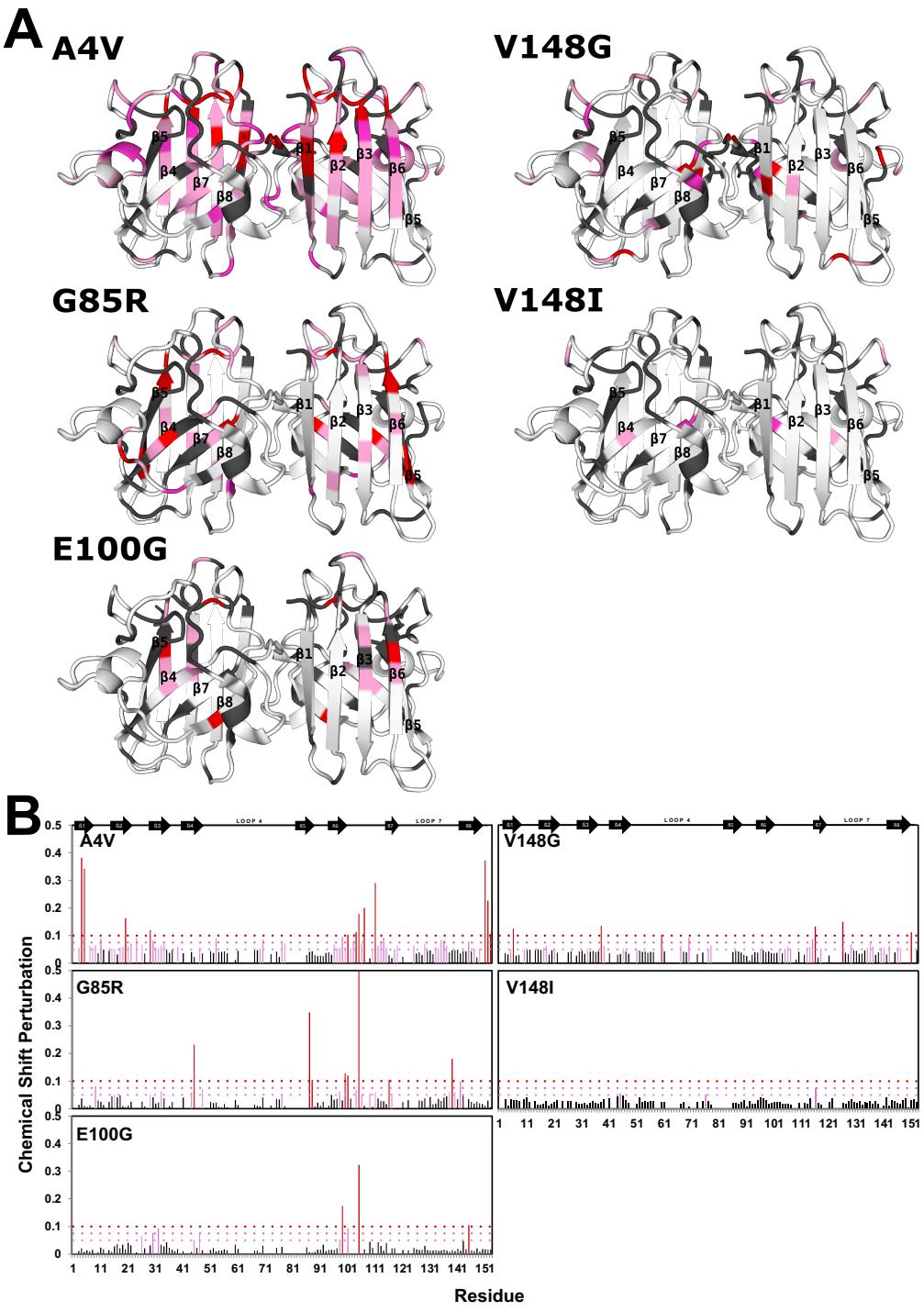



Figure 3.4: CSP for apoSH mutants relative to apoSH pWT at 25 °C (global cutoffs). CSP values were coloured based on the same cutoff for all mutants. Residues are coloured to aid in visualization according

to the relative magnitude of CSP values. The pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1, respectively. (A) CSP values are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). White denotes CSP values smaller than 0.05, and grey indicates residues where no data are available. (B) CSP values are shown as vertical bars, where black bars are values that are smaller than 0.05.

#### 3.3.1.2 CSP values for apoSH mutant at 25 °C can highlight local restructuring due to unfolding

Moreover, the same global cutoff is applied to CSP values calculated at 25 °C (Figure 3.4). Most of the mutants do not show significant changes or increases in perturbation, despite A4V and E100G apparently beginning to unfold according to their DSC thermograms.[65] However, V148G, a mutant that also starts to unfold by 25 °C, shows increased areas of perturbation around the native-helix in loop 7, an area that helps stabilize metal binding.[215] Although it is limited to a few residues, these perturbations reflect an area of V148G that becomes highly dynamic or unfolded prior to the rest of the protein. This is of particular interest since in Chapter 7, our data suggest the zinc shuffling from the unfolded zinc-bound SOD1 to the second metal-binding site on a folded zinc-bound SOD1, due to its higher zinc affinity, is perturbed in V148G. This was most fascinating since only metal-binding mutants were expected to lack zinc shuffling, and V148G is not considered a metal-binding mutant because apoSH V148G showed comparable zinc affinity to pWT.

Now, we can hypothesize that the lack of zinc shuffling observed in the DSC endotherm may be due to local unfolding initiating around the metal-binding site, such that upon availability of zinc from the unfolded zinc-bound state, the second binding site on the “folded” zinc-bound is no longer formed. This is also supported by *in vivo* aggregation propensity data for apoSH V148G.[25] In these *in vivo* experiments, the

aggregation propensity is high at both 25 and 37 °C. Upon addition of zinc, the 25 °C and 37 °C samples experienced a 51% and 12% rescue from aggregation, respectively. This suggests that folded V148G apoSH has intact metal-binding sites, and can experience the stabilizing impact of the zinc binding at 25 °C. However, at 37 °C the metal-binding sites are likely unfolded, and hence V148G cannot experience the full stabilizing impact of zinc binding, and thus do not experience aggregation rescue. In addition to the local unfolding impacting metal binding and potentially protein maturation, these early increased dynamics could also serve as an aggregation hotspot.

Overall, the trend observed for apoSH SOD1 mutants where sites of mutations are most perturbed relative to pWT aligns with the observation made for the holoSS SOD1 state.[66] In holoSS mutants, the sites of mutation were typically most perturbed with effects propagating to neighbouring residues. Prior studies also suggest that in the apoSH SOD1, chemical shift differences between mutants and pWT are localized to mostly the site of mutation.[55,66] In addition, their chemical shift difference profiles determined from TROSY-HSQC spectra align strongly with the CSP plots described here, specifically enhanced differences throughout the A4V sample. Although Sekhar *et al.* do not comment on A4V's distinct profile, our data here along with results Sekhar *et al*., 2016, suggest there may be less native-like structure present in the apoSH state of A4V.[55] In summary, the CSP data suggest there are significant structural perturbations in the apoSH mutants around the site of mutation with propagating effects to neighbouring residues.

### 3.3.2 Extent of structure along the folded temperature range

While the temperature coefficient data collected are well in the folded temperature range of all the mutants, given the diverse nature of the mutations, understanding the extent that the apoSH state and apoSH mutants are structured is important for interpreting temperature coefficients. Individual mutations along with lack of key maturation events can disrupt or create structure, which may not be present in the fully mature control. Here, urea-denatured SOD1 chemical shifts collected at 25 °C are used as a reference for unfolded SOD1[216], to which calculated random coil temperature coefficient values for SOD1 sequence are applied to extrapolate the expected random coil chemical shifts at different temperatures (Appendix C, see Chapter 5).[217]

The experimental values at the lowest (~9 °C) and highest temperature (~18 °C) along the folded temperature range are compared to the calculated unfolded SOD1 chemical shifts at the same temperatures. If the chemical shift difference between the calculated and experimental is smaller at 18 °C than at 9 °C, the residue is becoming random coil-like throughout the temperature range. If the chemical shift difference does not decrease with temperature, the residue is considered not shifting to random coil-like in the temperature range. Some assumptions are required for this method to be valid, including that the thermally unfolded chemical shifts would be similar to the chemically denatured chemical shifts, the calculated random coil SOD1 temperature coefficients are the same regardless of solution conditions, and unfolded SOD1 behaves random coil-like.

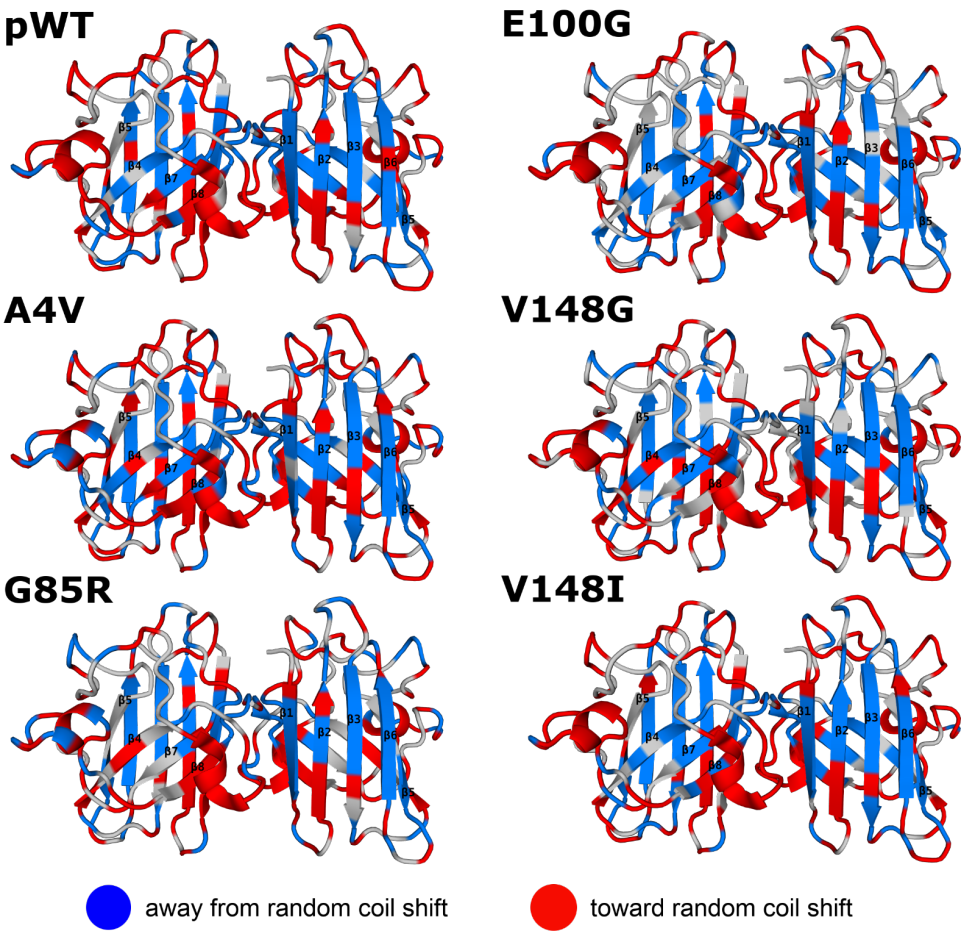



Figure 3.5: Comparing experimental chemical shifts in the folded temperature range (9–18 °C) to unfolded SOD1 chemical shifts. Chemical shift values are compared to unfolded SOD1 chemical shifts; calculated unfolded SOD1 random coil temperature coefficients are used to extrapolate the unfolded SOD1 chemical shifts to different temperatures. If chemical shift differences between the folded and unfolded chemical shifts at 9 and 18 °C decrease, it is interpreted as the residue is becoming unfolded/random coil-like (red). However, absence of decrease means the chemical shift is not moving toward unfolded/random coil (blue). Grey indicates residues where no data are available.

Along the folded temperature range, general trends on extent of structure emerged. Most of the ß-strand residues do not shift closer to their calculated random coil shifts, whereas the large loops (loop 4 and 7) and the ends of ß-strands (ß2, ß3) shift toward their random coil shifts (Figure 3.5). Additionally, most residues near the dimer interface (e.g. ß8) are shifted toward their random coil chemical shifts. These trends largely hold throughout the mutant set, however there are smaller differences, but due to

the many assumptions, their significance cannot be validated. Instead, the data can be interpreted on a more global scale, where residues that move toward random coil chemical shifts are considered becoming unstructured, and residues that do not, remain structured. Although the lack of decreased differences in chemical shift values do not inherently imply presence of structure, and instead negligible differences could suggest random coil structure at both temperatures, the interpretation here is validated by literature findings (discussed below).

The above trends suggest that most of the ß-strands are structured, albeit shortened for ß2 and ß3, and the loops and dimer interface are becoming or may be less structured throughout the folded temperature range. Similarly, previous findings have reported that most apoSH mutants have the native ß-barrel present but loops 4 and 7 are highly dynamic.[55] These large loops are involved in critical maturation steps such as metal-binding and dimerization, and in the mature enzyme are more structured. However, without those maturation events, the loops in the apoSH state are understandably less structured[54], and here show more random coil-like behaviour with temperature. Altogether, a comparison of the experimental apoSH chemical shifts in the folded range to that of the chemically denatured SOD1 suggests the ß-barrel is mostly intact with less structure formation in areas where maturation events have not yet taken place.

### 3.3.3 Proton amide temperature coefficients for ApoSH SOD1 mutants

The temperature coefficient data reported here are for the folded temperature range from 9–18 °C, as in sections 3.3.1 and 3.3.2. In the temperature coefficient figures, temperature coefficient data are binned according to prior cutoffs[66,203], where warmer (cooler) colours indicate large (small) absolute temperature coefficient values, which correspond to more dynamic (more structured) areas. Most temperature coefficients are negative, and here the smaller (i.e. more negative) temperature coefficients are referred to as large temperature coefficients due to their large absolute temperature coefficient value, whereas values closer to 0 are considered small regardless of their sign. The goal of collecting the temperature coefficient data for apoSH mutants are to offer insight on how the lack of maturation of apoSH SOD1 along with point mutations impact the local structural stability. Could areas of local change suggest critical regions responsible for protein misfolding and subsequently initiating aggregation?

Analysis of temperature coefficient data can be viewed through multiple lenses. Binning the data initially is important for deciphering patterns but due to the wide range of temperature coefficient values, multiple bin sizes are necessary (Figure 3.6). However, with multiple bin sizes, trends are not as easily discernible. Previously, holoSS SOD1 temperature coefficients could be reliably averaged based on secondary structure due to a sound crystal structure. On the other hand, when a reliable structure is not available for a protein, such as apoSH SOD1, multiple types of temperature coefficient averaging must be considered when mapping onto a homologous structure such as holoSS SOD1; with

multiple types of averaging, limitations due to structural inconsistences between native apoSH SOD1 and the holoSS SOD1 structure used for value mapping can be minimized due to different structural assumptions. In this section, apoSH mutant temperature coefficients will be assessed by looking at rolling averages, averaging based on structured components, and averaging based on hydrogen bonds of the structured components expected in the mature form (holoSS).

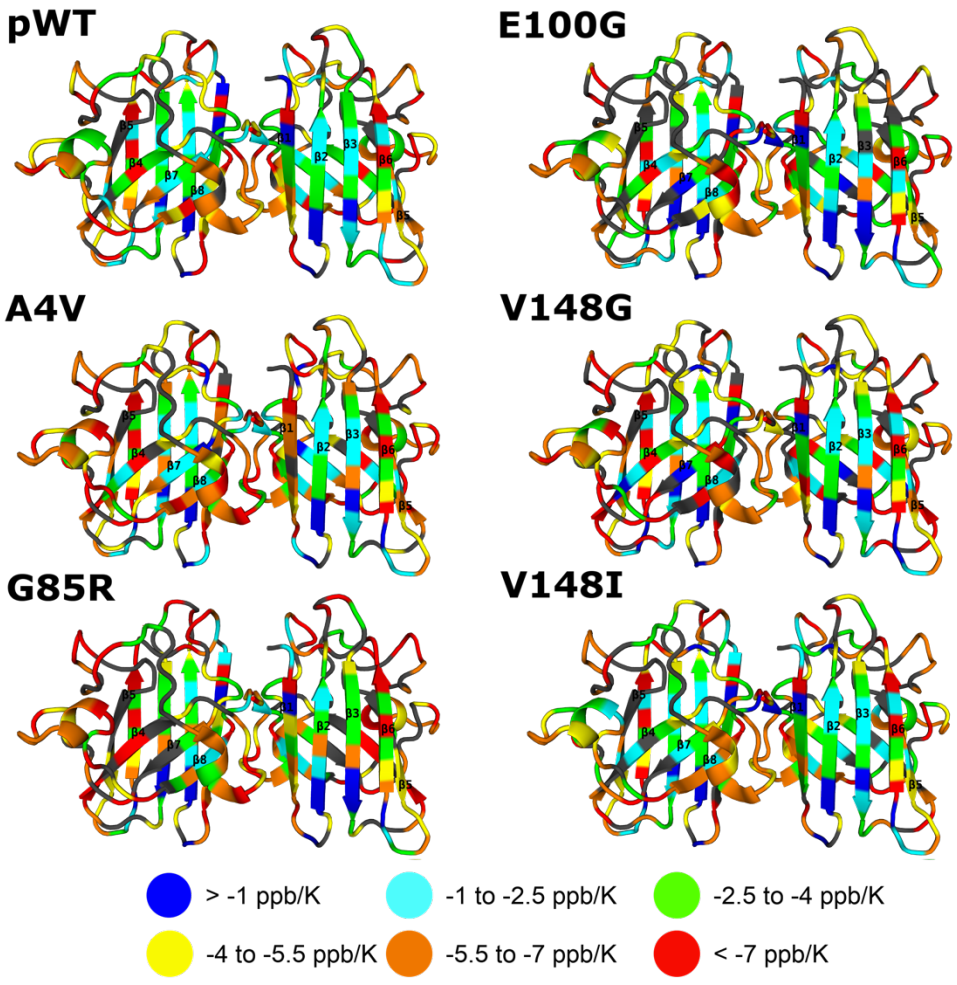



Figure 3.6: Amide proton temperature coefficients for apoSH pWT and mutants from 9–18 °C at pH 7.4. Temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol).

Residues are coloured by a red to blue gradient, with red representing less stable residues and blue the more stable. The colour scheme is given in the legend. Grey indicates residues where no data are available.

#### 3.3.3.1 Rolling averages suggest the temperature coefficients reflect values are grouped based on their secondary structural element

Since assessing individual residues and their temperature coefficients is difficult to deconvolute for trends, various averaging methods can be applied instead (Figure 3.6). Rolling averages is a type of averaging used to assess the mean over n values. By using rolling averages over five residues, local structural differences can be easily discerned here. Fortunately, since the rolling average is over multiple residues, outliers have a lower impact on the average temperature coefficient value. Thus, rolling averages are useful for resolving small changes in temperature coefficients throughout the protein, and identifying local differences in apoSH mutants without assuming any secondary structural elements that may or may not be present.

The rolling averages highlight that the temperature coefficients for SOD1 residues have grouped behavior that is consistent with known SOD1 secondary structural components, aside from some outliers (Figure 3.7). Residues near or involved in ß-strands show lower temperature coefficient values whereas residues near or in loops have higher temperature coefficient values. In pWT, ß2, ß3, and ß7 have the lowest temperature coefficients with cool colours (Figure 3.7), and ß1, ß4-ß6, ß8 and loops 4 and 7 have higher temperature coefficients. Although, when the data are shown as a dimer, ß1 and ß8 should not be considered as dynamic as the values suggest, however

the apoSH state exists predominantly as a monomer. Based on these results, ß1, ß4-ß6, and ß8 are considered edge strands, and ß2, ß3, and ß7 are considered core strands.

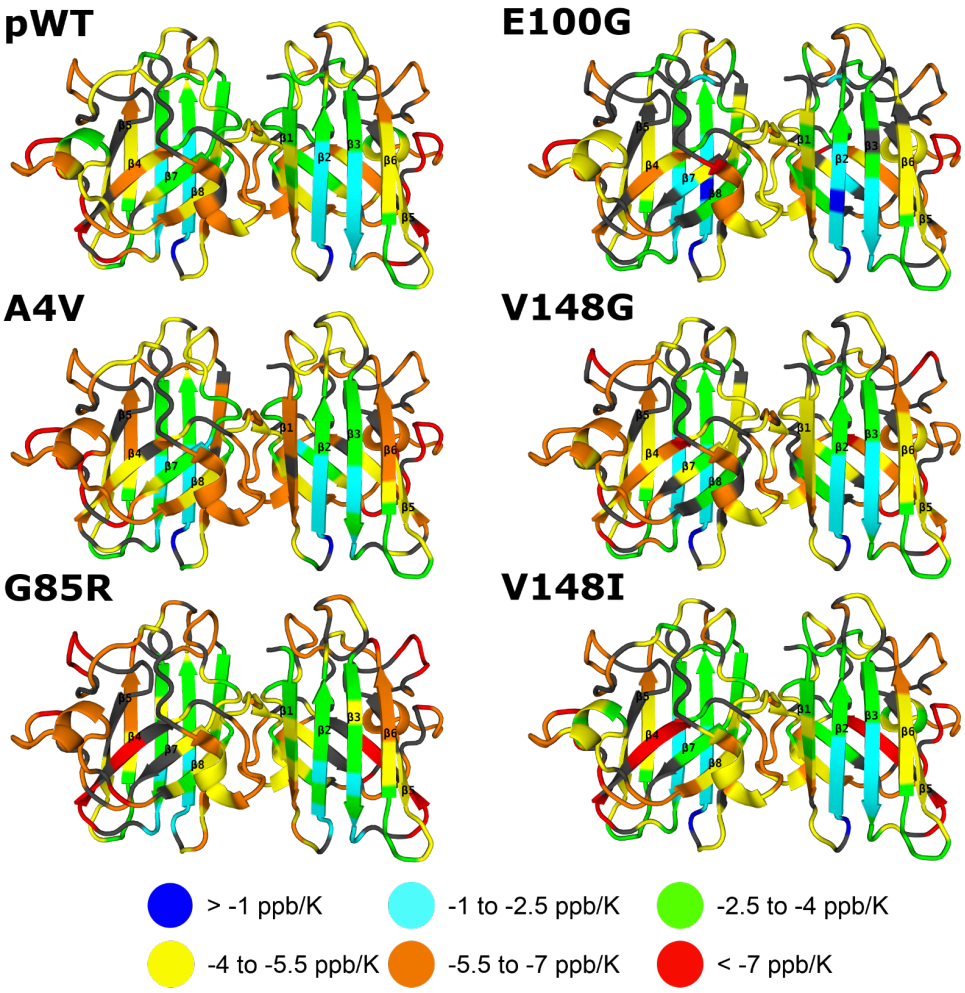



Figure 3.7: Four-residue rolling average amide proton temperature coefficients for apoSH pWT and mutants from 9–18 °C at pH 7.4. Rolling average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue the more stable. The colour scheme is given in the legend. Grey indicates residues where no data are available.

In addition to the periphery of the protein having lower structural stability than the core strands, rolling averages are useful for identifying small, local areas of perturbation. Each mutant shows subtle differences throughout the structure, differences are considered notable only if their difference relative to pWT is +/- 1 ppb/K (Figure 3.8).

V148I temperature coefficients resemble pWT-like behaviour except for differences in some residues near and in loop 7, as well as the dimer interface. This dispersed pattern of both higher structural stability (smaller temperature coefficient than pWT) and lower structural stability residues (larger temperature coefficient than pWT) throughout these regions could reflect a low population of dimer formed by apoSH V148I, as has previously been reported for apoSH V148I.[12,54] In contrast, V148G shows changes, mostly increases in temperature coefficients, throughout the structure. These changes are consistent with decreased global stability of the V148G monomer and V148G consistently having a weakened dimer interface.[53,65,214] Overall, it is remarkable to see how differently point mutations at the same position manifest.

On the other hand, A4V, G85R and E100G are different mutations at different locations but share similar perturbations in temperature coefficients. For instance, the residues near or involved in forming the native alpha-helix in loop 7 in the mature enzyme are more dynamic in these mutants than in pWT. Notably, these mutants have previously shown decreased formation of this helix relative to apoSH pWT[66], which is consistent with the enhanced dynamics of these residues seen here. Additionally, in A4V and G85R we see enhanced dynamics at the end of ß6 and in loop 6, but despite the location of E100G at the end of ß6, it does not show differences relative to pWT. In summary, the rolling averages of temperature coefficients are important for finding small local changes throughout the protein and can highlight trends otherwise overlooked by individual or average temperature coefficients averaged by structural components.

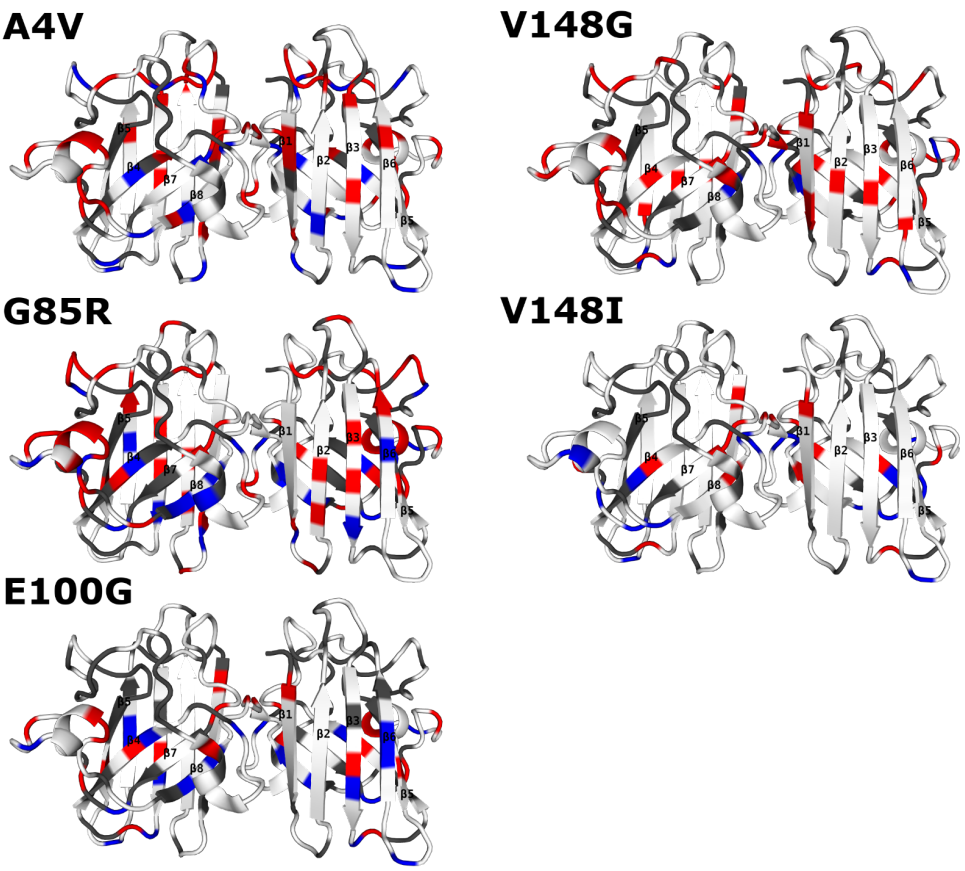



Figure 3.8: Temperature coefficient differences between apoSH mutants and pWT. Differences in individual temperature coefficients in the folded range (9–18 °C) are calculated between the mutants and pWT. Residues are coloured based on their differences, and absolute differences greater than 1 ppb/K are considered significant. If the absolute temperature coefficient value in the mutant is larger than 1 ppb/K relative to pWT, then the residue is less stable in the mutant (red). However, if the absolute temperature coefficient value for the mutant is smaller than - 1 ppb/K then the residue is more stable in the mutant (blue). The white residues represent absolute differences smaller than 1 ppb/K, and grey indicates residues where no data are available.

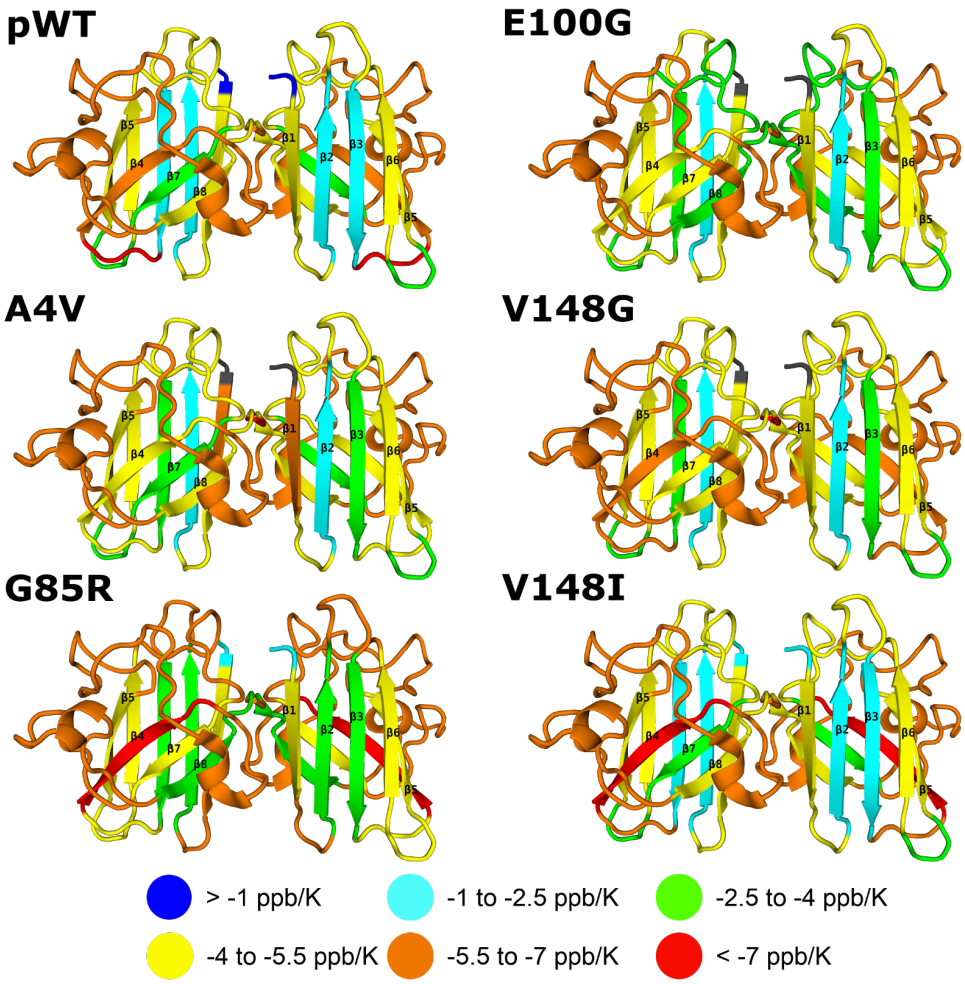



Figure 3.9: Secondary structural average amide proton temperature coefficients for apoSH pWT and mutants from 9–18 °C at pH 7.4. Secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue for the more stable. The colour scheme is given in the legend. Grey indicates residues where no data are available.

### 3.3.3.2 Averaging temperature coefficients based on their secondary structural elements reveal apoSH monomers have a dynamic periphery with stable core strands

Additionally, averaging secondary structure features formed in the mature enzyme can further support which secondary structural components are more or less structured between apoSH SOD1 mutants. Similar to the rolling averages in pWT, the average for secondary structure features indicates ß2, ß3 and ß7 are considered less dynamic

relative to the loops and peripheral or edge strands, ß1, ß4-ß6, and ß8 (Figure 3.9). All mutants except for V148I show some degree of enhanced dynamics in the core strands which coincides with these mutants also having lower melting temperatures ($T_{0.5}$) than pWT (Table 3.1).[65] V148I is the only mutant with core strands that remain pWT-like which agrees with their comparable $T_{0.5}$ (Table 3.1).

From the CSP data in section 3.3.1, mutations cause a change in the local environment at or near the site of mutation, which could manifest as changes in dynamics at the sites of mutation relative to pWT. However, only secondary elements that house the point mutations for A4V reflect significant differences in averaged temperature coefficients (+/- 1 ppb/K) relative to pWT, while G85R, E100G, V148G and V148I show no significant differences (Figures 3.8 and 3.9). The variable impact of mutations on the averaged secondary structure temperature coefficient could be due to the location of the mutation with respect to their secondary structural component. A4V, G85R, V148G/I are residues that are located centrally in their respective strands, whereas E100G is located toward the ends of ß6. However, with the apoSH state being more immature than the mature holoSS state, the strands in apoSH may not be fully formed, and the critical hydrogen bonds involved may not be present.[66] Thus, if the ends of strands are less structured, introduction of mutations in a flexible region may not be as disruptive as introducing mutations into the centre of a structured strand that in all likelihood is forming stabilizing bonds.[66]

Moreover, there are also other secondary structural features in mutants other than the structural element housing the mutation that show changes relative to pWT, that could be suggesting additional differences in dynamics for mutants. In both G85R and E100G, ß7 and ß8 are more and less dynamic relative to pWT, respectively (Figure 3.9). Additionally, only G85R shows the ß5-6 cleft being more dynamic which aligns with the enhanced dynamics Sekhar *et al*., 2016 report for this cleft.[55] These perturbations suggest the impact of point mutations is not only localized to the site of mutation, rather propagates throughout the structure.

With these averaged temperature coefficients based on secondary structural elements, the need to address the influence of a large outlier is critical. Although from the rolling average data, most of the data show grouped behaviour meaning no large discrepancies between a residue and its neighbouring residues exist, residues in ß4 are a notable exception. In ß4, most of the residues are surprisingly very similar, however F45 is consistently a larger value than the rest of the residues. With F45 considered, G85R and V148I's ß4 averaged temperature coefficient is significantly different (+/- 1 ppb/K) than pWT, but without F45 only G85R remains significantly different (Figure 3.9, Figure 3.10). G85R is the only mutant with expected pronounced differences in ß4 since the G85R mutation is on the neighbouring ß5 strand; aside from propagating effects due to the mutation on the neighbouring strand, the mutation has an additional impact by likely disrupting the hydrogen bond between G85 and F45.[218] By averaging based on secondary structure, edge beta strands tend to have values lower than core beta strands

because only some of their residues are involved in hydrogen bonds. Therefore, the values of the averaged secondary structure temperature coefficients are useful for identifying edge strands that are only partially involved in hydrogen bonds and core strands that are sandwiched between beta-strands.

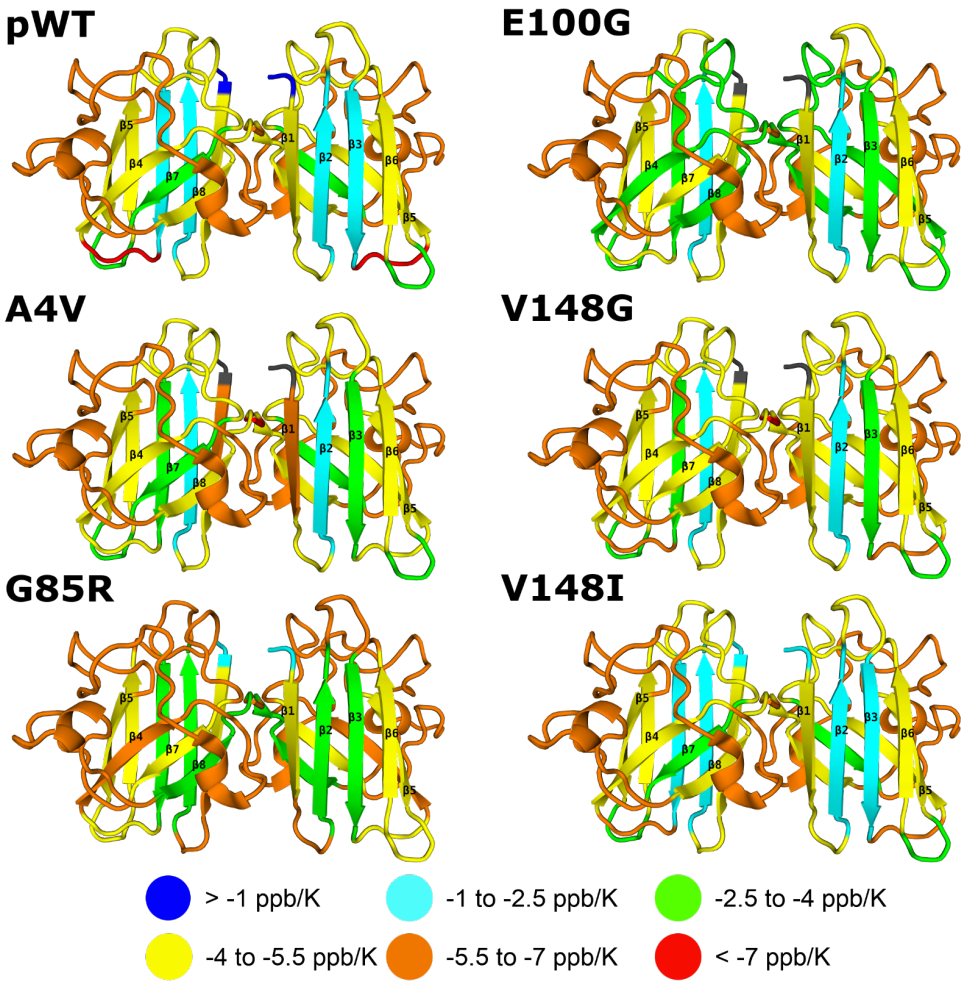



Figure 3.10: Secondary structural average amide proton temperature coefficients for apoSH pWT and mutants from 9–18 °C at pH 7.4, without residue 45 considered. Secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue for the more stable.

The colour scheme is given in the legend. Grey indicates residues where no data are available. In ß4, the outlier, residue 45, is not considered. Averaged values are also given in Table 3.3.

It is understandable that the edge beta strands have higher temperature coefficients compared to core strands; most of the residues in core strands are involved in hydrogen bonding to multiple neighbouring strands, whereas residues in edge strands will only have one neighbouring strand to hydrogen bond with. Therefore, while averaging the whole strand is useful for assessing how different their stabilities are between mutants, averaging based on residues that would be hydrogen bonded in the mature enzyme is required to understand changes in hydrogen bonding in the immature state.

#### 3.3.3.3 Hydrogen bonding in apoSH pWT is similar to holoSS SOD1

Previously, amide proton temperature coefficients were used to assess intramolecular hydrogen bonding[199,219], and although more than hydrogen bonds play a role in amide proton chemical shift temperature dependence[66,202], temperature coefficients previously correlated loss of structure to changes in the hydrogen bonding. ApoSH SOD1 lacks metal cofactors and the intrasubunit disulfide bond, and subsequently hydrogen bonding, that help stabilize structural components in the mature enzyme. The lack of these maturation steps may result in loss of hydrogen bonds, and with the addition of point mutations, apoSH SOD1 may show altered protein dynamics that influence its folding and/or misfolding mechanisms.

Hydrogen bonded residues are defined as residues involved in hydrogen bonding in the mature holoSS SOD1 (Figure 3.11A, Appendix A). By averaging only the expected

hydrogen bonded residues by secondary structural elements, differences in hydrogen bonding across beta-strands can be probed for local differences as well as between mutants. Temperature coefficient values more than -4.6 ppb/K are assumed to be hydrogen bonded.[199,219] Based on this value, most of the hydrogen bonds seem intact in apoSH pWT across the beta strands, which agrees with the apoSH beta-barrel being mostly holoSS-like (Section 3.3.1)[55], with the notable exception of ß4 (Figure 3.11). ß4 has two histidines (46 and 48) that are involved in coordinating the copper ion, and additionally this strand involves hydrogen bonding to both ß5 and ß7, both strands that also coordinate copper and zinc ions, respectively.[220] The lack of metal cofactors could enhance the local dynamics of each strand, which inadvertently impacts the hydrogen bonding.

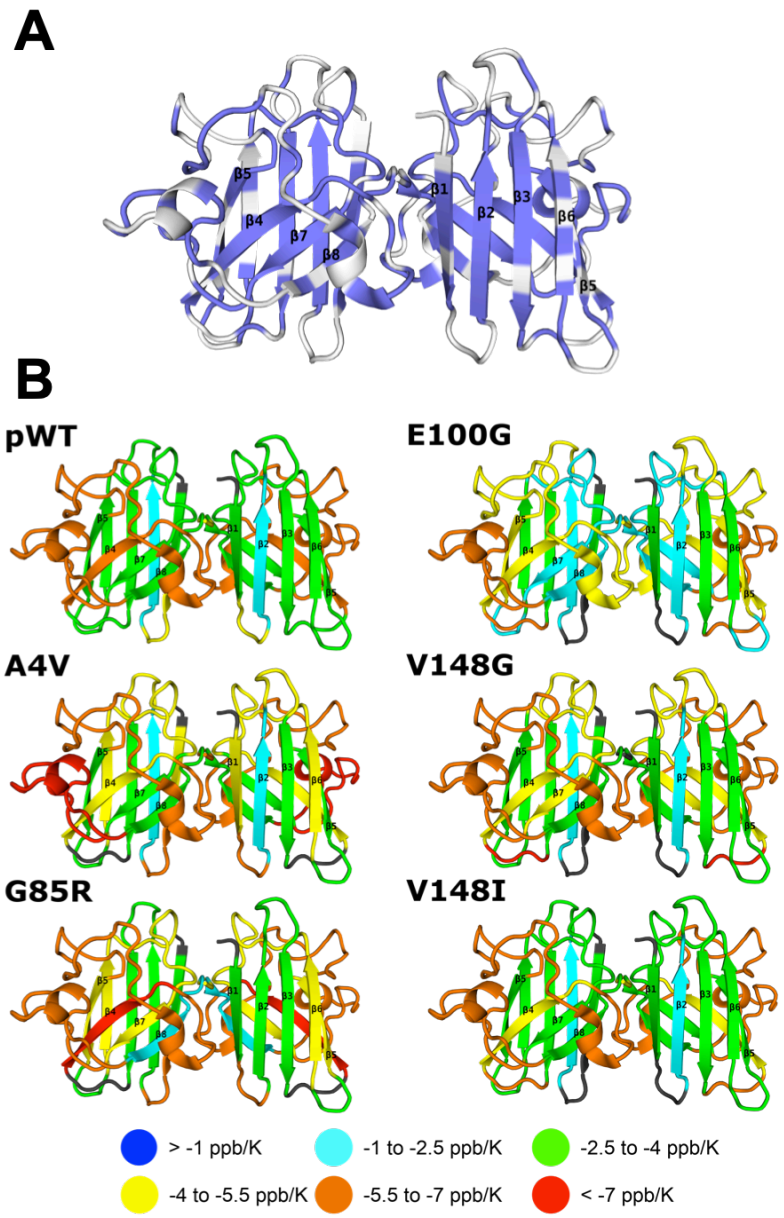



Figure 3.11: Secondary structural average amide proton temperature coefficients for hydrogen bonded residues in apoSH pWT and mutants from 9–18 °C at pH 7.4. (A) Residues that are hydrogen bonded in holoSS SOD1 are mapped onto a 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). (B) Averaging only hydrogen bonded residues, secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue for the more stable. The colour scheme is given in the legend. Grey indicates residues where no data are available.

Interestingly, H46 and H48 residues that coordinate the metal cofactors are not the ones with larger temperature coefficients (-3.6 and -4.2 ppb/K, respectively). Instead, G41

and F45 have the largest values (-6.9 and -11.7 ppb/K, respectively), and they are residues involved in reciprocal hydrogen bonding to G85 and A89 in ß5, respectively (Appendix A).[221] The G85 amide peak is not assigned, however A89 has a value of -5.0 ppb/K, which is lower than the -4.6 ppb/K value expected for a hydrogen bonded residue.[199,219] This could suggest that hydrogen bonding is heavily impaired between ß4 and ß5, likely due to the missing cofactors. Additionally, between ß4 and ß5, the hydrogen bonded residues in loop 4 also suggest perturbed hydrogen bonding, with only residues 68 and 69 (-3.1 and -3.7 ppb/K, respectively) having values below -4.6 ppb/K. The high dynamics in this area is likely due to the absence of metal cofactors and intra-subunit disulfide bond, which would otherwise anchor structural components together.

Additionally, residues in loop 7 that are hydrogen bonded in the mature enzyme also may not be intact in apoSH SOD1. Loop 7 residues not near the short helix have large temperature coefficients, whereas the values for residues responsible for forming native helix (131-136) suggest the hydrogen bonds are partially intact in apoSH SOD1.[222] This could suggest the helix is partially formed in apoSH SOD1, which aligns with previous studies suggesting apoSH pWT can transiently access the helix.[54,55] In summary, hydrogen bonding in apoSH pWT seems heavily impacted in ß4, ß5, loop 4 and loop7, all areas that are involved either directly or indirectly with maturation events and subsequent structuring in holoSS SOD1.

3.3.3.4 ApoSH mutants have perturbed hydrogen bonding that may explain their different dynamics

Additionally, apoSH mutants also show small differences in hydrogen bonded residues relative to pWT that are not exclusively located near the site of mutation (Figures 3.8 and 3.12). V148I is pWT-like but shows small differences, such as F45, which is discussed above. V148G interestingly shows significant perturbations around the site of mutation that cascade through ß8. Residues in loop 7 (128, 136, 138, 139) are more dynamic in V148G than in pWT, whereas near the site of mutation, G147 decreases from -4.1 to -2.1 ppb/K from pWT to V148G. G147 hydrogen bonds to G114 which changes from -2.6 to -6.5 ppb/K. These contrasting changes might suggest the native hydrogen bonds between 114 and 147 might be perturbed, and 147 is bonded differently. Additionally, G150 seems to no longer form a hydrogen bond with V5, but the V5 temperature coefficient suggests V5 is still hydrogen bonded. The hydrogen bonded temperature coefficient differences between V148I and V148G highlight the impact a point mutation can make on the hydrogen bonding pattern throughout the protein.

Similarly to V148G, A4V also shows changes throughout the structure relative to pWT (Figures 3.8 and 3.12). For A4V, temperature coefficient values for residues 4 and 5 in ß1 (-6.2 and -5.5 ppb/K, respectively) suggest these residues no longer form hydrogen bonds. However, both residues 20 and 150 reciprocally form hydrogen bonds with 4 and 5 in holoSS SOD1, respectively, and their temperature coefficients reflect values of intact hydrogen bonding (F20 is -1.5 ppb/K for F20 and -1.4 ppb/K for G150).

Similar to V148I and V148G, ß4 in A4V is also different from pWT (discussed above). Where only pWT shows one residue with a disrupted hydrogen bond in ß4 (F45), in A4V V47 also loses its hydrogen bond. In mature SOD1, V47 hydrogen bonds to G82 and F45 bonds to G85, both near or in ß5. In ß6, although both residues 95 and 99 in A4V increase in temperature coefficient values relative to pWT, only the value for I99 suggests a hydrogen bond is no longer present in A4V, compared to A95 which suggests it is not present in both pWT and A4V. Overall, there appears to be a significant disruption in the hydrogen bonding pattern for A4V (expanded on below).

Furthermore, E100G and G85R both have different thermal stabilities, but where G85R is the more stable mutant of the two, it shows significantly more perturbations in hydrogen bonded residues than E100G relative to pWT (Figures 3.8 and 3.12). G85R shows significant perturbations in ß4, ß5, loop 6, ß7 and ß8. It seems introducing G85R in a metal-binding region causes propagating effects to strands that otherwise do not seem related to metal binding. Specifically, residues 144 and 147 in ß8 both show decreases in temperature coefficient values, which could be suggesting hydrogen bonds are forming due to breaking of critical bonds in ß4 (residues 44, and 45), ß5 (residue 87), and near loop 6 (residues 101, 105, and 106). Although ß7 does not show increased hydrogen bonding relative to pWT like ß8 does, the missing assignments in ß7, specifically 119 and 120, would be vital to understanding if ß7 is impacted by residues in ß8 with enhanced hydrogen bonding. Moreover, E100G does not show significant differences relative to pWT, aside from ß4 (residues 45 and 46) and loop 4 (residues 52

and 60). However, the change in values in either region is not large enough to suggest changes in hydrogen bonding relative to pWT. Overall, the subtle changes in temperature coefficient values in hydrogen bonded residues point to potential changes in dynamics due to changes in bonding.

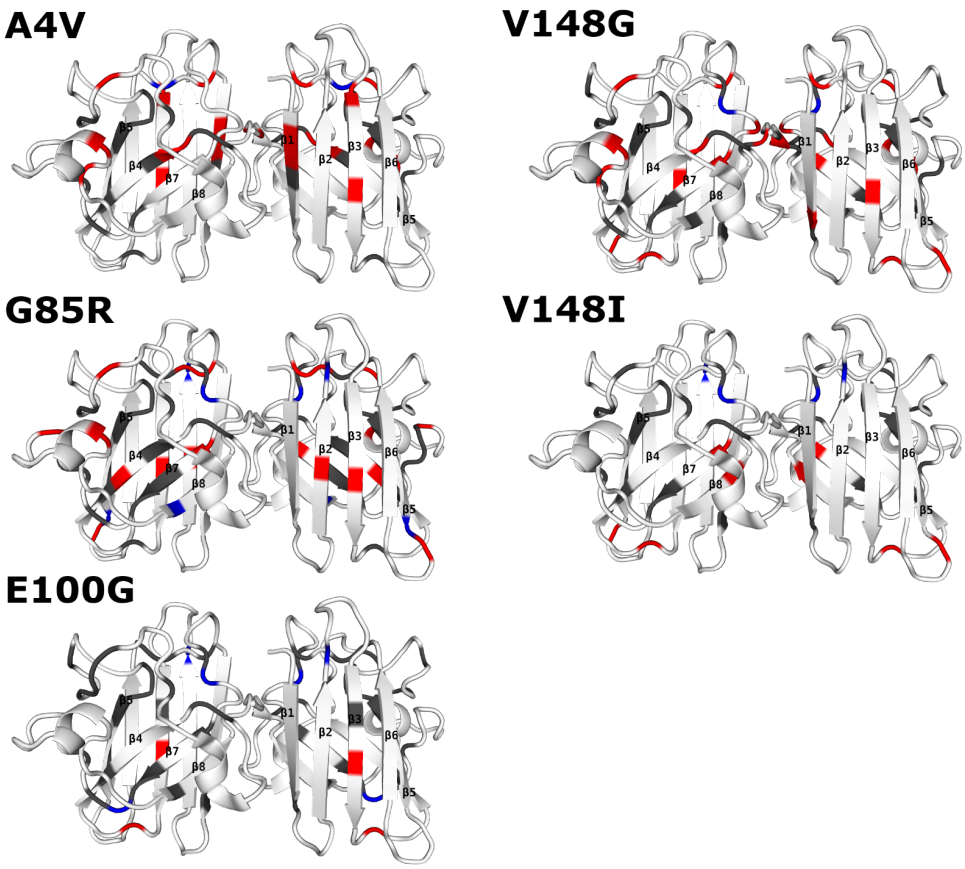



Figure 3.12: Hydrogen bonding differences based on temperature coefficient values between apoSH mutants and pWT. Residues are assumed to be hydrogen bonded if their temperature coefficient value is larger than -4.6 ppb/K. Blue indicates the residue is considered hydrogen bonded in apoSH mutant and not in pWT, and red indicates a loss of hydrogen bonding in apoSH mutants that is present in pWT. White represents no change in the residue's hydrogen bonding status, and grey indicates residues with no data available.

Earlier the rolling averages for temperature coefficients highlighted the enhanced dynamics in loop 7 for A4V, G85R, V148G and to a lesser extent E100G (Figure 3.7). From looking at only values for the residues expected to form hydrogen bonds, it is clear

the increased dynamics relative to pWT occur in mostly hydrogen bonded residues (Figure 3.8, 3.11A, and 3.12). For three of these mutants, previous studies indicates transient formation of the native helix is diminished.[55] Conformational disorder in either loop 4 or 7 is known to result in exposure of the edge strands in monomer SOD1, which leads to non-native protein association.[223,224] Although some of these residues are not considered hydrogen bonded in apoSH pWT (based on their value being less than -4.6 ppb/K), the overall enhanced dynamics in these regions for apoSH mutants could have consequences in later structural steps upon maturation and/or play a role in protein aggregation.

The hydrogen bonded residues are expected to hydrogen bond in the mature holoSS state based on structural data, not necessarily in the apoSH state.[218] Again, temperature coefficient values smaller than -4.6 ppb/K are not considered as hydrogen bonded. Upon comparison of the hydrogen bonded residues between apoSH pWT and apoSH mutants, A4V, G85R and V148G show loss of hydrogen bonds throughout the structure relative to pWT (Figure 3.12). In contrast, E100G and V148I show both residues that gain and lose hydrogen bonds, and both mutants are considered pWT-like in terms of structure.[55] However, A4V, G85R and V148G show extensive perturbations which may align with these mutants showing impaired maturation. A4V and V148G are both dimer interface mutants that weaken the dimer interface[87] (see Chapter 7) whereas G85R perturbs metal binding[225] (see Chapter 7). These decreased hydrogen bonds in the apoSH mutants may impact the folding pathways which inevitably may alter maturation.

Table 3.2: Different temperature coefficient averaging methods and their uses.

| Different Averaging Methods | Uses |
|---|---|
| Rolling average | Small, local structural differences can be discerned without any assumption of secondary structure<br>Ideal for proteins with no reliable structure |
| Averaging based on secondary structure | Highlights perturbations in secondary structure<br>Assumes residues are involved in secondary structure |
| Averaging based on hydrogen bonded residues present in secondary structure | Highlights perturbations in hydrogen bonding in secondary structure elements<br>Assumes secondary structure |

3.3.3.5 Temperature coefficient data correlate with previously reported changes in dynamics for apoSH mutants

With the abundant amide proton temperature dependence data considered above, we now consider the trends among mutations and their consequences on protein folding. Due to the location of the mutation, A4V has a more dynamic dimer interface, specifically residues in ß1 (-6.1 ppb/K compared to -4.3 ppb/K for pWT). These residues are also involved in hydrogen bonding, and their change in values relative to pWT suggest these hydrogen bonds are no longer formed. This loss of structure near the dimer interface may explain previous reports that highlight perturbed dynamics in the dimer interface as well as weakened dimerization for A4V in multiple states (see Chapter 7).[55,87]

Table 3.3: Amide proton temperature coefficients for apoSH variants.

| | pWT | A4V | G85R | E100G | V148G | V148I |
|---|---|---|---|---|---|---|
| **average** | -4.78 (125) | -5.16 (126) | -5.20 (113) | -4.56 (107) | -5.10 (116) | -4.82 (121) |
| **standard dev** | 2.73 | 2.53 | 2.43 | 2.71 | 2.93 | 2.95 |
| **average ddTC** | - | -0.30 (120) | -0.58 (110) | 0.11 (106) | -0.45 (111) | -0.08 (119) |
| | | | | | | |
| **N-terminal B-sheet** | -3.43 (31) | -4.16 (30) | -3.94 (31) | -3.32 (29) | -3.79 (31) | -3.49 (31) |
| **C-terminal B sheet** | -4.49 (24) | -4.71 (27) | -5.22 (19) | -4.57 (20) | -4.84 (23) | -5.09 (23) |
| **N** | -0.93 (1) | - | -1.50 (1) | - | - | -1.48 (1) |
| **B1** | -4.26 (6) | -6.06 (5) | -4.58 (6) | -4.24 (6) | -5.18 (6) | -4.41 (6) |
| **loop I** | -5.27 (4) | -4.21 (4) | -5.55 (4) | -4.94 (3) | -5.17 (3) | -4.90 (3) |
| **B2** | -2.24 (9) | -2.35 (9) | -2.82 (9) | -2.05 (9) | -2.35 (9) | -2.28 (9) |
| **loop II** | -5.30 (3) | -5.00 (4) | -6.32 (3) | -5.42 (3) | -4.70 (2) | -5.28 (3) |
| **B3** | -2.32 (8) | -2.93 (8) | -2.99 (8) | -2.55 (7) | -2.62 (8) | -2.26 (8) |
| **loop III** | -7.71 (3) | -4.22 (2) | -5.03 (2) | -4.04 (2) | -6.91 (3) | -6.24 (3) |
| **B4** | -5.52 (6) | -5.45 (7) | -7.20 (5) | -5.40 (6) | -5.68 (8) | -7.01 (6) |
| **loop IV** | -5.76 (19) | -6.07 (20) | -6.02 (20) | -5.61 (17) | -6.10 (18) | -5.64 (19) |
| **B5** | -4.58 (4) | -5.03 (5) | -4.79 (3) | -4.84 (4) | -4.87 (5) | -4.55 (4) |
| **loop V** | -3.80 (4) | -3.62 (4) | -4.01 (4) | -3.01 (4) | -3.46 (4) | -3.59 (4) |
| **B6** | -4.89 (8) | -5.30 (8) | -5.37 (8) | -4.45 (7) | -4.99 (8) | -5.02 (8) |
| **loop VI** | -4.80 (13) | -5.30 (13) | -5.53 (10) | -3.94 (8) | -5.13 (12) | -4.67 (12) |
| **B7** | -3.28 (5) | -3.44 (6) | -4.94 (3) | -4.05 (3) | -4.48 (5) | -3.93 (4) |
| **loop VII** | -5.89 (21) | -6.93 (20) | -6.64 (17) | -5.90 (19) | -6.54 (19) | -5.84 (20) |
| **B8** | -4.57 (9) | -4.93 (9) | -3.96 (8) | -3.98 (7) | -4.33 (5) | -4.86 (9) |
| **C** | -6.80 (2) | -7.74 (2) | -6.87 (2) | -6.45 (2) | -9.39 (1) | -6.54 (2) |

Temperature coefficient values are averaged across all assignments, the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets. Average ddTC is the difference in average temperature coefficients for the mutant relative to pWT and is calculated by subtracting the pWT average temperature coefficient value from the mutants' average temperature coefficient value.

Moreover, the stark differences between V148I and V148G temperature coefficients highlight the impact a nonconservative mutation like V148G has. Although the CSP data in section 3.3.1 do not show considerable differences (Figure 3.3), the temperature coefficients suggest, on average, V148G (-5.1 ppb/K) is less structured/more dynamic than V148I (-4.8 ppb/K) (Table 3.3). Additionally, the V148G temperature coefficient data also suggest there are less hydrogen bonded residues in V148G relative to pWT, which could explain its lower thermal stability (34 °C), while V148I remains pWT-like, both in terms of its stability (51°C) and its temperature coefficient data. Throughout

previous findings, V148I has consistently shown pWT-like behaviour in terms of structure[55,66], monomer stability[65,87], metal-binding affinity (Chapter 7), dimerization (Chapter 7)[87], and aggregation propensity.[57] In contrast, V148G has always shown destabilized behaviour in all those respects. Aside from V148 being highly conserved[25], understandably, a conservative mutation such as valine to isoleucine on the protein surface could be considered less impactful, especially in a dynamic protein like apoSH SOD1, compared to a nonconservative mutation like valine to glycine.

On the other hand, where V148I and V148G temperature coefficient data agree with their global stability, E100G and G85R temperature coefficient data do not. The more thermally stable G85R (40.7 °C) shows an average temperature coefficient of -5.2 ppb/K and the least thermally stable apoSH mutant studied here, E100G (33.2 °C) has an average value lower than pWT (-4.6 ppb/K, discussed more below) (Table 3.3). Aside from A4V, G85R is the only other mutant that has a N-terminus face (consists of ß1-ß3, ß6) significantly more dynamic than pWT (see Table 3.3). Also, both A4V and G85R showed significant CSP values on the N-terminus face (see section 3.3.1), suggesting these mutants may be forming less native-like interactions in the monomer that manifest in different local environments relative to pWT.

Similar to A4V, G85R has previously shown to have issues with maturation. However, where A4V fails to natively dimerize (see Chapter 7)[53], G85R fails to metalate properly, form the conserved disulfide bond, and perturb enzymatic activity (see Chapter

7).[221,226,227] Notably, this propagating impact reported for G85R aligns with the temperature coefficient data suggesting the mutation has more of a global impact on protein structure and maturation. Overall, the temperature coefficient data highlight critical areas of changed dynamics relative to pWT, that align well with previous findings.

### 3.3.4 Comparing ApoSH SOD1 to HoloSS SOD1

With the addition of metal cofactors, disulfide bond, and dimerization, it is understandable the holoSS SOD1 (~91 °C) is more thermally stable than the apoSH state (48 °C) (Table 3.1).[65,66] This increased thermal stability is reflected in the average temperature coefficient for holoSS pWT (-3.2 ppb/K) being smaller than apoSH pWT (-4.8 ppb/K), which aligns with holoSS pWT being more structured than apoSH pWT (Tables 3.2 and 3.3).[66] Interestingly, the holoSS and apoSH pWT N-terminus face (ß1-ß3, ß6) does not show large differences between the states, -3.6 and -3.4 ppb/K, respectively, but the C-terminus face (ß4, ß5, ß7, ß8) shows larger differences (-3.0 and -4.5 ppb/K, respectively) (Tables 3.2 and 3.3). Since the C-terminus face has residues involved in metal binding and disulfide bonds[215], it is not surprisingly the C-terminus is more destabilized not only in the apoSH state compared to the holoSS state, but also relative to the N-terminus face.

Table 3.4: Comparing amide proton temperature coefficients for apoSH and holoSS SOD1.

| | ApoSH | | | HoloSS | | |
|---|---|---|---|---|---|---|
| | rapWT | raE100G | raV148I | hpWT | hE100G | hV148I |
| average | -4.78 (125) | -4.56 (107) | -4.82 (121) | -3.24 (97) | -3.37 (93) | -3.02 (92) |
| standard dev | 2.73 | 2.71 | 2.95 | 2.16 | 2.31 | 2.03 |
| average ddTC | - | 0.11 (106) | -0.08 (119) | - | -0.16 (83) | 0.14 (81) |
| | | | | | | |
| N-terminal B-sheet | -3.43 (31) | -3.32 (29) | -3.49 (31) | -3.57 (27) | -3.44 (21) | -3.48 (23) |
| C-terminal B sheet | -4.49 (24) | -4.57 (20) | -5.09 (23) | -2.83 (26) | -2.96 (27) | -2.66 (22) |
| N | -0.93 (1) | - | -1.48 (1) | - | -8.97 (1) | - |
| B1 | -4.26 (6) | -4.24 (6) | -4.41 (6) | -3.66 (6) | -3.92 (4) | -4.48 (3) |
| loop I | -5.27 (4) | -4.94 (3) | -4.90 (3) | -4.73 (1) | - | - |
| B2 | -2.24 (9) | -2.05 (9) | -2.28 (9) | -3.15 (7) | -3.05 (7) | -2.52 (7) |
| loop II | -5.30 (3) | -5.42 (3) | -5.28 (3) | -5.12 (1) | -5.32 (1) | -4.93 (1) |
| B3 | -2.32 (8) | -2.55 (7) | -2.26 (8) | -2.76 (7) | -2.81 (6) | -2.46 (6) |
| loop III | -7.71 (3) | -4.04 (2) | -6.24 (3) | -4.02 (2) | -4.22 (2) | -3.98 (2) |
| B4 | -5.52 (6) | -5.40 (6) | -7.01 (6) | -3.08 (7) | -2.95 (8) | -2.57 (6) |
| loop IV | -5.76 (19) | -5.61 (17) | -5.64 (19) | -3.50 (18) | -4.08 (18) | -3.25 (21) |
| B5 | -4.58 (4) | -4.84 (4) | -4.55 (4) | -3.05 (7) | -3.54 (6) | -3.29 (6) |
| loop V | -3.80 (4) | -3.01 (4) | -3.59 (4) | -2.76 (3) | -2.77 (3) | -3.15 (4) |
| B6 | -4.89 (8) | -4.45 (7) | -5.02 (8) | -4.69 (7) | -3.96 (4) | -4.47 (7) |
| loop VI | -4.80 (13) | -3.94 (8) | -4.67 (12) | -2.61 (7) | -2.86 (7) | -2.70 (8) |
| B7 | -3.28 (5) | -4.05 (3) | -3.93 (4) | -1.78 (5) | -1.90 (6) | -1.71 (5) |
| loop VII | -5.89 (21) | -5.90 (19) | -5.84 (20) | -3.09 (11) | -2.70 (11) | -2.14 (9) |
| B8 | -4.57 (9) | -3.98 (7) | -4.86 (9) | -3.39 (7) | -3.46 (7) | -3.08 (5) |
| C | -6.80 (2) | -6.45 (2) | -6.54 (2) | -0.70 (1) | -4.31 (2) | -3.72 (2) |

Temperature coefficient values are averaged across all assignments, the beta-strands on N-terminal (ß1-3,6) and C-terminal face (ß4,5,7,8) and secondary structural elements, and each average also has the number of residues averaged in brackets. Average ddTC were calculated by subtracting the pWT value from the mutants for the same SOD1 state. HoloSS SOD1 values were obtained from Doyle *et al*., 2016.[66]

Previously, average temperature coefficient values for holoSS mutants showed some correlation to their global thermal stabilities.[66] The averaged values are further expanded by averaging values that correspond to all the same residues in either apoSH or holoSS (see Appendix K), as well as residues that are measured in both apoSH and holoSS (51 residues) (Table 3.5, Appendix L). These averages are further filtered by only using residues that are hydrogen bonded in the holoSS state as considered above. All the above averages are then ranked across their respective states, and correlations are not used since only 3 protein variants are considered.

Table 3.5: Averaging temperature coefficients for apoSH and holoSS SOD1 based on different conditions.

| | ApoSH | | | HoloSS | | |
|---|---|---|---|---|---|---|
| | pWT | E100G | V148I | pWT | E100G | V148I |
| $T_{(0.5)}$ | 47.60 | 33.20 | 51.00 | 96.30 | 91.10 | 97.00 |
| Avg (all res) | -4.78 | -4.56 | -4.82 | -3.24 | -3.37 | -3.02 |
| Avg (state) | -4.64 | -4.52 | -4.48 | -2.93 | -3.09 | -2.81 |
| Avg (same res) | -4.52 | -4.46 | -4.29 | -3.11 | -3.24 | -2.97 |
| Avg (H-bond res) | -4.09 | -3.79 | -3.95 | -2.89 | -2.89 | -2.75 |
| Avg (H-bond state) | -3.90 | -3.79 | -3.75 | -2.83 | -2.99 | -2.69 |
| Avg (H-bond same res) | -4.43 | -4.40 | -4.14 | -2.90 | -3.04 | -2.75 |

Avg (all res) is averaging all residues assigned for each mutant and SOD1 state. Avg (state) is averaging all the same assigned residues across pWT, E100G and V148I for the same state (apoSH or holoSS). Avg (same res) is averaging the residues that are assigned across all 3 variants over both states, which is 54 residues. Avg (H-bond res) is averaging all hydrogen bonded residues assigned for each respective mutant. The hydrogen bonded residues are residues assumed to be hydrogen bonded in the holoSS state based on Jonnsson *et al*.[221] Avg (H-bond state) is averaging all hydrogen bonded residues that are assigned in all mutants across each state. Avg (H-bond same res) is averaging all hydrogen bonded residues that are assigned in all mutants across both states. HoloSS SOD1 values were obtained from Doyle *et al*., 2016.[66]

For all holoSS averages, the order remains unchanged, with V148I being most structured, then pWT and lastly E100G, which agrees with their thermal stabilities (Table 3.5). However, apoSH mutant thermal stabilities do not correlate with their averaged temperature coefficients - while the relative apoSH mutant thermal stabilities are the same as in holoSS (V148I, pWT, E100G), the average temperature coefficients take on the opposite trend, with E100G appearing the most stable and V148I appearing the most dynamic. Although the least thermally stable apoSH E100G having the highest local structural stability convolutes the trend, it is important to consider global stability is not always a dependable predictor of protein dynamics.[204,206,207,228] Additionally, apoSH E100G specifically has the fewest assignments (107) relative to pWT (125) and V148I (121), which suggests there could be residues that have shifted in E100G relative to pWT which could also be less structured in E100G. With no assignments for these potentially

shifted residues, the averaged values could be inappropriately weighted toward residues that are more structured in E100G.

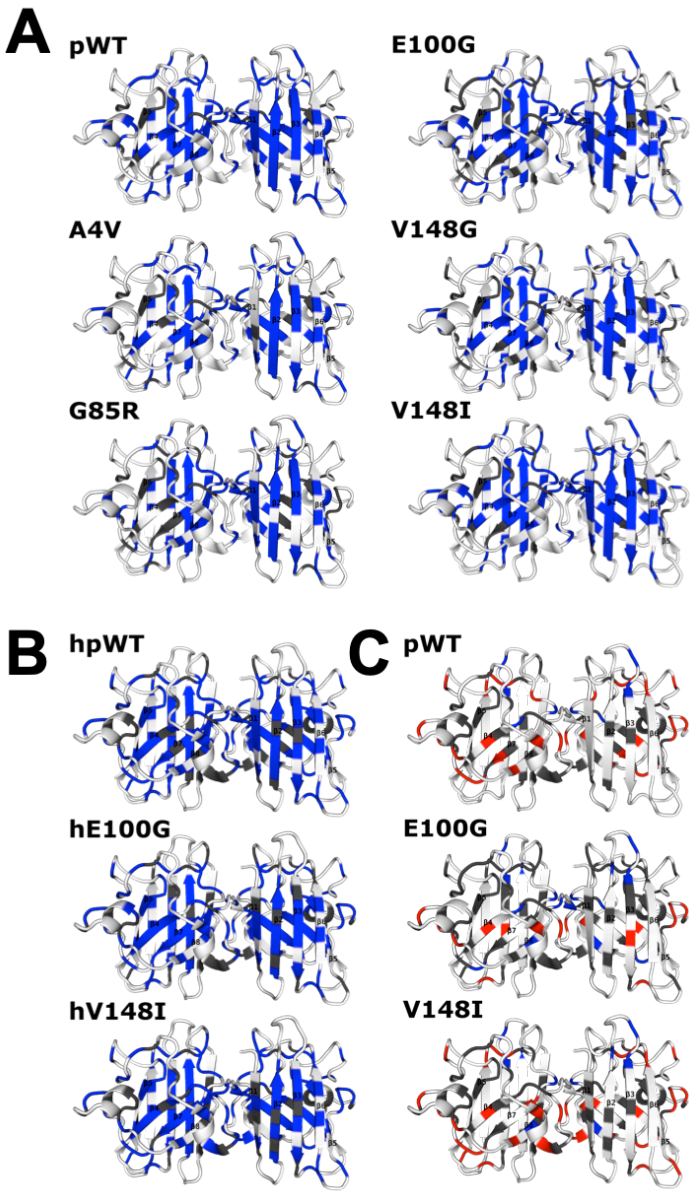



Figure 3.13: Comparison of hydrogen bonded residues in apoSH and holoSS SOD1 based on amide temperature coefficients. (A) and (B) residues are considered hydrogen bonded (blue) if their temperature coefficient value is larger than -4.6 ppb/K for apoSH and holoSS SOD1, respectively. (C) Differences in hydrogen bonds between apoSH and holoSS SOD1 are considered. Blue and red residues indicate the residue gained a hydrogen bond in the apoSH and holoSS state, respectively. White represents no change in the residue's hydrogen bonding status, and grey indicates residues with no data available.

Upon inspection of differences in hydrogen bonding between apoSH and holoSS pWT, apoSH SOD1 shows fewer hydrogen bonds than holoSS SOD1 (Figure 3.13). More specifically, residues near metal-binding sites show loss of hydrogen bonds (temperature coefficient value becomes smaller than -4.6 ppb/K) in strands and loops. Without the metals to serve as structural anchors, the enhanced dynamics suggest critical bonds are no longer formed. This disruption in hydrogen bonds due to lack of maturation events is also occurring in E100G and V148I, and interestingly in similar regions seen in the pWT comparison (Figure 3.13). Overall, introducing metal cofactors, and maturation events in general, leads to increased structuring in SOD1 due to hydrogen bond formation.

Additionally, the impact mutations have on different SOD1 states is worth considering when discussing aggregation prone states. In section 3.3.3., hydrogen bonding pattern differences between mutants and pWT are compared. Where the apoSH E100G and V148I show some differences across hydrogen bonding, the holoSS E100G and V148I show either minimal or no changes relative to pWT. Although both E100G and V148I are considered pWT-like[55,66], their differences in behaviour may highlight subtle changes in dynamics across maturation states.

It is clear mutations impact the apoSH state more than the holoSS state, which can be explained by maturation events serving as points of stability such that impact of mutations is minimized. This was also previously observed when assessing thermal stability using differential scanning calorimetry (DSC) for both states.[52,65] The range of

melting points, and thus the extent of destabilization relative to pWT, was larger for apoSH mutants than for holoSS mutants. The differences noted here suggest changes in local stability and/or altered dynamics in apoSH SOD1, which in turn impact the protein folding landscape and further support more immature states being important in protein aggregation.[55]

### 3.3.5 Amide proton temperature dependence changes upon protein unfolding

Since certain apoSH SOD1 mutants are considered thermally destabilized both relative to pWT and at physiological temperature, the extent of unfolded structure and subsequently the mutant dynamics at physiological temperatures is vital to understand. However, due to limited data exploring protein unfolding through proton amide temperature dependence, it is unclear how protein unfolding impacts proton amide temperature dependence. Here, we chose four thermally destabilized SOD1 mutants (A4V, G85R, E100G, and V148G), to assess temperature driven protein unfolding via HSQC spectra. For each mutant, spectra were collected at a range of temperatures; spectra were collected at temperatures prior to any protein unfolding, (henceforth referred to as the folded range) as well as temperatures where the protein has begun to unfold based on their respective DSC thermogram (henceforth referred to as the unfolding range) and intensity profiles of Trp-32 sidechains (Table 3.1, see Chapter 4).

Unfolding temperature ranges varied in an attempt to acquire different extents of protein unfolding in addition to the variable melting temperatures for each mutant (Tables

3.1 and 3.5). G85R had the least spectra collected (2) in the unfolding range that spanned from 29.3 °C - 31.5 °C, and with a melting temperature of 40.7 °C, 31.5 °C is at the start of protein unfolding. Clearly with only 2 spectra, slopes are considered more error prone. For A4V, the unfolding range is from 19.8 °C - 34.5 °C (9 spectra), which is a few degrees before its melting temperature of 36.3 °C. The 6 unfolding spectra for E100G were collected from 22.3 °C until the melting temperature (33.2 °C), and 9 spectra for V148G from 20.1 - 36.7 °C, well past V148G's melting temperature (34 °C).

Table 3.6: Folded and unfolded temperature ranges for the apoSH mutants and their respective number of spectra acquired in brackets.

| Mutants | Folded Range | Unfolded Range |
|---|---|---|
| pWT | - | - |
| A4V | 5.65 - 17.94 (9) | 19.8 - 34.5 (9) |
| G85R | 6.20 - 27.04 (10) | 29.29 - 31.50 (2) |
| E100G | 5.17 - 24.50 (9) | 22.23 - 33.65 (6) |
| V148G | 1.04 - 19.88 (12) | 20.11 - 36.70 (9) |
| V148I | - | - |

- mutants not considered.

3.3.5.1 Variant CSP values in the unfolding range identify key initial unfolding regions

Previously in section 3.3.1, CSP data suggest destabilized mutants that start unfolding at 25 °C show more local areas of perturbations, which is thought to suggest protein unfolding. Extrapolating those data, CSPs are calculated at the highest temperature of measurement for each mutant, relative to the pWT spectra at the

corresponding temperatures. Where 19 °C CSP data in section 3.3.1 represent the folded structures of each protein and included local changes relative to pWT due to point mutations, the highest temperature CSP data (from now on referred to as high temperature CSP) represent areas of local changes due to protein unfolding. Since pWT remains folded throughout the unfolding ranges for the destabilized mutants, additional shifts at the high temperature, relative to those at 19 °C, are assumed to reflect protein unfolding.

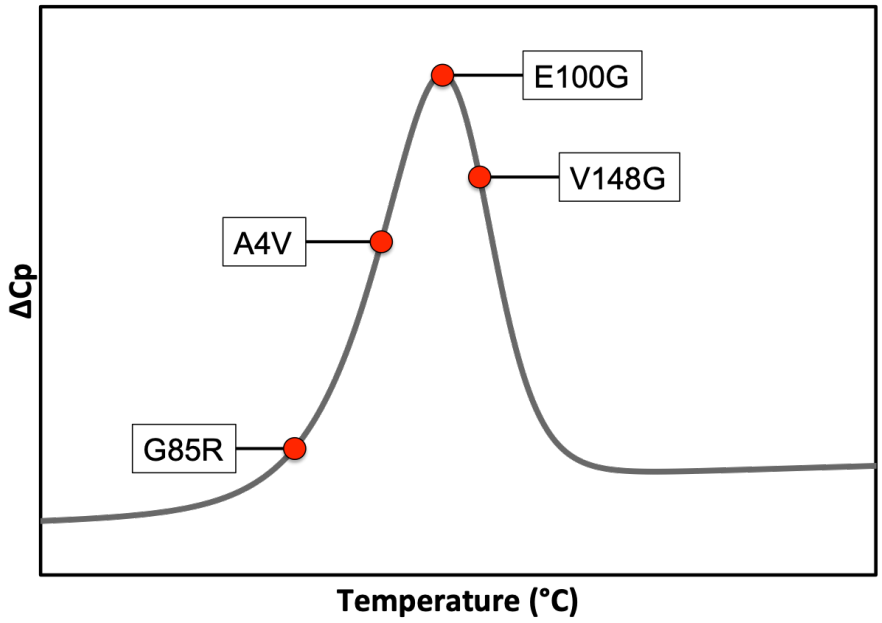



Figure 3.14: Representative DSC endotherm depicting the extent of unfolding for the spectrum collected at the highest temperature for each apoSH mutant. This is a representative DSC endotherm with apoSH mutants labeled along the unfolding to indicate the highest temperature a spectrum was collected relative to the mutant's endotherm.

Similar to the significant CSP values established by local standard deviations for each mutant discussed in section 3.3.1, sites of mutation and neighbouring residues are more locally perturbed than the rest of the structure. However, with the same global

cutoffs as section 3.3.1, apoSH mutants show more areas of perturbations at the high temperature than at 19 °C. Since G85R is the least into its unfolding transition relative to its DSC endotherm (Figure 3.14), the minimal CSP changes observed at its highest temperature aligns with the minimal extent of unfolding expected (Figure 3.15A).[65] It is anticipated that as proteins become more unfolded, the CSP values will vary more relative to the folded pWT.

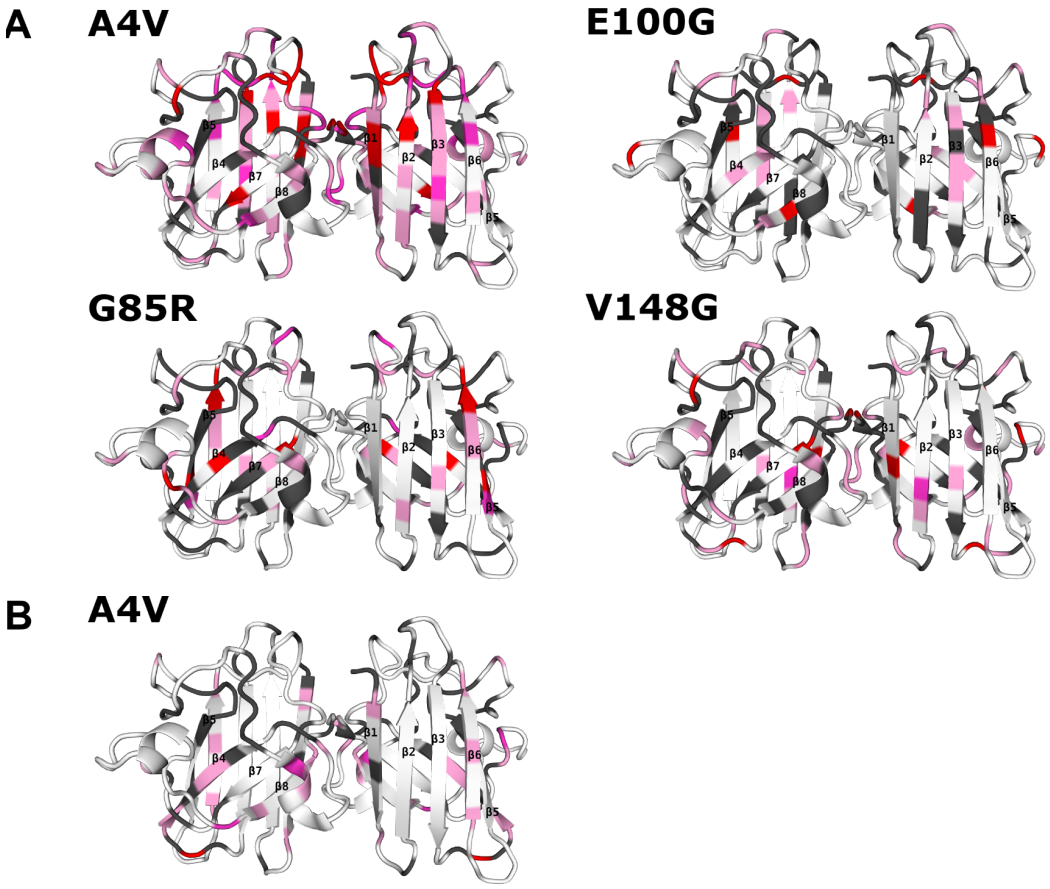



Figure 3.15: CSP for apoSH mutants relative to apoSH pWT at the highest temperature (global cutoffs) for the respective mutants. CSP values were coloured based on the same cutoff for all mutants. Residues are coloured to aid in visualization according to the relative magnitude of CSP values. The pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1 ppm, respectively. White denotes CSP values smaller than 0.05, and grey indicates residues where no data are available. (A) CSP values are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). Refer to Tables 3.1 and 3.5 for melting temperatures and temperature ranges spectra are acquired. (B) CSP values for apoSH A4V between 19 °C and high temperature (34.5 °C).

Surprisingly, E100G showed minimal increases in high temperature CSP data (Figure 3.15A). This is particularly unexpected since E100G spectra were collected until the melting temperature (Figure 3.14). Most of the changes are dispersed throughout the loops which may hint at reasons why minimal changes were observed. As previously discussed, E100G has the least number of assignments, and additionally, for all mutants the large loops (4 and 7) are largely missing assignments. Thus, in addition to missing residues, if most of the changes are expected to be in the loops, those missing assignments may be the explanation for the unexpected results.

More in line with the expected results, V148G shows enhanced CSP values at the high temperature. As stated earlier (section 3.3.1), the CSP data at 19 °C suggest differences at the site of mutation and near the metal-binding region, specifically loop 7. Since the highest temperature for V148G (36.7 °C) is well past its melting temperature (34 °C) (Figure 3.14), these effects further propagate throughout loop 7, and now include loop 4. From the 25 °C and high temperature CSP data, loop 7 seems to be the structural feature perturbed earliest by thermal unfolding. The gradual impacts, initially to loop 7 and then loop 4, are logical considering critical hydrogen bonds are formed between loop 7 and loop 4 that may no longer be present in apoSH SOD1.[229] Additionally, the perturbed metal-binding issues in the unfolding V148G discussed earlier (section 3.3.1) now seem valid. V148G initially unfolding around the metal-binding sites also explains the inability to see aggregation rescue at 37 °C[25] as well as the lack of zinc shuffling, described in Chapter 7.

Unlike V148G, or the other apoSH mutants, the high temperature CSP data for A4V behave uniquely. With its highest temperature being well into the unfolding transition (Figure 3.14), it is expected CSP values would only increase, implying the structure is moving away from the natively folded state. However, A4V CSP data at 19 °C show significant perturbations which lead to the hypothesis that A4V is likely forming a less native protein fold than other apoSH mutants. With a different fold, the new structure will unfold differently as well, meaning CSP values are not always expected to increase as would occur for a natively folded structure. A4V high temperature CSP data also suggest the non-native structure is still prevalent (Figure 3.15A), and instead of exclusively seeing certain areas become more perturbed, there are also areas that become less perturbed with temperature relative to pWT. With the structure fold likely being less native in comparison to pWT, conclusions on what the perturbation values mean are difficult.

3.3.5.2 CSP values between folded and unfolding A4V suggest the periphery of the protein unfolds initially

With A4V showing widespread differences relative to pWT in the folded temperature range, deciphering differences upon unfolding is difficult. However, determining the A4V CSP data between 19 °C and its highest temperature (34.5 °C) identifies areas that vary upon thermally unfolding A4V relative to its own folded structure. Unlike most of the CSP values for the mutants that show differences only at the site of mutation, the relative comparison highlights the edge strands in A4V (ß4 and ß6) as well as near the dimer interface are perturbed with temperature (Figure 3.15B). The negligible

CSP values at the site of mutation are expected since CSP data are not compared to pWT, as it was earlier, rather A4V at two different temperatures. Overall, the relative comparison of A4V at the two temperatures suggest the periphery of the protein unfolds prior to ß2, ß3 and ß7 (previously identified as a stable core in section 3.3.3.4). While A4V highlights the importance of a native fold when comparing mutants to pWT to assess protein unfolding, comparing mutants at each of the temperatures may highlight local areas of perturbations that could suggest unfolding.

3.3.5.3 Folded and unfolding temperature coefficients show differences that correlate with the extent of protein unfolding

With these structural differences upon protein unfolding, comparing the temperature coefficient data in the folded and unfolding temperature ranges (which will be referred to as the folded temperature coefficients and unfolding temperature coefficients, respectively), might further identify initial structural regions thermally unfolding. As noted above, the ranges of temperatures and the number of spectra collected are summarized in Table 3.6. The folded and unfolding temperature coefficients can also be assessed by the different averaging methods discussed earlier; however, the data in the unfolding range has less temperature coefficients because the peaks become difficult to track in the unfolding range, thus averaging of available residues may be misleading. Instead, the folded and unfolding temperature coefficients are compared by using both differences (difference = TC unfolding – TC folded) and ratios (ratio = TC unfolding/TC folded, where a difference of +/- 1 ppb/K or a ratio change of 50 % is

considered substantial. The ratios are plotted by both their value as well as binning the data by the 50 % cutoff. By comparing the folded and unfolding temperature coefficients, the data may highlight local structural changes due to protein unfolding.

The two points of comparison, differences and ratios, show similar results, but where the differences and ratios are binned and thus show only the points of significant differences, the ratio range shows gradual changes throughout structural regions (Figure 3.16). Due to only having two spectra in the unfolding region, G85R results are more error prone. Nevertheless, the folded and unfolding temperature coefficient for G85R show, although minimally, differences in the core strands (ß2, ß3) and edge strands (ß4, ß6, ß8). These changes, aside from ß2, are due to the temperature coefficient value becoming smaller (blue, less negative), suggesting more local structuring upon unfolding. The areas of changes also correlate with the residues that show larger CSP values. With two spectra for the unfolding temperature coefficients, making conclusions is difficult, but the stabilization in temperature coefficient values for G85R could be forming local areas of structuring due to restructuring and/or aggregation.

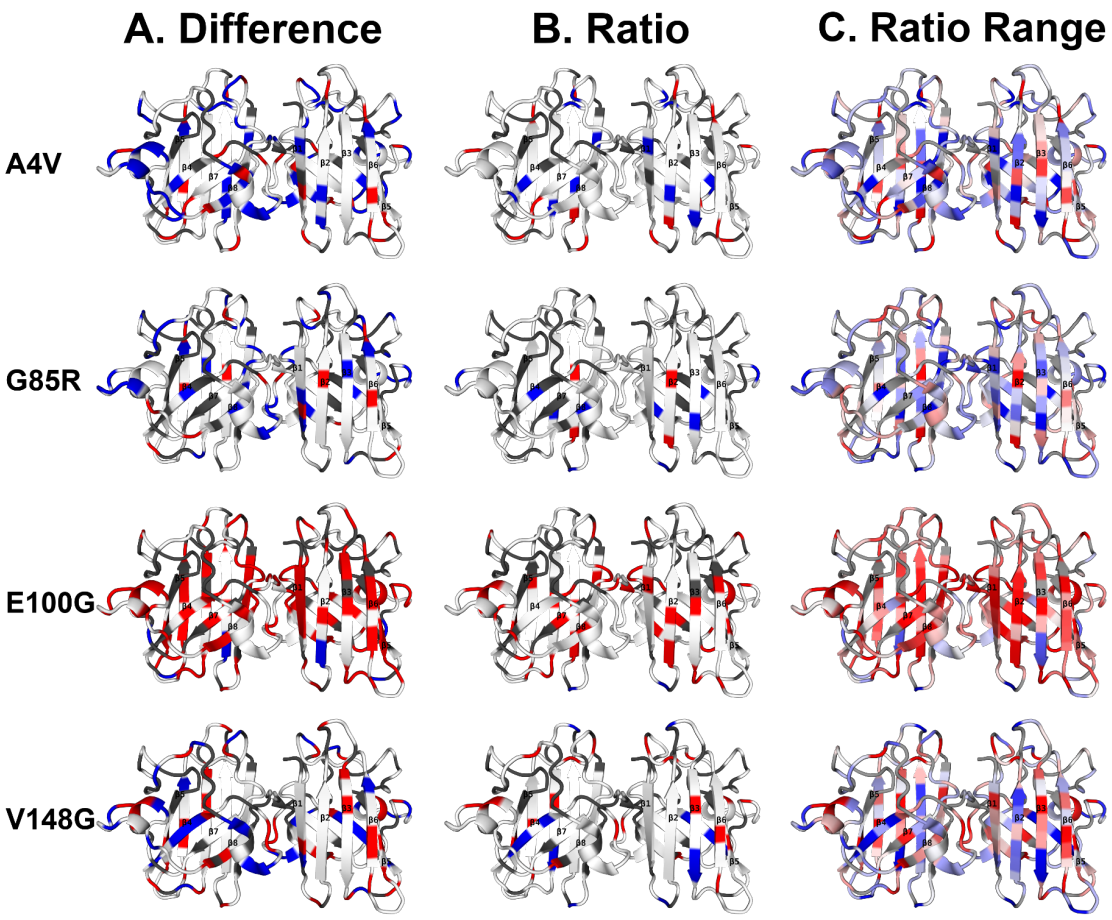



Figure 3.16: Comparison of folded and unfolding temperature coefficients for apoSH mutants. Differences, ratios and the ratio range are mapped onto 3D structures of holoSS SOD1 (1HL5, rendered with PyMol). (A) Differences in temperature coefficients are considered significant if their absolute difference is larger than 1 ppb/K. Red coloured residues are considered more dynamic in the unfolding temperature range, and blue is more stable in the unfolding temperature range. (B) For the ratios, red (blue) represents the unfolding temperature coefficient increases (decreases) by 50% (or 0.5) or more relative to the folded temperature coefficient. (C) Ratio range colours are the same as those for the ratio panel, however here any value smaller (larger) than 1, where 1 represents no difference between folded and unfolding values, is coloured blue (red). White represents no detectable change based on the respective cutoffs, and grey indicates residues with no data available.

Since A4V is further into its unfolding endotherm than G85R, there are more changes throughout the structure (Figure 3.14). Interestingly, A4V also shows both increases (red, more negative values) and decreases (blue, less negative values) in the temperature coefficient values (Figure 3.16), corresponding to loss and gain of structure upon protein unfolding, respectively. The beta-strands show some residues becoming

more structured, whereas near the ends of the beta-strands becomes less structured. The residues that show significant changes also happen to be residues near residues that showed changes in CSP when comparing 19 °C to 34.5 °C (Figure 3.15B). This suggests that the local environments are changing for residues due to local loss of structure, and it can manifest as CSP and/or temperature coefficient differences.

Along the same line as A4V, V148G also experiences both increases and decreases in temperature coefficients throughout the structure (Figure 3.16). The start of ß3, the entirety of ß6 and parts of the large loops seem to lose structure upon heating, while the end of ß3, the entirety of ß4 and ß8 seems to gain structure. Although it is sensible to consider that temperature coefficients would only increase (more negative) due to loss of structure and become more random coil-like, it is also not entirely surprising local areas experience structuring. Decreased temperature coefficients (less negative) may be a result of changes in the intrasubunit bonding in an attempt to find a new local minimum due to unfolding in other regions, and/or could also include stabilization due to intersubunit association.

However, E100G, one of the mutants furthest along its unfolding transition, shows mostly increases (more negative) in temperature coefficient values in the unfolding range than the folded range (Figure 3.16). While the CSP data do not highlight these changes, the global cutoffs could mask the subtle changes occurring in the structure. Here, the comparison of the folded and unfolding temperature coefficients suggests a more global

increase in dynamics occurring throughout the structure, which is likely due to global protein unfolding. Considering E100G and V148G unfolding temperatures are well into the DSC endotherm but show differences in CSP and temperature coefficient data, E100G's data could be highlighting that it may unfold differently relative to the other mutants assessed here.

It is interesting to consider the minimal decreases (smaller absolute value) in temperature coefficients observed for E100G relative to the other mutants. The comparison of the folded and unfolding temperature coefficients highlights the local changes in dynamics that occur because of protein unfolding. Additionally, the comparison at different points along their DSC endotherm may be emphasizing the consideration that other processes like restructuring and aggregation may arise. Deconvoluting the temperature coefficients may be more complicated upon unfolding, as now the residues may access additional states. The observed chemical shift is due to weighted averaging of all states the residues can access. With unfolding, although the residues may no longer access some folded states, additional states arise due to unfolding, and consequently protein association may also initiate. Thus, comparing the folded and unfolding temperature coefficients must be done with the important consideration to all possible states available.

Overall, local structural unfolding is identifiable by comparing the unfolding and folded temperature coefficients and CSP data. Since the highest unfolding temperature

is close to physiological temperature for these destabilized mutants, the unfolding data presented here may suggest the dynamics available to the protein *in vivo*. For instance, loop 7 in V148G shows enhanced CSP values relative to pWT with temperature, as well as significant changes in the temperature coefficients between the folded and unfolding ranges. Previously loop 7 has been reported as an aggregation prone region[59,222], and with the lack of structuring from metal cofactors and the disulfide bond, in addition to the shift in dynamics for V148G, loop 7 could reasonably serve as a structural hotspot for aggregation. Identifying local areas of enhanced or different dynamics at physiological conditions may be key in understanding protein misfolding, and consequently protein aggregation.

### 3.3.6 Curvature

Amide proton temperature dependences that are nonlinear can report on alternative states accessed by proteins throughout the temperature ramp. Since assessing extent of curvature is heavily dependent on the quality of the dataset, all spectra in the folded range and unfolding range for each mutant are used (temperatures ranges summarized in Table 3.6). However, ensuring curvature is statistically significant is important to remove false curvature, thus residues are only considered curved if they meet the stringent cutoffs of an $R^2 > 0.8$ and a mean p-value $< 0.05$ for a quadratic fit. With that in mind, apoSH mutants are evaluated for curvature in hopes to identify regions accessing alternative conformations that may be related to either native maturation processes or non-native association.

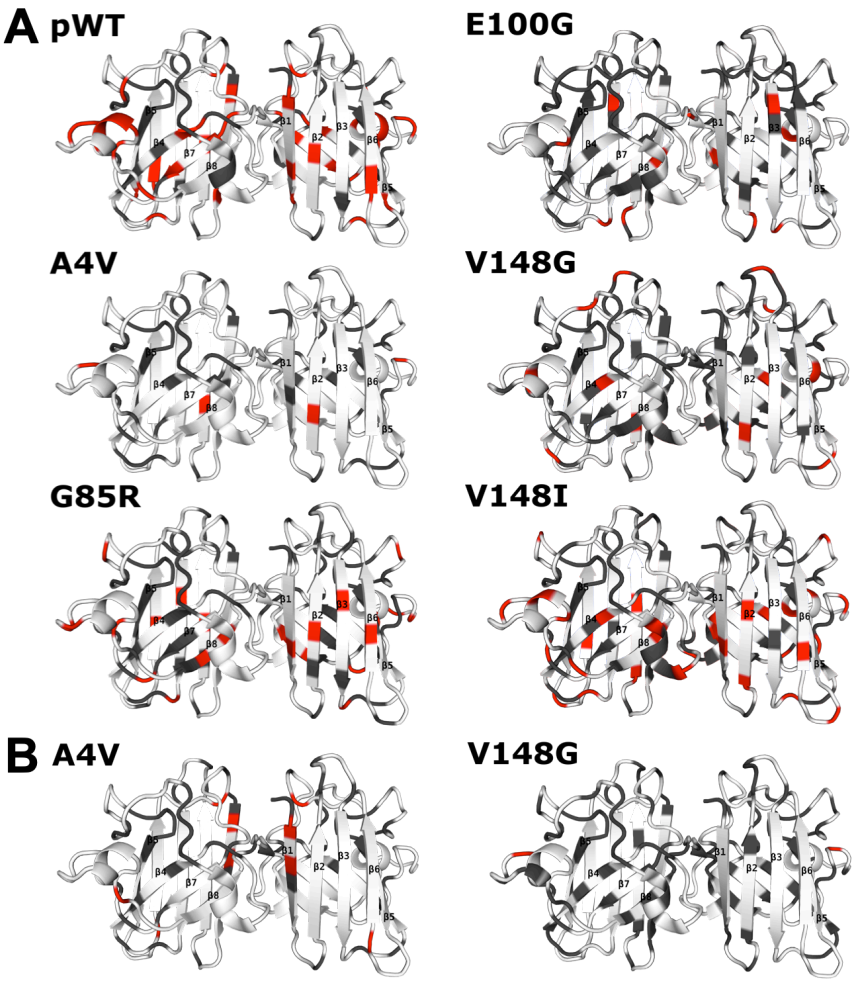



Figure 3.17: Curvature of amide proton temperature coefficients in apoSH pWT and mutants in folded and unfolding temperature ranges. Residues are considered curved if they show a $R^2$ value > 0.8 and a mean p-value < 0.05 for a quadratic fit. For folded temperature range (A) and the unfolding temperature range (B), residues are coloured red for residues that show curvature, white for non-curved residues, and grey for residues with no data available.

### 3.3.6.1 Curvature is mostly present in hydrogen bonded residues in apoSH pWT

Since apoSH SOD1 lacks all the maturation steps, it is more dynamic than its other states, which could suggest apoSH SOD1 has more conformations available and subsequently more curvature. However, apoSH pWT only shows 22 residues with curvature, of which 16 are residues known to hydrogen bond in the holoSS state (Figure 3.17A).[221] The hydrogen bonded residues specifically could be suggesting alternative

states that may be changing the hydrogen bonding. These residues in general are located on edge strands (ß1 and ß6), at the ends of other beta strands, and residues involved in forming the short helix in loop 7 in the mature enzyme. However, the location of these residues may suggest differences in the alternative state being accessed.

The position of the curved residues based on their secondary structural element in the holoSS state may suggest that some residues in apoSH pWT are involved in non-native processes, while others may be hinting formation of hydrogen bonds similar to the holoSS state. For the residues that show curvature at the edges of beta strands and at the short helix in loop 7, this curvature could be an indication of structuring. The beta strands may be transiently frayed at the ends in the apoSH state and similarly the short helix is not formed in the apoSH state due to missing maturation events. Since both the ends of beta strands and the short helix are structured in the holoSS state, the curvature for these residues may suggest the two states being accessed are the unstructured and structured states.

Meanwhile, the curvature observed for ß1 and ß6 could also align with the conclusions drawn above, but it may also suggest an additional process of strand unfolding. The residues that show curvature in ß1 and ß6 in apoSH pWT are also residues near or involved in disrupted hydrogen bonding based on their temperature coefficient values (discussed in section 3.3.3.). With the disrupted hydrogen bonds, ß1 and ß6 might exist in an alternative state that involves detaching from the beta-barrel. Overall, apoSH

pWT shows extensive curvature in hydrogen bonded residues where the alternative state arises from change in structuring.

Similarly to apoSH pWT, holoSS pWT has 28 curved residues, and most of them (24) are hydrogen bonded residues.[66] For both apoSH and holoSS pWT, the curvature seen in predominantly hydrogen bonded residues may point to altered hydrogen bonding compared to the ground state. However, where in holoSS pWT, the altered hydrogen bonding could be forming a different hydrogen bond, due to its enhanced dynamics, apoSH pWT curvature may have an alternative state where the hydrogen bond no longer exists. Additionally, ß5 and ß6 edge strands in holoSS pWT show extensive curvature, which is interpreted as an alternative state where the edge strands are no longer attached to the beta barrel, akin to the observations for ß1 and ß6 in apoSH pWT. Collectively, the trends between holoSS and apoSH pWT suggest hydrogen bonded residues are likely to access alternative states, and the alternative states involve changes in hydrogen bonding.

3.3.6.2 ApoSH mutants have less curved residues due to their enhanced dynamic nature relative to pWT

For the apoSH mutants, curvature may be highly dependent on the dynamics of the specific mutants, and the hydrogen bonding pattern. Since apoSH V148I is pWT-like in almost all respects discussed above, it is unsurprising the curved residues in V148I are similar in position and number (16) to pWT (Figure 3.17A). However, V148I does show more curvature for residues near the site of mutation and near the dimer interface relative to pWT, and the ratio of hydrogen bonded and non-hydrogen bonded curved residues is

more even. These subtle differences may suggest an alternative state available to V148I that is not seen in pWT. Although V148I is consistently pWT-like in many experiments, perhaps these subtle differences in curvature may be key in understanding its role in disease.

Alternatively, the rest of the destabilized mutants show a significantly reduced number of curved residues relative to pWT. A4V, G85R, E100G and V148G had only 2, 12, 6 and 6 curved residues, respectively (Figure 3.17A). Aside from A4V, the number of curved residues correlates with the mutant's respective melting temperatures (Table 3.1). A4V only shows two residues as curved, which may be related in part to its less native-like fold as discussed earlier. Additionally, A4V, G85R, and E100G temperature coefficient data suggest these mutants have more dynamic residues around the short helix in loop 7, and previous literature supports their inability to transiently access the native helix from the apoSH state.[55] Interestingly, these 3 mutants are the only mutants to show no curvature in the native-helix, which is present to some extent in pWT, V148G, and V148I. Perhaps the curvature observed in the native helix is hinting at the ability to structurally mature, and lack of curvature in some mutants suggests perturbed maturation.

Moreover, where the apoSH mutants show dramatically decreased curved residues, the holoSS mutants, aside from V148I, show increased curved residues, mostly in hydrogen bonded residues.[66] A potential explanation may involve hydrogen bonding

and dynamics. While holoSS mutants experience alternative states around their hydrogen bonded residues accessing potentially new hydrogen bonded patterns, the apoSH mutants are less structured than holoSS mutants, and are missing critical hydrogen bonds. These missing hydrogen bonds in apoSH SOD1 may subsequently impact the ability to access the alternative state seen in holoSS SOD1. However, an important consideration to discuss is lack of curvature does not suggest lack of alternative conformations. Curvature is only detectable when there are only two prominent states fast exchanging and have comparable energy. With multiple accessible states, or slow exchanging states, or states with larger differences in energy, the residues will not show detectable curvature. Thus, the decrease in curved residues for apoSH mutants may not be due to lack of curvature, but lack of detectable alternative states within the constraints.

3.3.6.3 Curvature between the folded and unfolding temperature ranges show complex behaviour between A4V and V148G

Changes in curvature are also evaluated between the folded and unfolding ranges of temperature. However, due to insufficient number of spectra to statistically evaluate curvature, curvature could not be assessed for G85R and E100G unfolding ranges. Folded and unfolding curvature data for A4V and V148G show differences in curvature both in the number of residues and the residues that present curvature (Figure 3.17B). Only one residue is curved in V148G, whereas A4V has 5. The loss or gain of curvature could be related to the extent of protein unfolding and/or protein association. The unfolding A4V curvature data show residues predominately in ß1, and coincidently near

the site of mutation, are curved. This may be hinting at an alternative state where ß1 may be either unfolding and/or associating, which aligns with ß1 having both more structured and dynamic residues in the unfolding range relative to the folded range, and otherwise very few changes. However, V148G may possess significantly more dynamics since it is further into its unfolded endotherm, and the alternative states may no longer be determined for reasons mentioned above. Overall, unfolding curvature introduces many more considerations, such as processes that may only exist due to local unfolding.

## 3.4 Conclusion

Collectively the data presented here highlight the local dynamic changes in apoSH mutants both relative to apoSH pWT and holoSS SOD1. The CSP data show sites of mutations being important for assessing local structural changes relative to pWT. Additionally, extending the previously published urea denatured SOD1 chemical shifts with calculated SOD1 random coil temperature coefficients to assess the extent of random coil in SOD1 in the folded range, the beta-barrel for most mutants is intact, whereas the large loops (4 and 7) have random coil-like chemical shifts. The temperature coefficients also further identify local changes in dynamics that correlate strongly with literature and may ultimately suggest hotspots for protein misfolding and consequently aggregation.

Here, we also explored temperature dependence in the folded and unfolding temperature ranges. The CSP for comparing to folded pWT helped identify areas for

potential local structural changes due to thermal unfolding, while the temperature coefficients highlighted changes in dynamics upon protein unfolding. Lastly, curved residues are predominately hydrogen bonded residues, and the number of curved residues decreases in destabilized mutants. In summary, using amide proton chemical shifts can offer an atomistic insight on local dynamics that are otherwise overlooked by more crude biophysical methods.

# Chapter 4: Exploring changes in local protein dynamics through $^{1}H$-$^{15}N$ cross-peak shape for immature Superoxide Dismutase-1 variants

Author Contributions:

Harmeen Kaur Deol: conceptualization, methodology, data acquisition, resources, validation, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft; Bao Khanh Nguyen: methodology, resources, validation, formal analysis; Jeffrey Alexander Palumbo: NMR data acquisition; Dalia Naser: NMR data acquisition, investigation; Duncan Mackenzie: NMR data acquisition, investigation; Michael Vincenzo Tarasca: methodology, NMR data acquisition resources, investigation; Elizabeth M. Meiering: conceptualization, supervision, project administration, funding acquisition.

## 4.1 Context

With all these different aggregate polymorphisms, it is clear point mutations and solutions conditions influence the aggregate structure. However, it is unclear how either influence the protein dynamics and in turn the aggregation pathway(s). In this chapter, the protein dynamics for apoSH pWT SOD1 using intensity profiles and linewidth changes throughout the folded temperature range are investigated. In addition to the temperature-influenced dynamics, we also probe for the differences in dynamics between apoSH pWT and apoSH mutants. Additionally, the temperature at which maximum intensity occurs

correlates to the temperature the residue shifts from its folding to unfolding environment, which allows for identifying areas of local protein unfolding throughout the temperature ramp. Lastly, linewidths are used to understand potential dynamics the protein may access throughout the heat ramp. However due to the marginal stability and decreased structure of apoSH SOD1, the crowded spectra are a limiting factor in these analyses. Collectively, the NMR peak shape methodology and data analyses of apoSH SOD1 and its variants presented here highlight critical local differences in dynamics through the structure, as well between apoSH mutants.

## 4.2 Introduction

Protein structure and dynamics are central to protein stability, protein folding and protein association. However, protein structural dynamics can be hard to characterize due to the timescale of the dynamics or lack of high-resolution detail. Fortunately, protein NMR offers atomistic insight on protein structure and dynamics at a range of timescales. Each NMR peak consists of a chemical shift, peak intensity and linewidth, all reflecting the local chemical environment, population and local dynamics, respectively, of the corresponding residue.[78] In Chapters 3 and 5, we explored the differences in chemical shift, and here peak intensity and linewidths methodology and data will be explored. While chemical shift and temperature coefficient data report on local structure differences, it fails to consider different chemical environments a residue can be exchanging between; a chemical shift can show no changes throughout a changing solution condition, but the exchange processes available to the residue can change and alter the peak shape. Simply, peak intensity can report on the number of nuclei resonating at a given frequency, whereas

linewidths change based on relaxation rates. However, both peak intensity and linewidth are influenced by additional dynamic processes, such as a residue exchanging between two distinct chemical environments.

Protein dynamics are difficult to capture in a single 2D HSQC experiment, and rather require multiple HSQC spectra over changing conditions. The condition is used to probe/amplify protein dynamics and could include time, temperature, protein concentration, pH, and ligand titration among others.[66,203,211,230] By collecting multiple spectra over changing conditions, depending on the timescale of the event, peaks in the NMR spectra may change as their associated residues sample different states throughout the varying conditions.[231–233] For instance, this could be done for a ligand titration, where residues that may be involved in binding to the ligand can access both a bound and unbound state and depending on the relative exchange between the two states, the chemical shift, intensity and linewidths may change to reflect the shift in dynamics. In general, peak shape analysis has been used for complex reaction mechanisms, such as non-2-state binding[234,235], protein folding[83,236], and allostery.[237,238]

Here, we collect 2D $^{1}H^{15}N$-HSQC spectra throughout a temperature range to characterize structural dynamics for apoSH SOD1. Initially, the protein dynamics are assessed for apoSH pWT SOD1 using intensity profiles and linewidth changes throughout the folded temperature range. In addition to the temperature-influenced dynamics, we also probe for the differences in dynamics between apoSH pWT and apoSH mutants. Additionally, the temperature at which maximum intensity occurs correlates to the temperature the residue shifts from its folding to unfolding environment, which allows for identifying areas of local protein unfolding throughout the temperature ramp. Lastly,

linewidths are used to understand potential dynamics the protein may access throughout the heat ramp. However due to the marginal stability and decreased structure of apoSH SOD1, the crowded spectra are a limiting factor in these analyses. Collectively, the high-resolution analyses of apoSH SOD1 and its variants presented here highlight critical local differences in dynamics through the structure, as well between apoSH mutants.

## 4.3 Methods

### 4.3.1 Protein preparation for NMR

Proteins for NMR samples were expressed using M9 minimal media under different growth conditions. Briefly, BL21 non DE3 *E. coli* were used, which do not confer any antibiotic resistance. Once transformed with the pHSOD1ASlacI$^{q}$, plasmid for expressing SOD1, which confers ampicillin resistance, cells were grown overnight in LB broth with ampicillin. Next, 10 mL of the overnight culture was used to inoculate 1 L of M9 minimal media, which contains $^{15}NH_4Cl$ as the sole nitrogen source. Cells were grown to $OD_{600}$=~0.6, and then induced with IPTG. All protein samples underwent a copper charging and heat treatment before being purified by hydrophobic column chromatography to ensure proper metalation. Copper was added (final concentration of 1.4 mM) and then the sample were heated for 20 minutes at 70 °C. The solution had ammonium sulfate (final concentration of 3 M) added and then was filtered through a nitrocellulose filter (0.45 µm cutoff) prior to applying onto a hydrophobic-interaction column. The protein was eluted and placed in dialysis tubing (6-8 kDa) to be dialyzed against milliQ $H_2O$ for 4x8-hour exchanges. With the Amicon ultrafiltration membrane, the

volume of the sample was reduced, and the concentration was checked with the use of the UV-vis spectrophotometer at 280 nm. Aliquots were then syringe-filtered into Eppendorf tubes and flash frozen in liquid nitrogen. Metals were removed via dialysis of holo SOD1 against ethylenediaminetetraacetic acid (EDTA) at pH 3.8.[208] Another set of exchanges were done with 50 mM sodium acetate, 100 mM NaCl at a pH of 3.8. The last set of exchanges were against milliQ water with one final exchange against 1 mM HEPES at pH 7.8. Metal contamination was assessed by PAR assay, initially developed by Crow *et al*.,[209] and further modified by Colleen Doyle and Heather Primmer as described previously.[210]

Reduction of the disulfide bond was performed as described previously.[65] The method used the reducing agent tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl). Initially, oxidized apo SOD1 was denatured using 2 M guanidine hydrochloride (GdnHCl) and 20 mM HEPES for a pH of 7.8. TCEP was then added to the sample, and it was placed in a desiccator for one hour before the sample was exchanged into buffer (1 mM TCEP and 20 mM HEPES at pH 7.4). The success of the reduction was checked by running an SDS-PAGE gel. Preparation of the NMR sample required concentration of the sample to ~15-20 mg/mL SOD1, with a final sample volume of ~450 mL. 50 µL of $D_2O$ and sodium trimethylsilylpropanesulfonate (DSS) (final concentration of 1 mM) was added and the sample was transferred to an NMR tube (final sample volume of 500 mL). Upon final sample preparation, we also employed an initial $^1H$ 1D NMR experiment (p3919gp,

Appendix H) to confirm both the TCEP reduction state and the metalation status (Appendix D and E)

4.3.2 NMR experiments

$^{1}H^{15}N$-HSQC spectra were acquired at varying temperatures from 1.2 °C to 38.6 °C using a Bruker Avance 600 MHz spectrometer, using the acquisition parameters described in Appendix J with a D1 of 0.80. The sample was equilibrated for 20 minutes at each temperature prior to acquisition, as described by Doyle *et al*.[66] All spectra were referenced to DSS. Spectra were processed in TopSpin 4.0.7, and analyzed in CCPN 2.4.2. Sequence specific resonances were previously determined for pWT, A4V, G85R, E100G, and V148I apoSH SOD1 at pH 7.4.[54,55]

In CCPN 2.4.2, the intensity and linewidths were fit only in the direct ($^{1}H$) dimension using Gaussian peak shape with a boxwidth of 1. Due to significantly decreased dispersion in the indirect ($^{15}N$) dimension relative to the direct dimension, only the direct dimension was considered. Peak intensities were normalized to the TRP-32 folded side chain and slopes were fit to a line using Microsoft Excel. Maximum peak intensity was assessed using Microsoft Excel. The $^{1}H$ linewidths were estimated from the width of the peak at half of the maximum height of the Gaussian peak fit.

## 4.4 Results and Discussion

Unlike the highly stable holoSS SOD1, apoSH SOD1 is marginally stable and cannot be heated past physiological temperatures without protein unfolding occurring. The mutant set chosen ranged in location, disease duration, and most importantly, stability (Table 4.1 and Figure 4.1). The variation in thermal stability means the folded temperature range varied significantly; stabilized mutants like V148I had a large, folded range similar to pWT, but destabilized mutants (melting temperatures less than pWT) like A4V had a smaller range.[65] However, to ensure enough data points are acquired, apoSH mutants were heated just prior to their unfolding, based on DSC thermograms[65] and tryptophan 32 side chain peaks (discussed in 4.4.1.1), with more intermediate temperature collected along the folded temperature range. For pWT, G85R and V148I, only the folded ranges are considered, but for A4V, E100G and V148G, both folded and unfolding temperature ranges are considered (Table 4.2).

Table 4.1: Summary of SOD1 mutant characteristics.

| Mutants | Location | Disease Duration (years)[a] | ApoSH $T_{0.5}$ (°C)[b] | HoloSS $T_{0.5}$ (°C)[c] |
|---|---|---|---|---|
| **pWT** | - | - | 48.5 | 92.7 |
| **A4V** | ß1 | 1.2 | 36.3 | 86.7 |
| **G85R** | ß4 | 6.0 | 40.7 | 77.5 |
| **E100G** | ß6 | 4.7 | 33.2 | 86.2 |
| **V148G** | ß8 | 2.1 | 34.0[d] | 86.9 |
| **V148I** | ß8 | 1.7 | 51.0 | 92.7 |

[a]Average disease durations are from Wang *et al*.[212]

[b]Melting temperatures are from Vassall *et al*.[65]

[c]Melting temperatures are from Stathopolous *et al* and Doyle *et al*.[52,66]

[d]Based on visual inspection.

- not applicable.

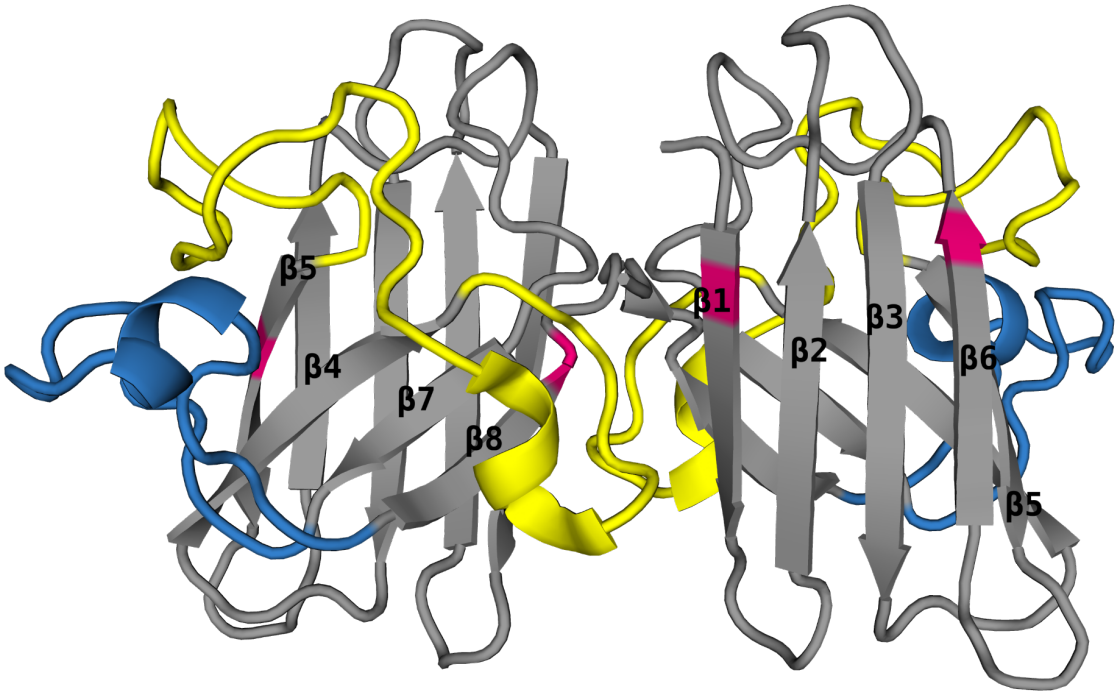


Figure 4.1: ALS-associated SOD1 mutations mapped onto holoSS SOD1. Mutations studied here are mapped onto the 3D structure of holoSS SOD1 (PDB 1HL5, rendered with PyMol) in magenta. A4V, G85R, E100G and V148G/I are in ß1, ß4, ß6 and ß8, respectively. Loops 4 and 7 are loops that become structured with maturation events, and are coloured yellow and blue, respectively.

Table 4.2: Different temperatures and spectra ranges collected for apoSH mutants.

| Mutants | $T_{0.5}$ (ºC)[a] | $T_{unfolding}$ (ºC)[b] | Folded[c] | Unfolded[d] |
|---|---|---|---|---|
| pWT | 48.5 | 36.0 | 4.85 - 35.4 (13) | - |
| A4V | 36.3 | 18.0 | 5.65 - 17.94 (9) | 19.8 - 34.5 (9) |
| G85R | 40.7 | 31.5 | 6.20 - 27.04 (10) | 29.29 - 31.50 (2) |
| E100G | 33.2 | 24.0 | 5.17 - 24.50 (9) | 22.23 - 33.65 (6) |
| V148G | 34.0 | 25.0 | 1.04 - 19.88 (12) | 20.11 - 36.70 (9) |
| V148I | 51.0 | 40.0 | 9.30 - 39.55 (13) | - |

[a]Melting temperatures are from Vassall *et al*.[65]
[b]Start of unfolding temperature from Vassall *et al*.[65]
[c,d]Folded and unfolding temperature ranges, respectively, for the apoSH mutants and their respective number of spectra acquired in brackets.

Additionally, NMR lineshape analysis is a useful tool commonly used for assessing binding of ligands to a biomolecule.[79–81] It can help assess thermodynamics, kinetics and structural information of binding events. In addition to these processes, lineshape analysis has been explored for comparing protein variants for differences in dynamics.[82] Unfortunately, lineshape analysis can be convoluted due to the numerous factors that impact peak shape and position; two of these key factors include exchange rate between processes and relaxation rate (Figure 4.2).[78,83–85] Glaves *et al.* highlight the complexities and limitations associated with peak shape analysis for a self-associating protein; the NMR assignments for aggregate-intermediates must be known for reliable interpretation on self-association.[239] Processes such as ligand binding that may be simplified to exclusively a bound and unbound state are easier to interpret. However, assessing lineshape changes throughout multiple spectra where processes such as dimerization,

aggregation and/or population of transient conformations are occurring convolute and limit lineshape analysis greatly.

For apoSH SOD1 mutants, population of both non-native and native dimers, aggregates, and transient conformations are possible along the temperature ramp.[12,54,56,65] Aside from these processes convoluting the sample, each of these species impact peak shape in different ways. Dimers, aggregates and transient conformations will have their own respective exchange rates with the monomer, whether that is the folded or unfolded monomer is an additional complication, as well as different relaxation rates. Impacts to exchange rates and relaxation rates will subsequently influence peak position, intensity and linewidths. Thus, with these potential complications, our aim is to attempt to analyze peak shape where conclusions are only drawn with multiple residues behaving comparably as opposed to isolated residues.

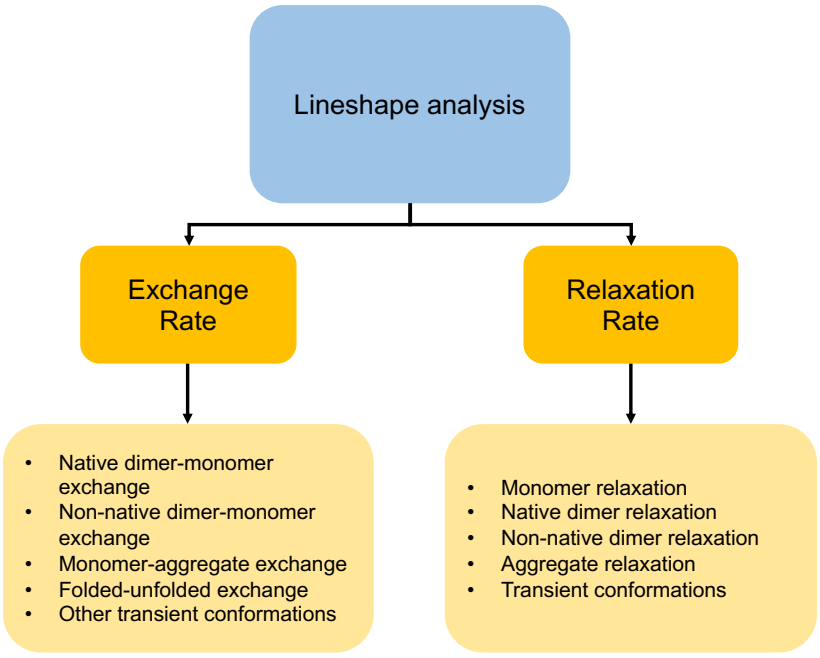



Figure 4.2: Different factors and processes that may impact lineshape analysis.

### 4.4.1 Intensity profiles

NMR signal intensity is directly correlated to the number of nuclei existing in the same chemical environment. If a residue under certain conditions experiences multiple chemical environments in response to accessing additional states, the signal intensity will change to reflect the additional chemical environments, as discussed above. However, so many factors influence peak intensity and therefore it is difficult to classify peak intensities from a single spectrum. In these cases, comparing peak intensity changes by collecting multiple spectra throughout a changing condition with a reference peak offers a more controlled analysis. Additionally, spectral crowding is a significant limitation, but any conclusions made are based on multiple residues which will limit errors; this limitation is further addressed in section 4.4.3. Previously, variant studies on RFP used deviation from average peak intensity for each residues as a threshold for assessing differences in dynamics[82], but here, the intensity of the apoSH pWT and mutant peaks are being monitored throughout a temperature ramp (from now on referred to as the intensity profile) using a side chain peak to normalize.

#### 4.4.1.1 Folded and unfolded Tryptophan-32 side chain correlate to extent of protein unfolding

Previously, Sekhar *et al*., identified two different chemical environments for the tryptophan-32 (Trp-32) sidechain that are in slow exchange, and thus result in two different chemical shifts (Figure 4.3A).[55] The two chemical shifts are considered to correspond to folded and unfolded conformers (Trp-32F and Trp-32U, respectively), and their relative peak volumes are roughly correlated with protein stabilities assessed by

DSC thermogram.[65] Here, the Trp-32F and Trp-32U intensity profiles are compared to the entire DSC thermogram, to assess if these peaks can also probe for the extent of unfolding.

The Trp-32F and Trp-32U peaks do not show up for all the apoSH SOD1s studied here (Figure 4.3B). For apoSH pWT, only a folded peak appears throughout the temperature range in the spectra collected. However, apoSH V148I does show a weak Trp-32U peak at higher temperatures (>37 °C). Both pWT and V148I have higher thermal stability relative to the destabilized mutants (Table 4.1) and should not show signs of protein unfolding, and accordingly they only present a Trp-32F peak at most temperatures. The Trp-32F intensity only increases for the entire temperature range, which is likely a consequence of peak shape changes due to the protein tumbling differently at different temperatures, as it occurs for all SOD1 mutants.[75] However, since the temperature range for both SOD1s is in their folded range based on their DSC thermograms, this increase in intensity for Trp-32F supports the protein remaining folded throughout the temperature range (Figure 4.3B).[65]

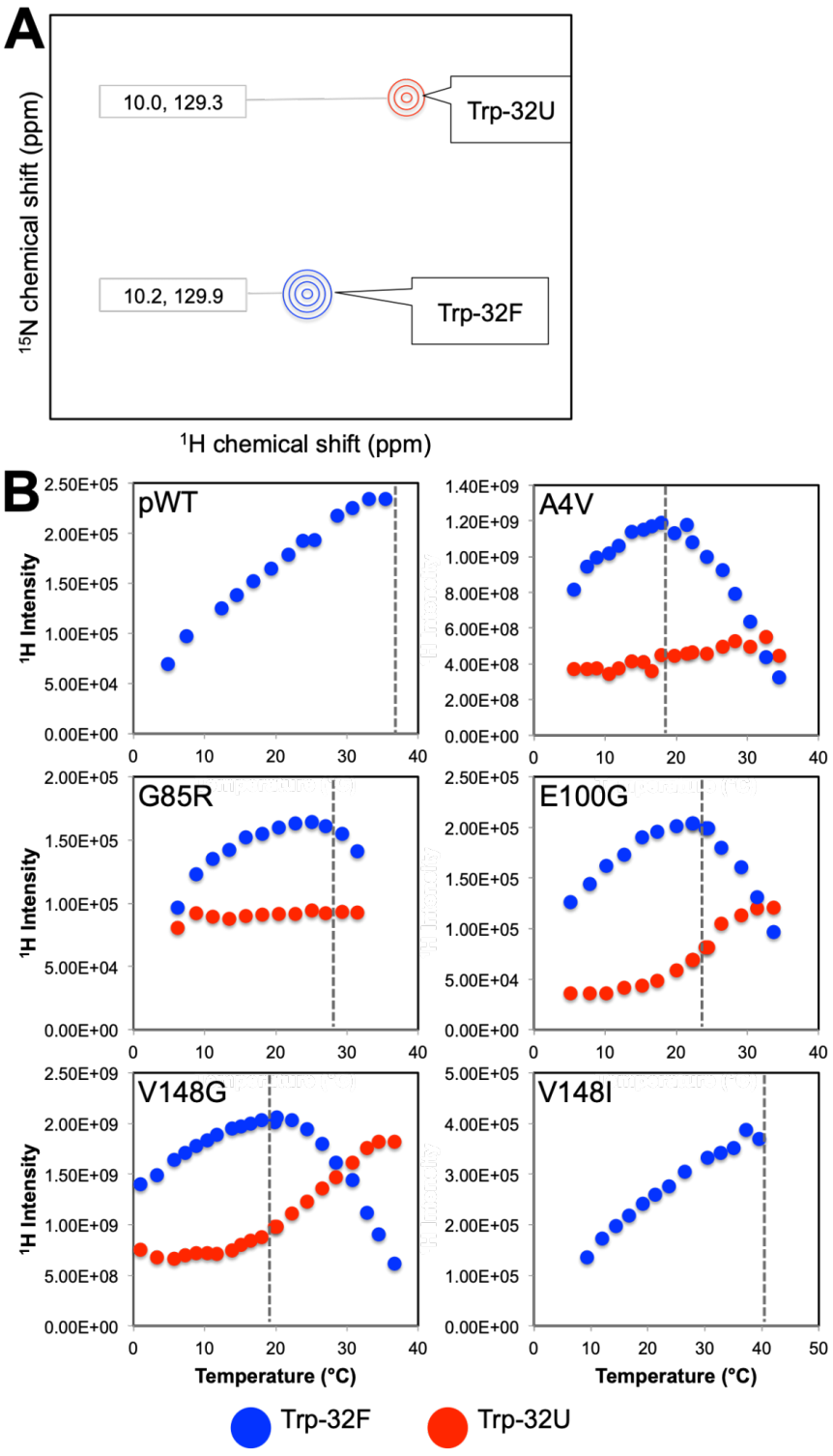



Figure 4.3: Trp-32F and Trp-32U chemical shifts (A) and peak intensities (B) throughout the temperature ramp. All peak intensities were normalized by receiver gain. **(A)** Trp-32's sidechain has an amide peak whose chemical shift varies based on whether it is in the folded or unfolded conformer, Trp-32F and Trp-32U, respectively. Both sidechain peaks are isolated and show up downfield from the amide region, making

them easy to locate and follow, as well as ideal reference peaks for assessing the intensity profiles. The chemical shift values for each peak is for illustrating their differences, and the respective values vary based on the mutant and temperature. **(B)** The Trp-32F (blue) and Trp-32U (red) peak intensities are evaluated throughout the temperature ramp. For pWT and V148I, no protein unfolding is observed and subsequently no Trp-32U peak is observed. However, for the destabilized mutants, A4V, G85R, E100G and V148G, both Trp-32F and Trp-32U show up at all temperatures, even in the folded temperature range. Throughout the folded temperature range, the TrpF-32F peak intensity increases, whereas the Trp-32U peak shows minimal changes. However, the Trp-32F and Trp-32U intensity begins to decrease and increase, respectively, when the protein is in the unfolding temperature ranges.

The destabilized apoSH mutants' data suggest the change in Trp-32F intensity reports on protein unfolding. For all mutants, the Trp-32F intensity increases for the temperatures prior to the protein unfolding observed by the DSC thermogram. However, for the destabilized mutants, A4V, G85R, E100G and V148G, their profiles show initial increases and then the intensity decreases at the higher temperatures. Interestingly, the temperature at which the initial decrease in intensity occurs coincides with the temperature at which the protein begins to unfold in their DSC thermograms.[65] This suggests there is a shift in the protein from the folded to unfolded conformer due to thermal unfolding. Thus, the Trp-32F is a great peak for tracking when the protein remains folded (intensity increases), and when the protein starts to unfold (decrease in intensity).

Since the Trp-32F intensity change can report on shifts from the folded to unfolded conformers, the Trp-32U profile should also agree based on previous findings.[55] We expect the temperature that the Trp-32F intensity starts to decrease, would also be the same temperature the Trp-32U intensity would increase since it corresponds to the unfolded conformation. For most of the destabilized mutants, the signal intensity does increase for the Trp-32U peak as the signal for the Trp-32F peak decreases, suggesting Trp-32U reports on the sample populating an unfolded state. However, unlike pWT and V148I, which do not show a reliable Trp-32U peak, the destabilized apoSH mutants show

a strong Trp-32U peak at all temperatures. Thus, our results suggest that Trp-32U intensity changes may be reporting on increases in unfolded population, but also potentially other non-native conformations.

With Trp-32U present in all spectra for destabilized mutants, it is unclear what conformation this peak is reporting on. One potential hypothesis is that the destabilized mutants may transiently access an unfolded conformer at all temperatures, and this change in chemical environment/state may not be detectable by DSC. As the protein is heated close to the start of its unfolding endotherm, the destabilized mutants start populating more of the unfolded state. In addition, apoSH V148I, a stabilized mutant, has the Trp-32U peak appear exclusively at temperatures close to the initial unfolding observed by DSC. Where the destabilized mutants may access this state due to their enhanced dynamics (see Chapter 3), V148I may only be able to transiently access this conformation at temperatures close to its unfolding. Additionally, in Chapter 6, Trp-32U peak intensity is discussed further and suggests its intensity changes can also be in response to protein aggregation. Overall, Trp-32F is exclusively an indicator for the folded population, and Trp-32U is useful for assessing unfolded protein as well as other processes (further discussed in Chapter 6).

#### 4.4.1.2 Trp-32F is a good candidate to normalize the intensity profile of SOD1 residues

Due to the large range of intensity profile behaviour, characterizing individual residues based on their intensity profiles becomes difficult. Hence, the need to normalize the residue intensity profiles is imperative for analyses, but choosing a consistently behaving residue for all mutants becomes difficult. An isolated peak from pWT cannot be

used since the same peak in mutants will behave differently due to the impact both temperature and mutations have on the local environment. Additionally, due to the extent that point mutations impact the protein stability and structure (see Chapter 3), backbone residues are not entirely reliable as their local dynamics may vary at certain temperatures. However, as discussed above, the Trp-32F is present in all spectra and since its intensity correlates well with DSC endotherms, it is understood to represent a folded structure.

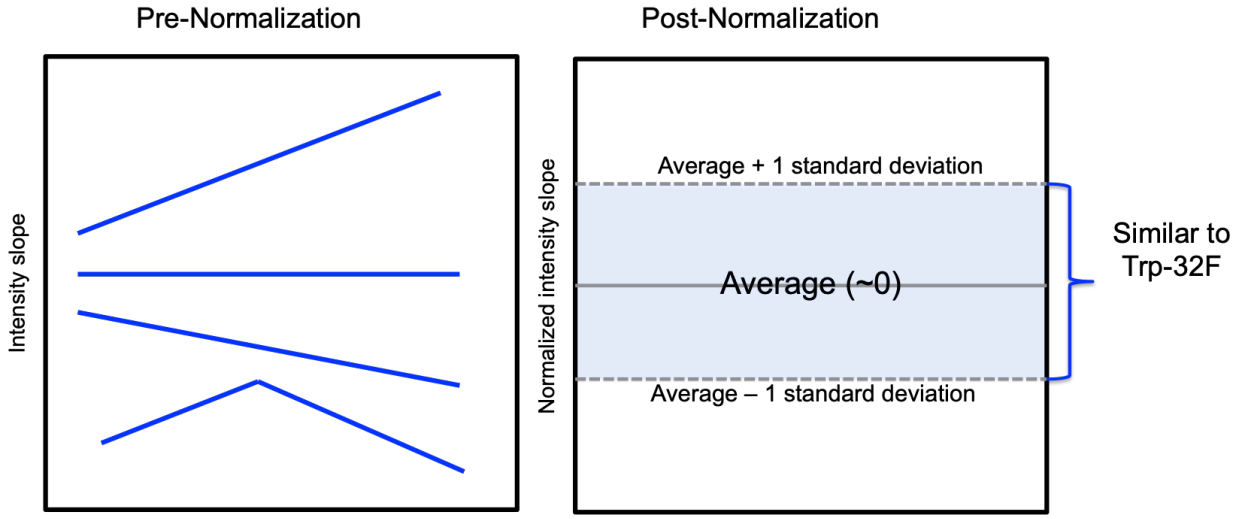



Figure 4.4: Pre-normalization intensity profiles and the interpretation of the post-normalized intensity slopes. The plot on the left shows all the potential intensity profiles seen for the residues before normalization. Due to the diversity in profile behaviours, interpretation is difficult. However, upon normalization to the Trp-32F peak, the interpretation of slope of intensity is now in reference to folded/structured conformer. The average slope value is close to 0 (solid grey line), and slope values within the average ± 1 standard deviation (blue shades box) are considered as behaving comparably to that of the Trp-32F peak, and thus are considered structured. Outside this range, the slopes are reporting on less structured/differently structured relative to the Trp-32F.

Due to the variability in intensity temperature dependence for all mutants, the need for an internal reference point is imperative. Thus, each residue intensity is normalized to their Trp-32F peak at all temperatures for apoSH pWT and SOD1 mutants' spectra, which removes the nonlinear influence of temperature or changes in protein concentration on peak shape, without changing the intensity profile shape. From here, the change in intensity throughout the folded and unfolding temperature range is considered (henceforth

referred to as the folded and unfolding intensity slopes). By normalizing all residue intensity profiles to Trp-32F, which represents the folded conformation, the signs and values of folded and unfolding intensity slopes now change. Post-normalized folded and unfolding intensity slopes that have values within the average ± 1 standard deviation mean their pre-normalized intensity slope is similar to the Trp-32F slope, and thus these residues are considered have comparable structure that responds similarly to temperature as the folded conformation (Figure 4.4).

However, larger differences between pre-normalized slopes and Trp-32F slopes result in a post-normalization intensity slope that has an absolute value larger than zero. For pre-normalization slopes that are larger, positive values relative to Trp-32F, the post-normalized intensity slope is now a positive value. In contrast, for a pre-normalization slope that is smaller than the Trp-32F, regardless if it is a small positive value, zero, or a negative value, the post-normalized intensity slope is negative. Where a post-normalized slope within the average ± 1 standard deviation suggests the residue behaves like the folded structure, a post-normalized slope with a value larger or smaller than the average ± 1 standard deviation cutoff suggests it deviates from the folded structure (Figure 4.4).

For this interpretation to hold true, we assume that the Trp-32F intensity profile is how all residue intensities behave in the folded structure. In reality, other folded intensity profiles likely exist that may be more or differently folded, and thus respond to temperature differently. With this acknowledgement and considering the average slope is centered closely around zero, post-normalized intensity slopes close to 0 are considered folded, and those that deviate from zero, are considered unstructured/less structured. This simplification is understandable considering the strong correlation of the Trp-32F intensity

correlation with the DSC endotherm, and its validation is considered further in the following sections.

4.4.1.3 Large loops deviate significantly from folded structure in apoSH pWT

For apoSH pWT, all spectra are collected in its folded range (Table 4.2). Hence, any data interpretation is for the native folded structure. However, as discussed above, without a point of reference, intensity profiles are difficult to interpret. Understanding the impact temperature has on the protein dynamics, subsequently the peak intensity, has multiple interpretations. Thus, by normalizing all residues' intensity at each temperature to their Trp-32F peak, the initial analysis is for assessing the extent of folded structure in pWT throughout the temperature range.

The range of normalized folded intensity slopes is large in apoSH pWT (Figure 4.5). By considering only slopes that are larger than the average +/-1 standard deviation as significantly different values, most of the residues have intensity slopes close to zero, thus like the Trp-32F value. This suggests that most of the structure has a folded environment similar to the Trp-32F peak. However, the large loops (4 and 7) are the only structural areas that show significant differences, meaning their intensity profiles deviate the most from the Trp-32F peak. This agrees with the extent of structure analysis done in Chapter 3 section 3.3.2 as well as literature suggesting the apoSH SOD1 beta-barrel is mostly intact, except for the large loops and dimer interface being highly dynamic.[55] This agreement in data further supports that the Trp-32F normalization is valid for assessing the extent of structure.

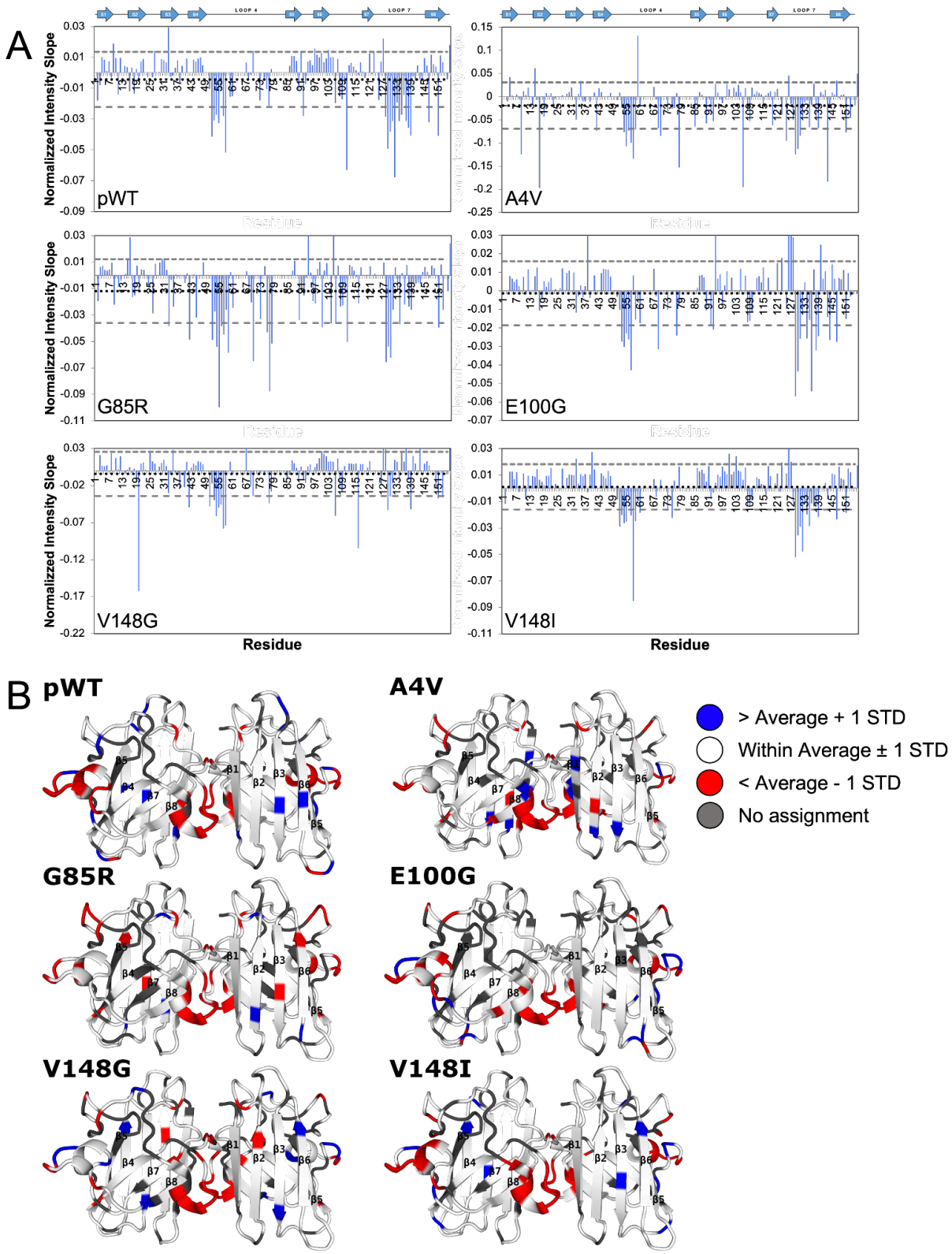



Figure 4.5: Post-normalized folded intensity slopes for apoSH pWT and mutants. **(A)** All the normalized intensity slopes (blue bars) are considered for the folded temperature ranges. The higher small grey dashed

line and the lower dashed grey line represent the average + 1 standard deviation and average - 1 standard deviation, respectively. The black dotted line represents the average, which shows up close to 0, the slope expected for the Trp-32F peak that the data are normalized by. **(B)** The folded intensity slopes are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). The residues are coloured based on colour legend provided, and each cutoff is based on the average and standard deviation values for individual mutants.

Furthermore, by averaging using rolling averages as done previously in Chapter 3 section 3.3.3.1, intensity profile values show clear patterns with less consideration for outliers. The rolling averages are averaged over 5 residues and highlight local differences throughout the structure without regarding secondary structural elements that may not be formed in the apoSH state. The most notable observation is that the rolling averages highlight the residue intensity profiles show grouped behaviour that correlates strongly with their secondary structural elements (Figure 4.6). The beta strands consistently show mostly small positive slope values similar to each other, but the large loops and loop 6 show large negative rolling intensity average slopes. This profile roughly matches that observed for the apoSH chemical shift-derived order parameters in Sekhar *et al*., 2015[54], which also further supports the normalization to Trp-32F is an appropriate measure of structure.

By assessing the intensity slopes using rolling averages, local areas of perturbation relative to the Trp-32F peak are identified. While the entirety of loop 7 deviates from the folded structure observed based on the similarity in slopes to Trp32F, only residues in the initial third of loop 4 seems significantly less structured (Figure 4.6). Although several residues are missing assignments in loop 4, there are enough in the loop to say there is a significant difference in behaviour throughout the loop. Additionally, loop 6 also shows a cluster of residues that have negative values. Even though these residues are not considered significantly different from the Trp-32F, most other residues

aside from loops 4 and 7 are positive with sporadic negative values, thus this difference in behaviour is worth considering. Similar to only parts of the loops behaving less structured or differently from the rest of the loop, the temperature coefficient data in Chapter 3 suggest only parts of the loops have disrupted hydrogen bonding.

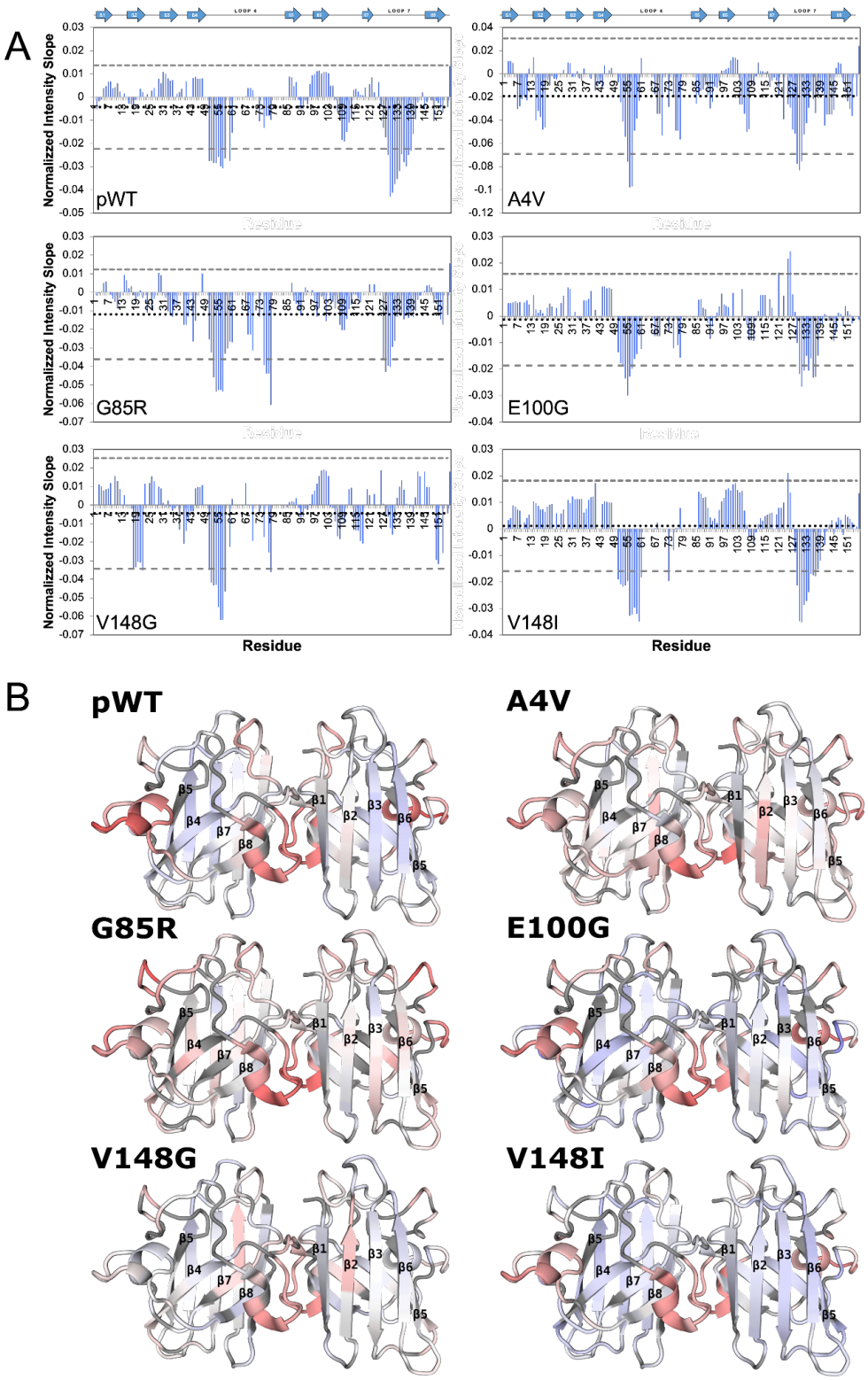



Figure 4.6: Post-normalized folded intensity slopes for apoSH pWT and mutants. **(A)** The rolling averages for normalized intensity slopes (blue bars) are considered for the folded temperature ranges. The higher small grey dashed line and the lower dashed grey line represent the average + 1 standard deviation and average - 1 standard deviation, respectively. The black dotted line represents the average, which shows

up close to 0, the slope expected for the Trp-32F peak that the data are normalized by. **(B)** The rolling averages of folded intensity slopes are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). The residues are coloured based on the average and standard deviation values for individual mutants, on a minimum (red) and maximum (blue) scale. Residues with values close to 0 are coloured white, and grey indicates residues where no data are available.

Without these maturation steps in apoSH, which subsequently impact the hydrogen bonds present in the mature enzyme (discussed in Chapter 3, section 3.3.3.3), the larger loops are less structured. The earlier parts of loop 4 are involved in forming a disulfide bond, parts of the dimer interface as well as metal-binding, and show the most disrupted hydrogen bonding in these residues relative to the rest of the loop. Loop 6 and 7 also shows heavily disrupted hydrogen bonding specifically around the residues that show large intensity slopes values. Overall, the lack of structuring suggested by the normalized intensity slopes can be explained by disrupted hydrogen bonding.

Additionally, while the beta-strands do not show significant intensity slope differences relative to Trp-32F, some strands consistently show differences. For pWT, ß3-ß6 show higher values than the rest of the beta strands, even though Trp-32 residue is in ß3. This suggests the folded environment is similar for all the beta strands, but there is slight variability for ß3-ß6. Aside from ß3, the other three beta strands (ß4-ß6) that show slightly higher intensity slope values relative to the other strands may be due to their enhanced dynamic nature since they are considered edge strands, as discussed previously in Chapter 3. Thus, the temperature dependence of the peak intensity may subtly vary across different folded structural components throughout the protein.

#### 4.4.1.4 ApoSH mutants show different extent of perturbations in their large loops based on their respective cutoffs

Each apoSH mutant shows very different ranges of post-normalized rolling average of intensity slope values, but similar areas of perturbations (Figure 4.6). For E100G and V148I, their profiles are very similar to pWT, with their large loops showing significant differences in the same residues as pWT. In contrast, A4V, G85R and V148G show different extents of perturbations in their large loops. ApoSH A4V does have higher intensity slope values in both loops, but fewer residues are considered significantly different relative to its Trp-32F peak. Similarly, G85R also has less residues that show significant differences in loop 7, but in loop 4 there is a gain of residues that show significant differences, specifically at the end of the loop. Lastly, V148G has no loop 7 residues that are considered significantly different relative to its Trp-32F intensity slope, but the same loop 4 residues show meaningful differences as pWT. Overall, the large loops are the key structural elements in each mutant that show the most perturbed behavior from their respective Trp-32F intensity slope.

#### 4.4.1.5 Comparing apoSH mutants by global cutoffs identifies differences in folded structure throughout mutants

Although looking at structural differences for individual mutants is useful for assessing local differences throughout the mutant relative to the rest of its structure, by comparing the mutant rolling average intensity slopes now using a global slope cutoff, we can identify critical differences due to point mutations. Even though all apoSH mutants show additional intensity profiles than pWT in their large loops, E100G and V148I rolling

averages are still considered pWT-like when compared on a global scale (Figure 4.7). ApoSH pWT, E100G and V148I rolling average intensity slopes suggest most of the structure, aside from the large loops, behaves similar to their Trp-32F peak when assessed using a global cutoff. E100G and V148I behaving pWT-like agrees with our findings in their temperature coefficient based hydrogen bonding pattern (Chapter 3, section 3.3.3.4), and literature findings suggesting their structure is pWT-like.[55]

Moreover, apoSH A4V, G85R and V148G have areas that behave less like their Trp-32F peak throughout the structure. For A4V and V148G, loop 4 residues near the dimer interface specifically and ß2 respond to temperature differently than their Trp-32F peak does. Furthermore, G85R shows differences throughout the structure specifically around the site of mutations, which agrees with its temperature coefficients being heavily perturbed there (Chapter 3 section 3.4.3.2). These differences highlight structural components that do not behave like their Trp-32F peak, and subsequently the folded conformer. Both A4V and V148G show inability to transiently see the native dimer as apoSH SOD1[55] and inability to form native-like dimers in other SOD1 states (Chapter 7).[87] Overall, A4V, G85R and V148G highlight that their Trp-32F peak is not representative of all folded structures as is the case with pWT, E100G and V148I; instead, these mutants may have larger differences in the extent of folded structural components throughout their structure, and thus Trp-32F does not reflect the global folded structure.

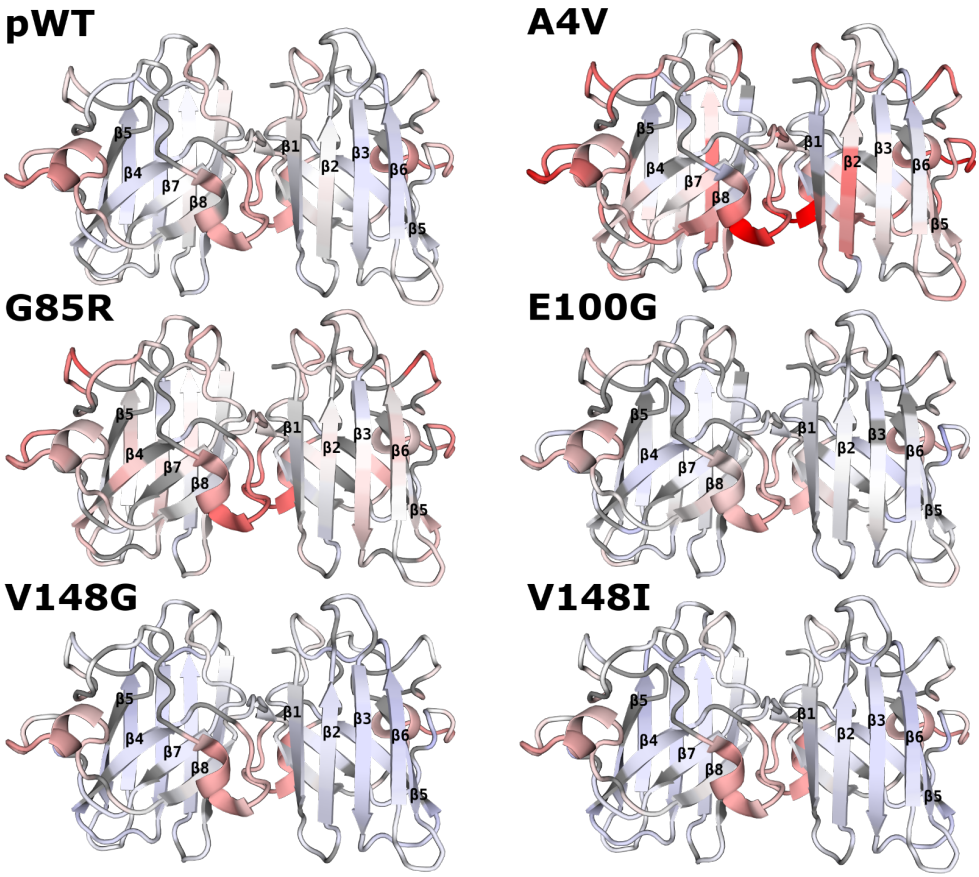



Figure 4.7: Post-normalized folded intensity slopes for apoSH pWT and mutants based on a global cutoff. The rolling averages folded intensity slopes are mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). The residues are coloured based on a minimum of -0.098 (red) and maximum of 0.098 (blue) scale, values that are based on the average standard deviations between pWT and mutants. Residues with values close to 0 are coloured white, and grey indicates residues where no data are available.

Furthermore, the average normalized intensity slope values are compared across apoSH pWT and apoSH mutants. Unsurprisingly, pWT, E100G and V148I average normalized intensity slopes are close to zero, -0.0043, -0.0013, and 0.0011, respectively, suggesting the Trp-32F peak intensity is reflective of most of the structure (Table 4.3). From A4V, G85R and V148G, only A4V and G85R have higher slope values, -0.0193 and -0.0119, respectively, while V148G has an average slope close to zero (-0.0045). A4V and G85R average slopes suggest the Trp-32F slope is not similar to most of the folded structure, which may be due to CSP values in Chapter 3 showing the highest deviation

from the native structure (pWT) for A4V and G85R. Additionally, most of this structure is potentially not as folded considering the high disruption of hydrogen bonding considered in Chapter 3. Moreover, V148G average intensity slope is close to zero, but unlike with pWT, E100G and V148I that also have values close to zero due to their Trp-32F peak being representative of most of the folded structure, V148G has larger absolute values, both positive and negative, throughout the structure that average to zero.

Table 4.3: Average folded intensity slopes for apoSH variants.

| | pWT | A4V | E100G | G85R | V148G | V148I |
|---|---|---|---|---|---|---|
| average<br>standard dev | -0.0043 (125)<br>0.02 | -0.0193 (126)<br>0.05 | -0.0013 (112)<br>0.02 | -0.0119 (117)<br>0.02 | -0.0045 (119)<br>0.03 | 0.0011 (124)<br>0.02 |
| | | | | | | |
| N-terminal B-sheet | 0.0051 (29) | -0.0027 (29) | 0.0065 (28) | -0.0022 (28) | 0.0017 (31) | 0.0098 (29) |
| C-terminal B sheet | 0.0031 (24) | -0.0129 (27) | 0.0047 (22) | -0.0042 (20) | -0.0022 (24) | 0.0074 (24) |
| N | -0.0179 (1) | - | - | -0.0194 (1) | - | -0.0118 (1) |
| B1 | 0.0027 (6) | 0.0069 (5) | 0.0047 (6) | 0.0059 (6) | 0.0106 (6) | 0.0070 (6) |
| loop I | 0.0060 (4) | -0.0226 (4) | 0.0059 (3) | -0.0089 (4) | 0.0130 (3) | 0.0041 (3) |
| B2 | 0.0000 (9) | -0.0206 (9) | 0.0028 (9) | 0.0016 (9) | -0.0173 (9) | 0.0089 (9) |
| loop II | 0.0028 (3) | -0.0058 (4) | 0.0048 (3) | -0.0095 (3) | 0.0119 (4) | 0.0057 (3) |
| B3 | 0.0066 (8) | 0.0020 (8) | 0.0022 (7) | -0.0036 (8) | 0.0032 (8) | 0.0105 (8) |
| loop III | 0.0018 (3) | -0.0109 (2) | 0.0099 (3) | -0.0034 (2) | -0.0167 (3) | 0.0114 (4) |
| B4 | 0.0081 (6) | -0.0040 (7) | 0.0101 (6) | -0.0136 (5) | 0.0018 (8) | 0.0108 (6) |
| loop IV | -0.0157 (19) | -0.0418 (20) | -0.0143 (18) | -0.0360 (20) | -0.0270 (18) | -0.0150 (19) |
| B5 | 0.0084 (4) | -0.0148 (5) | 0.0064 (4) | 0.0049 (3) | 0.0021 (5) | 0.0133 (4) |
| loop V | -0.0080 (4) | -0.0371 (4) | -0.0125 (4) | -0.0157 (4) | -0.0092 (4) | 0.0027 (4) |
| B6 | 0.0096 (8) | 0.0068 (8) | 0.0085 (7) | -0.0025 (8) | 0.0117 (8) | 0.0124 (8) |
| loop VI | -0.0040 (13) | -0.0113 (13) | -0.0006 (8) | -0.0086 (12) | -0.0006 (12) | 0.0052 (13) |
| B7 | 0.0009 (5) | -0.0077 (6) | 0.0063 (3) | -0.0055 (3) | -0.0147 (5) | 0.0054 (5) |
| loop VII | -0.0181 (21) | -0.0331 (20) | -0.0051 (20) | -0.0173 (18) | 0.0004 (19) | -0.0098 (20) |
| B8 | -0.0051 (9) | -0.0253 (9) | -0.0042 (9) | -0.0026 (9) | 0.0022 (6) | 0.0001 (9) |
| C | 0.0090 (2) | -0.0244 (2) | 0.0034 (2) | -0.0090 (2) | -0.0361 (1) | 0.0134 (2) |

Intensity slope values are averaged across all assignments, the beta-strands on N-terminal (ß1-3,6) and C-terminal face (ß4,5,7,8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

Moreover, when considering the average normalized intensity slopes for the N-terminal ß-sheet (ß1-3, 6) and C-terminal ß-sheet (ß4, 5, 7, 8), the values do not correlate with their corresponding values for temperature coefficients. In Chapter 3, the beta-strands on the N-terminal face and C-terminal face are averaged and show that the C-terminal face is less structured relative to the N-terminal face. However, the average

normalized intensity slope for both faces do not agree with this trend, rather the values between both faces are similar aside for A4V.

Interestingly for A4V, the C-terminal ß-sheet has a large value relative to its N-terminal face, -0.0129 and -0.0027, respectively (Table 4.3). The C-terminal face, specifically ß5 and ß8 have higher values than the other strands, but their temperature coefficient is not discernibly different. Since the C-terminal beta strands have similar chemical shift temperature dependence, it suggests there are similar extents of structure for these strands. However, the differences in temperature dependence of the normalized intensity values across the C-terminal beta-strands suggest the number of nuclei corresponding to that peak are impacted by temperature differently. In summary, the normalized intensity profile data are useful for identifying structural features that either do or do not behave like the folded conformer. Additionally, there are subtle differences in intensity profiles between apoSH mutants and pWT, which allude to changes in the SOD1 structure and dynamics that may play a role in its protein folding.

### 4.4.2 Max temperature: temperature corresponding to the maximum peak intensity throughout a temperature range

In section 4.4.1.2, the Trp-32F peak intensity changes correlated with protein unfolding. The peak intensity increases at temperatures where the protein remains folded, and then at temperature where the protein begins unfolding, the peak intensity starts to decrease. Thus, the maximum peak intensity corresponds to the temperature protein unfolding begins. We extend this analysis to all residues and their intensity profiles through their temperature ramps. Although not all intensity decreases correspond to

protein unfolding, here, any intensity decreases observed are considered a consequence of temperature-dependent dynamic changes. These changes are interpreted as the residue shifting from its folded conformation to an alternative conformation due to temperature, and even though the alternative conformation can also be folded, for simplification any decrease in intensity due to temperature is considered a consequence of local protein unfolding.

#### 4.4.2.1 The loop residues for apoSH pWT and SOD1 mutants show significantly different intensity profiles from the folded structure

By normalizing the intensity data to Trp-32F, the extent of structure is considered relative to Trp-32F as a representative of the folded conformation, rather than using the individual profiles, which varied too considerably to interpret. Upon analyzing normalized data, certain regions are identified to have less structure than the rest of the protein. For apoSH pWT, the large loops and parts of loop 6 are considered less structured and can be explained by disrupted hydrogen bonding (see Chapter 3). For the residues in these areas of differences, their profiles can now be individually considered.

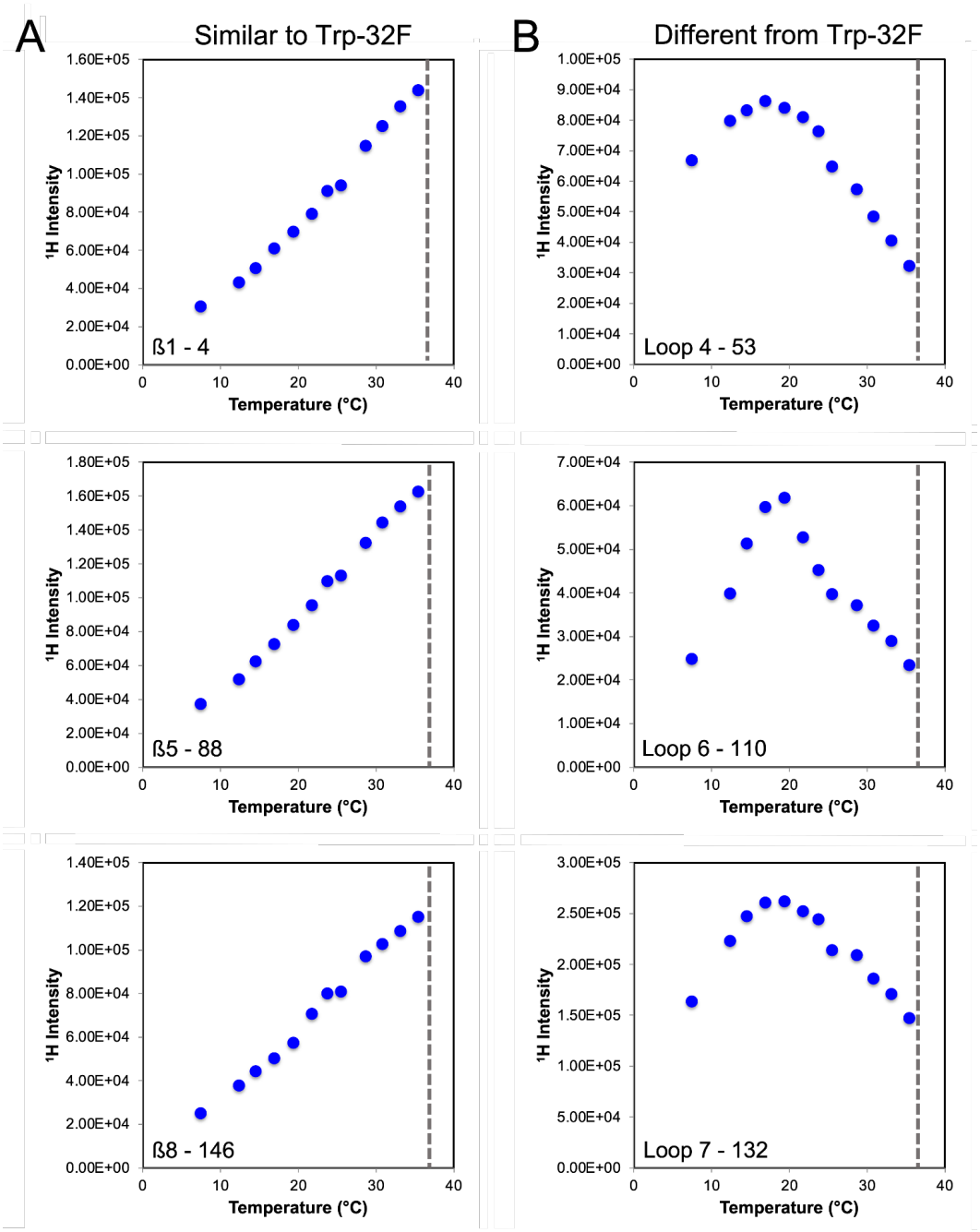



Figure 4.8: Intensity profile for residues that behave similar to (A) or differently from (B) Trp-32F for apoSH pWT. The dashed vertical line indicates when protein unfolding is expected based on the DSC endotherms reported in Vassal *et al*.[65] **(A)** Residues throughout the structure that showed similarity in intensity profiles to Trp-32F are chosen, which subsequently also have similar folded intensity slopes. Thus, these residues are considered to be in a folded/structured state similar to the one the Trp-32F is considered to report on. **(B)** However, the residues in the large loops, 4 and 7, as well as some in loop 6, show different profiles than Trp-32F, which explains the significantly different intensity slopes. The loops showing differences to the Trp-32F intensity profile as well as displaying behaviour reminiscent of protein unfolding at the high temperatures, suggests the loops are in a different folded state than Trp-32F reports on.

Since apoSH pWT spectra are collected throughout its folded temperature range, most residues show similar intensity profiles to the Trp-32F peak. The residues that do not show significant differences in intensity slopes to Trp-32F (Figure 4.4 and 4.6), also show a comparable increasing intensity profile throughout the temperature ramp to the Trp-32F peak (Figure 4.8A). In contrast, the residues that show the most differences in values relative to the Trp-32F peak, the larger loops and parts of loop 6, show very different intensity profiles (Figure 4.8B). While Trp-32F peak intensity only increases throughout its folded temperature range, these residues show initial increases in intensity followed by decreases at the higher temperatures.

Moreover, this intensity profile with an initial increase in intensity accompanied by a decrease in intensity is reminiscent of the Trp-32F for destabilized mutants that enter their unfolding temperature ranges (section 4.4.1.2, Figure 4.2B and 4.9). This is interesting since the temperatures that correspond to the decreased intensity are still within the folded range for pWT.[65] This decrease in intensity suggests the residues in these specific loops are populating an alternative state at the higher temperatures, which could correspond to local protein unfolding and/or protein association.

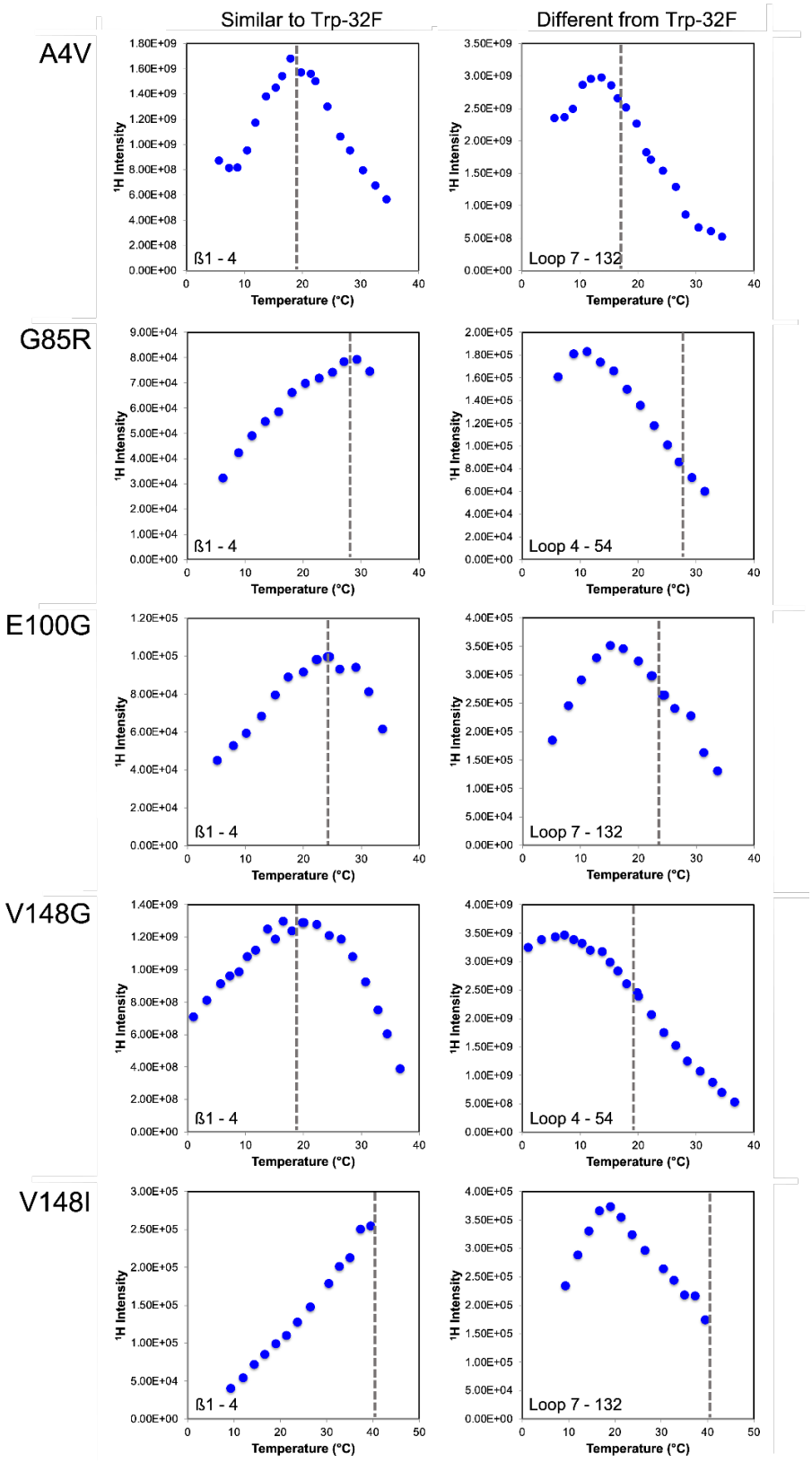



Figure 4.9: Intensity profile for residues that behave similar to (A) or differently from (B) Trp-32F for apoSH mutants. The vertical grey dashed line represented when protein unfolding begins based on DSC

endotherms reported in Vassal *et al*.[65] **(A)** Residue 4 for each mutant is shown as a representative peak that shows similarity in intensity profiles to the mutant's respective Trp-32F peak, which subsequently also have similar folded intensity slopes. Thus, these residues are considered to be in a folded/structured state similar to the one the Trp-32F is considered to report on. **(B)** However, the residues show different profiles than Trp-32F, mainly in the large loops (4 and 7), as well as some in loop 6, which explains the significantly different intensity slopes. The residues that show differences to their respective Trp-32F intensity profile as well as displaying behaviour reminiscent of protein unfolding at earlier temperatures than expected, suggests the loops are in a different folded state than Trp-32F reports on.

Furthermore, the loop residues in apoSH mutants that are considered significantly different from their Trp-32F slope show more patterns than the single profile observed for pWT (Figure 4.9). All mutants show two additional intensity profiles to the one observed for pWT, which includes a flat line with slope close to zero or a continuously decreasing intensity throughout the temperature ramp. Again, it is hard to decipher what each of these profiles could be reporting on, but it is worth noting that their additional intensity profiles are suggesting the loop residues are populating alternative states to different extents than observed for pWT.

These differences in peak behaviour throughout the dynamic loops are important to consider when trying to parse the misfolded and aggregation behaviour of apoSH mutants. For instance, previously Sekhar *et al*. reported on apoSH SOD1 mutants transiently accessing non-native dimers[54,55], and the residues involved in non-native dimerization are the same residues in loop 4, 6 and 7 that show less structuring relative to Trp-32F for some apoSH mutants. Even though pWT also shows residues in these areas deviate from Trp-32F slope value, these residues in addition to the loss of structure for destabilized mutants are considerable factors for changing folding dynamics and inevitably creating structural hotspots primed for aggregation.

4.4.2.2 Temperature corresponding to the maximum intensity may reflect local protein unfolding

Previously in Chapter 3, CSP values were useful for assessing extent of structural differences for apoSH mutants relative to pWT, but could not report on extent of structure effectively. Considering the loops for pWT show a different intensity profile compared to the rest of the folded structure, it suggests these loops respond to temperature differently, and highlight protein unfolding may begin at the loops. The temperature that corresponds to the maximum peak intensity for each residue will henceforth be referred to as the max temperature. Since both pWT and V148I do not enter their protein unfolding according to their DSC endotherms[65] and their Trp-32F peak profile, most of their residues only show signal increases (Figure 4.2B), and thus their max temperature corresponds to the temperature of the last spectrum acquired (Figure 4.10A).

However, as mentioned earlier, the large loops in pWT and V148I show a decrease in intensity at the higher temperatures in their temperature range relative to the destabilized mutants, but similar to the destabilized mutants, parts of the loops have max temperature that occurs as low as their second spectrum (~11 °C) (Figure 4.10B). While most of the structure has a high max temperature (average max temperature is 29.89 °C for pWT and 33.02 °C for V148I, Table 4.4), with over 69 and 53 residues for pWT and V148I showing a max temperature of 35.40 °C and 39.55 °C, respectively, loops 4, 6 and 7 show dramatically different values. Consistently, both pWT and V148I have either parts or most of their large loops that show signal intensity decreases at lower temperatures than expected (Table 4.4, Figure 4.10B), which corresponds to the loop dynamics changing differently through the temperature ramp. This dynamic change could be the

proteins accessing an alternative conformation that is beyond the detection of the instrument. Regardless, this suggests pWT and V148I loops are influenced by temperature earlier than the rest of the structure, which is interpreted here as local unfolding.

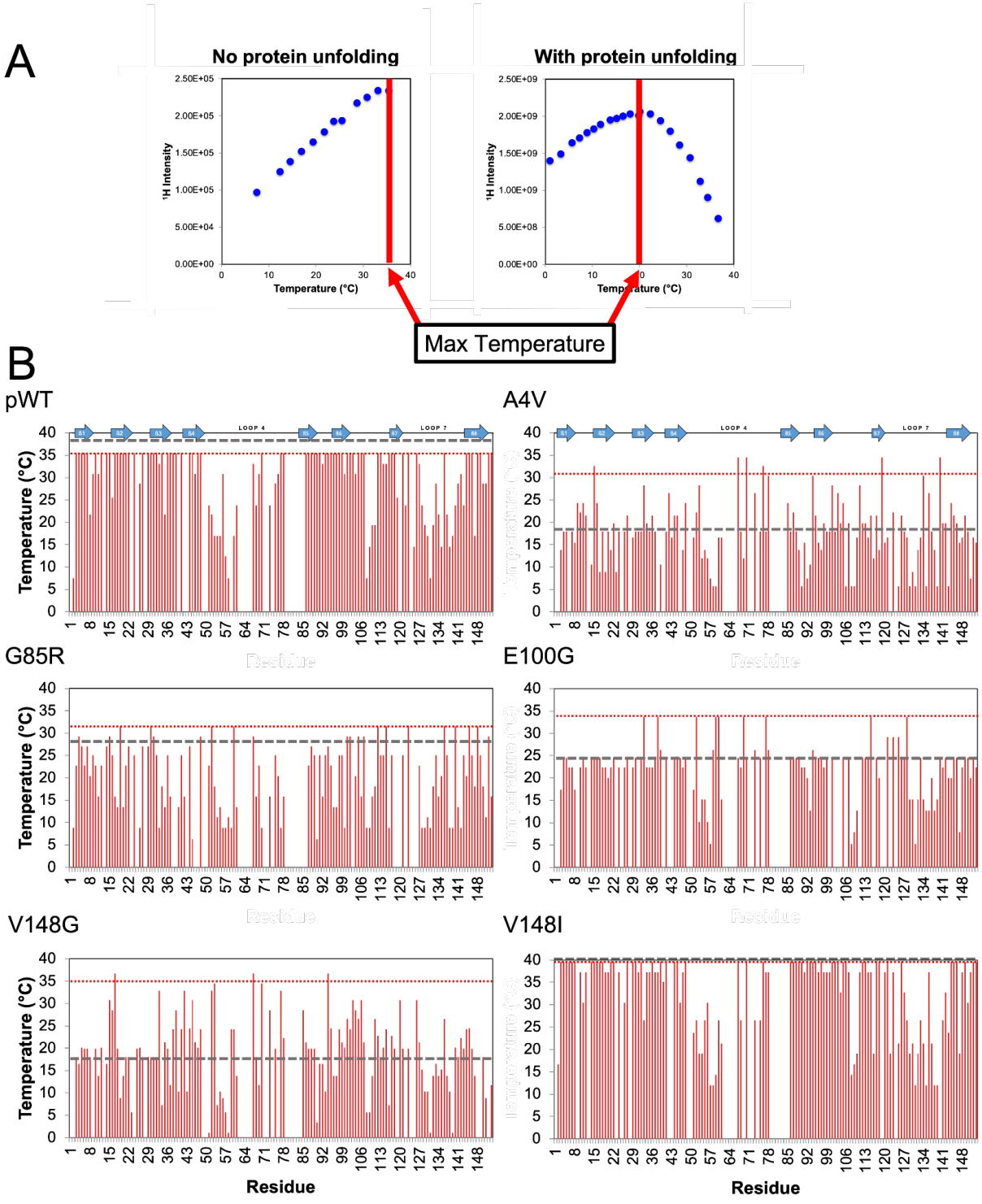



Figure 4.10: Maximum intensity temperature highlights local protein unfolding. **(A)** The temperature at which the highest peak intensity is observed is denoted as the max temperature (vertical red line). For pWT and V148I, spectra are only acquired in the folded temperature range, thus the maximum intensity for most of the folded residues corresponds to the highest temperature spectrum. However, for destabilized mutants where spectra are collected in both folded and unfolding temperature range, we extend the Trp-32F max temperature correlating to the start of protein unfolding to each residue, where the max temperature reports

on local unfolding. **(B)** Plots of max temperature for apoSH pWT and mutants (red bars). The horizontal red dotted line indicates the highest temperature spectrum collected, and the horizontal grey dashed line represents the start of protein unfolding based on the DSC endotherms reported in Vassal *et al*.[65]

Table 4.4: Average max temperature for apoSH variants.

| | pWT | A4V | E100G | G85R | V148G | V148I |
|---|---|---|---|---|---|---|
| **average** | 29.89 (125) | 17.92 (126) | 22.32 (113) | 20.29 (117) | 19.21 (123) | 33.02 (124) |
| **standard dev** | 7.90 | 7.07 | 6.09 | 7.72 | 8.16 | 8.79 |
| | | | | | | |
| **N-terminal B-sheet** | 34.22 (29) | 18.68 (29) | 23.32 (28) | 22.25 (28) | 21.53 (31) | 38.67 (29) |
| **C-terminal B sheet** | 33.39 (24) | 19.00 (27) | 23.90 (22) | 22.58 (20) | 21.21 (25) | 36.44 (24) |
| **N** | 7.45 (1) | - | - | 8.84 (1) | - | 16.69 (1) |
| **B1** | 33.12 (6) | 16.61 (5) | 21.41 (6) | 24.88 (6) | 18.73 (6) | 39.55 (6) |
| **loop I** | 33.11 (4) | 23.09 (4) | 23.07 (3) | 21.60 (4) | 17.93 (3) | 35.01 (3) |
| **B2** | 34.30 (9) | 17.18 (9) | 23.28 (9) | 22.82 (9) | 21.23 (9) | 39.05 (9) |
| **loop II** | 33.15 (3) | 16.56 (4) | 23.07 (3) | 20.31 (3) | 15.93 (4) | 35.76 (3) |
| **B3** | 33.40 (8) | 19.73 (8) | 24.58 (7) | 22.29 (8) | 19.17 (8) | 37.08 (8) |
| **loop III** | 35.40 (3) | 14.24 (2) | 28.15 (3) | 14.63 (2) | 18.70 (4) | 37.30 (4) |
| **B4** | 34.64 (6) | 20.30 (7) | 24.14 (6) | 20.68 (5) | 23.53 (8) | 37.66 (6) |
| **loop IV** | 24.46 (19) | 19.04 (20) | 22.71 (19) | 17.11 (20) | 19.39 (19) | 26.55 (19) |
| **B5** | 35.40 (4) | 19.25 (5) | 24.50 (4) | 24.95 (3) | 21.89 (5) | 39.55 (4) |
| **loop V** | 34.82 (4) | 9.76 (4) | 19.36 (4) | 18.03 (4) | 11.67 (4) | 38.99 (4) |
| **B6** | 34.83 (8) | 20.64 (8) | 24.45 (7) | 20.62 (8) | 22.62 (8) | 38.42 (8) |
| **loop VI** | 28.65 (13) | 17.41 (13) | 18.54 (8) | 21.44 (12) | 21.14 (13) | 33.23 (13) |
| **B7** | 32.96 (5) | 19.77 (6) | 26.05 (3) | 21.80 (3) | 20.98 (5) | 34.55 (5) |
| **loop VII** | 23.86 (21) | 16.25 (20) | 20.26 (20) | 18.02 (18) | 15.80 (19) | 24.46 (20) |
| **B8** | 30.55 (9) | 16.68 (9) | 20.92 (9) | 22.88 (9) | 18.44 (7) | 34.00 (9) |
| **C** | 35.40 (2) | 16.00 (2) | 23.43 (2) | 22.55 (2) | 11.74 (1) | 39.55 (2) |

The max temperatures are averaged across all assignments, the beta-strands on N-terminal and C-terminal face and secondary structural elements, and each average also has the number of residues averaged in brackets.

In contrast, since spectra are collected in the unfolding range for the destabilized mutants, their max temperatures can vary not only in the loops but throughout the structure. Considering E100G has a similar profile of normalized intensity slopes to pWT and V148I, it is unsurprising E100G's max temperature profile is also mostly similar to pWT and V148I (Figure 4.10B). More specifically, since E100G has spectra collected into its unfolding range (Table 4.2), the max temperature for most of the residues no longer corresponds to the highest temperature, as is the case with pWT and V148I where only the folded temperature range is assessed. However, 67 residues have a max temperature

of either 22.35 °C or 24.5 °C, which aligns with the start of unfolding expected for apoSH E100G, based on its DSC endotherm.[65]

For loops 4, 6 (site of mutation) and 7 in E100G, the average max temperature values are 22.71 °C, 18.54 °C and 20.26 °C (Table 4.4). Although the large loop max temperatures are not remarkably different than the other residues that have max temperatures of 22.35 °C and 24.5 °C, we must consider the folded range is much smaller for E100G, and that these loops still show smaller max temperature values relative to pWT and V148I. Specifically, loop 7 has the lowest max temperature compared to the rest of the mutants, and according to its temperature coefficient data, this loop is significantly less structured around the short helix that forms in loop 7 upon maturation (Chapter 3). This is particularly interesting considering apoSH E100G cannot transiently access the native-helix that occurs in loop 7.[55] Moreover, for a destabilized mutant, E100G shows subtle differences relative to pWT that may be critical in its protein misfolding, but overall, it is considered pWT-like when comparing relative max temperatures throughout the structure.

Contrary to pWT, V148I and E100G showing relatively high max temperatures throughout the protein apart from the larger loops, A4V and G85R show lower temperatures throughout the structure, instead of exclusively the loops (Figure 4.10B). A4V has the lowest average max temperature, 17.92 °C (Table 4.4), and instead of the large loops showing the lowest max temperature, loops 3 and 5 have the lowest values with 14.24 °C and 9.76 °C, respectively. In addition, by comparing the relative rankings of the average max temperature for secondary structural elements across the mutant set, A4V has a significantly different rank order than pWT (Table 4.4). This agrees with data

supporting apoSH A4V is not forming a native-like fold similar to pWT (see Chapters 3 and 7)[55], thus it is not expected to unfold like pWT.

Similarly, G85R also shows rank order that is different to that observed for pWT (Table 4.4). This mutant also shows significantly perturbed structure relative to pWT when assessing CSP and hydrogen bonded residues (Chapter 3), further supporting our conclusions on its less native-like fold. Additionally, both A4V and G85R loop 7 max temperature values are significantly lower than that of pWT, 16.25 °C and 18.02 °C, respectively (Table 4.4). Analogous to E100G, these apoSH mutants also cannot transiently access the native helix observed in loop 7[55], and show residues responsible for forming the native helix are considerably more dynamic (Chapter 3). To summarize, max temperature data for A4V and G85R suggest their protein fold is not native-like and support local areas of significant differences that may serve as hotspots for protein misfolding and aggregation.

Lastly, V148G shares similar characteristics to A4V and G85R, but shows interesting differences. ApoSH V148G has a low average max temperature similar to A4V and G85R, and a comparable ranking of averaged secondary structure max temperature to mostly A4V, and partially V148I. Additionally, loop 4 is not the area with the lowest max temperature, rather similar to A4V, loop 5 has the lowest max temperature of 11.67 °C followed by loop 7 (15.80 °C). Naser *et al*. report loop 5 buried in their apoSH V148G inclusion bodies, while loops 4 and 7 remain solvent accessible, and since both loops 5 and 7 show destabilized behaviour, it is only loop 5 that may play a role in protein aggregation. These differences in max temperature across the protein structure and between mutants reveal key structural changes throughout the temperature ramp.

Overall, the max temperature reveals critical areas that are unfolding sooner than other parts, but also highlight areas of the protein that are dynamic at physiological temperature. By assessing the protein changes throughout the temperature ramp, we can extrapolate those areas that show protein unfolding prior to physiological temperature, are also unfolded at physiological temperature. In addition to the large loops, apoSH pWT and mutants, except for G85R, show from loop 7 to ß8 has considerably lower max temperatures relative to the rest of the protein. The C-terminal sheet being more susceptible to unfolding agrees with temperature coefficient data showing less structured residues in this area (Chapter 3). Understandably, the C-terminal sheet becomes more structured upon maturation, and without these maturation events, the residues are not as structurally stabilized, and thus can serve as structural hotspots for protein aggregation. In summary, the max temperature data are important for understanding structural areas that are less structured and subsequently more dynamic at physiological temperatures, which will ultimately alter the native folding pathways.

### 4.4.3 Exploring linewidths: Less structured regions show peak broadening through the temperature ramp

With peak shape changes, we considered intensity earlier, but another important aspect of assessing peak shape is linewidths. Linewidths with peak intensity can help identify potential processes occurring in the sample. For example, based on their normalization, Sengupta *et al*. saw decreases in peak intensity coupled with minimal changes in their linewidth represented residues showing monomer-loss kinetics.[230] Additionally, they also observed peaks disappearing and appearing, both with increasing

linewidths, that corresponded to residues involved in oligomerization and peaks corresponding to the oligomer formation, respectively. In addition to assessing oligomerization, NMR lineshape analysis is used for other complex reactions such as allostery[237,238], cooperative binding[234], protein folding[83] and many more.

However, to understand the limits of line shape analysis, there are two considerations necessary to discuss. The first, with 2D NMR experiments, the peak shape changes in each dimension differently, thus simply concluding on fast and slow exchange is difficult. Fortunately, for the experiments discussed here, only the direct dimension is considered as the indirect dimension is considerably more crowded. A second more important limitation is that linewidths and peak intensity are only reliable if the peak does not overlap with another throughout the duration of the experiment. Unfortunately, due to the apoSH SOD1 being significantly more unstructured than its mature form, a lot of residues overlap in the chemical shift regions that correspond to random-coil shifts (discussed in Chapter 3). Thus, these limitations are important to consider when making conclusions.

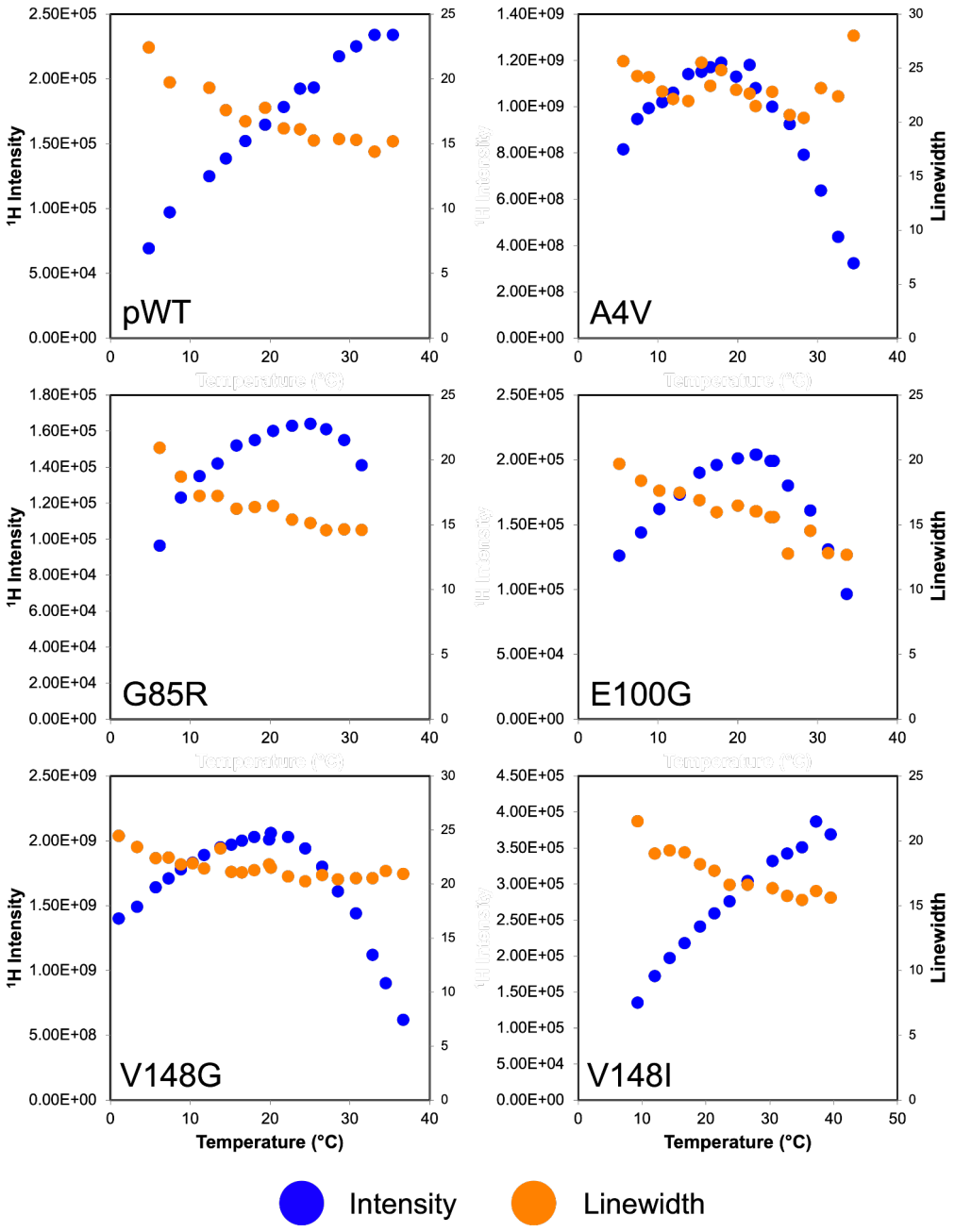



Figure 4.11: Trp-32F peak intensity and linewidth profiles throughout the temperature ramps for apoSH pWT and mutants. The Trp-32F intensity profiles show minimal scatter whereas the Trp-32F linewidths are more prone to scatter, thus overall trends must be considered for conclusions. Although both experience the impact of crowding, linewidths are more error prone due to peak overlapping, which may not be impacting the peak height. Overall, as the intensity increases or decreases throughout the heat ramp, the linewidth has minimal changes, suggesting the population corresponding to the Trp-32F peak is either increasing or decreasing, respectively.

Similar to the intensity data, the Trp-32F peak in pWT might offer another useful reference point for considering linewidth changes. The Trp-32F peak shows the linewidths decrease as the peak intensity increases with temperature (Figure 4.11). This suggests this peak is experiencing peak sharpening, either due to temperature impacts on the shims, or due to shifts in dynamics such that more nuclei populate the local environment corresponding to this chemical shift. However, the range of linewidth values is small relative to the intensity values and shows significant scatter throughout the temperature ramp. Thus, normalization is not possible as it was for the intensity profiles, but linewidths are still useful to assess for their general trends.

For the mutants, the Trp-32F linewidths highlight the difficulty of assessing lineshape. Similar to pWT, A4V, G85R, E100G and V148G show decreasing linewidths within their folded temperature range, but with varying degrees of considerable scatter (Figure 4.11). With Trp-32F being an isolated peak, the extent of scatter observed across pWT and the mutants emphasizes how error prone the linewidth fits can be, even when the intensity profiles show minimal scatter. Moreover, in the unfolding temperature range for each mutant, although the intensity decreases, the linewidths show a slight decrease, much less than the folded temperature range. This suggests that upon entering the unfolding, similar to Sengupta *et al*. the intensity decreases coupled with the linewidth not changing considerably aligns with monomer-loss.[230] Collectively, Trp-32F cannot be used as a normalization reference for linewidths as it was with the intensity profiles; but the Trp-32F linewidths further support the intensity decrease is a result of folded monomer loss due to thermal unfolding.

Similarly, Trp-32U peaks across the mutants also support the unfolding conclusions. Since pWT and V148I have no unfolding spectra collected, they are excluded from this analysis. For all the destabilized mutants, their Trp-32U intensity profiles show increases in intensity only upon entering their unfolding temperature range. Interestingly, the linewidths, similar to Trp-32F, show minimal changes with temperature, which supports this peak corresponding to gain of unfolded species (Figure 4.12). In Chapter 6, we further discuss Trp-32U can also report on additional species. Overall, the Trp-32F and Trp-32U linewidths further support these peaks appropriately representing folded and unfolded species, respectively, based on their intensity profiles.

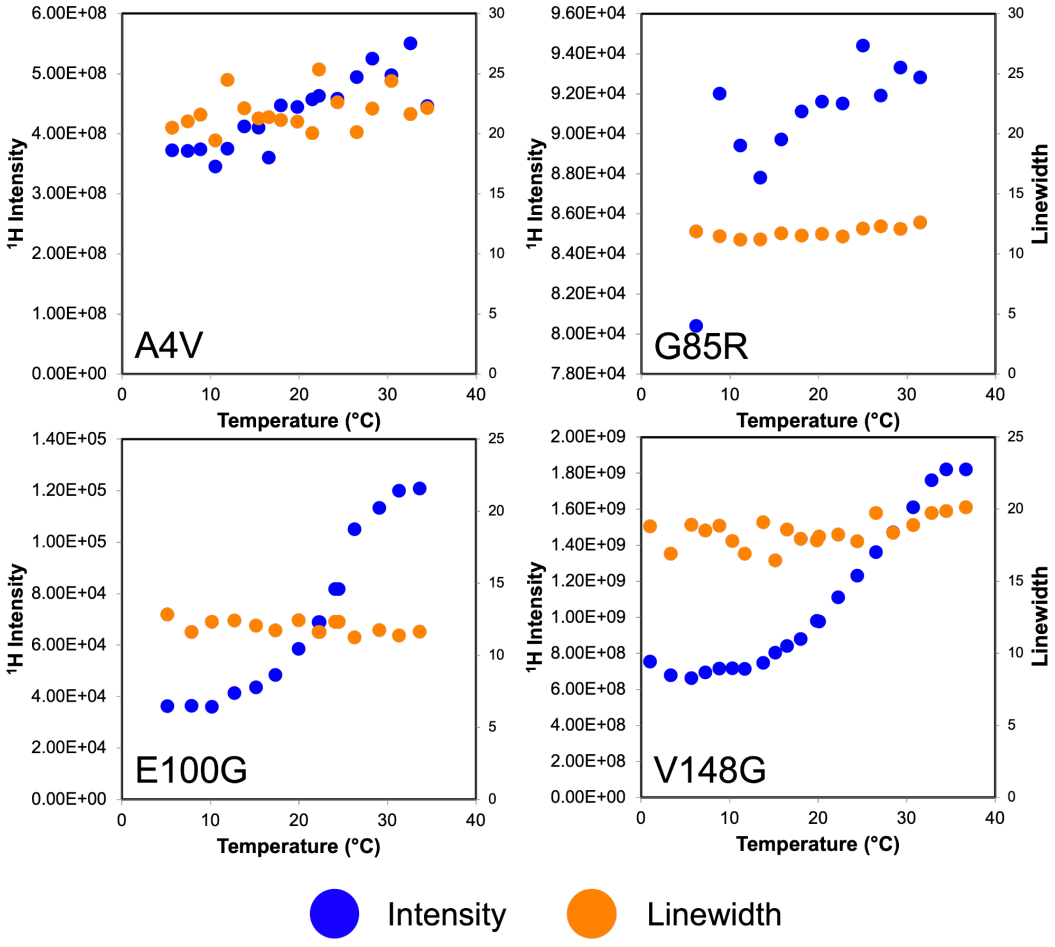



Figure 4.12: Trp-32U peak intensity and linewidth profiles throughout the temperature ramps for apoSH pWT and mutants. The Trp-32U intensity profiles show minimal scatter whereas the Trp-32U linewidths are

more prone to scatter, thus overall trends must be considered for conclusions. Although both experience the impact of crowding, linewidths are more error prone due to peak overlapping, which may not be impacting the peak height. Overall, as the Trp-32U peak intensity increases as the protein unfolding begins, and linewidths change minimally suggesting a gain of population corresponding to the Trp-32U peak.

Similar to the intensity profiles, the linewidths show variable changes throughout the structure (Figure 4.13). In addition to the high scatter in the linewidth profiles for isolated peaks, the linewidths are sensitive to overlapping peaks, which unfortunately is a majority of the apoSH peaks. For the sake of simplicity, trends in linewidths are considered based on clusters of residues showing similar trends, and visually assessing outliers. Linewidth trends are categorized as either increasing, decreasing or no discernible change.

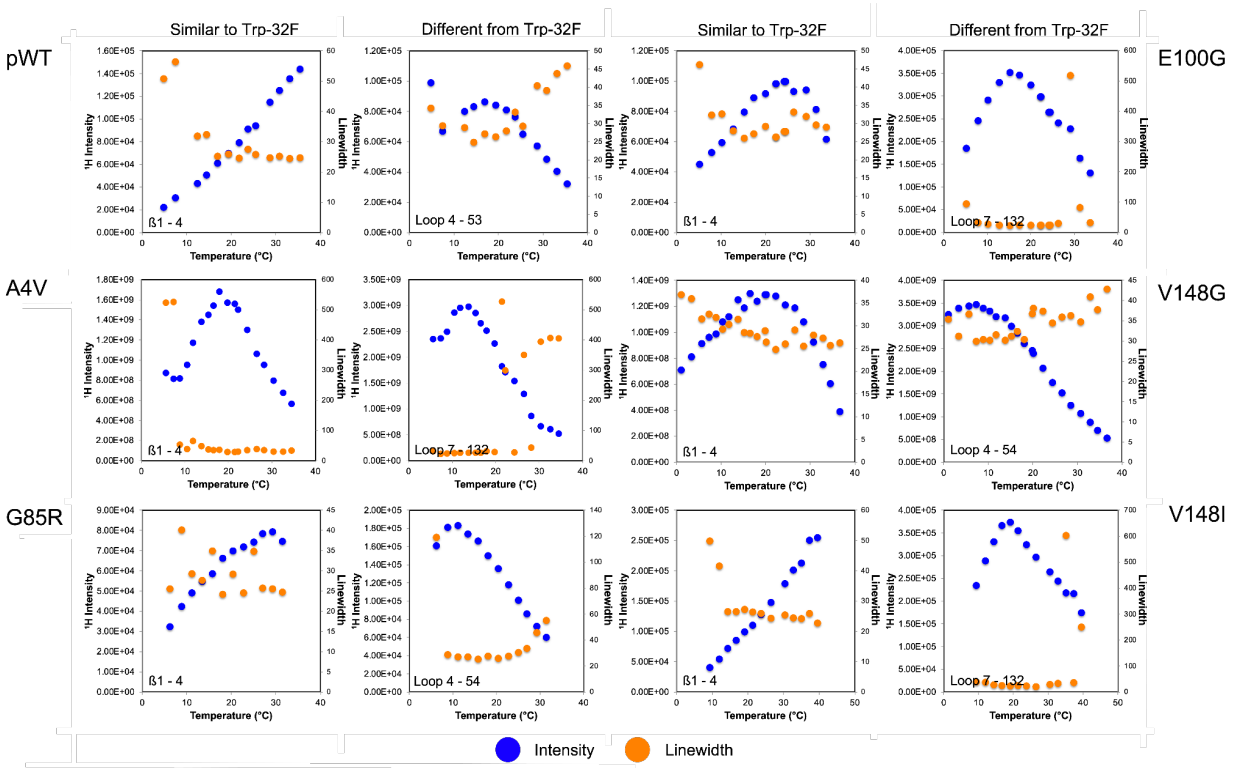



Figure 4.13: Intensity and linewidth profile for residues that behaved similar to (A) or differently from (B) Trp-32F for apoSH pWT and mutants. Residue 4 for each mutant is shown as a representative peak that shows similarity in intensity and linewidth profiles to the mutant's respective Trp-32F peak, which subsequently also have similar folded intensity slopes. Thus, these residues are in a folded/structured state similar to the one the Trp-32F is considered to report on. Upon entering the unfolding temperatures, the intensity decreases, and the linewidth does not change suggesting a decrease in the population. However, the residues show different profiles than Trp-32F, mainly in the large loops (4 and 7), as well as some in loop 6, which explains the significantly different intensity slopes. Upon intensity decrease, the linewidths increase suggesting the residues have different dynamics than those observed for the folded/structured

parts of the protein. The residues that show differences to their respective Trp-32F intensity and linewidth profile as well as displaying behaviour reminiscent of protein unfolding at earlier temperatures than expected, suggests the loops are in a different folded state than Trp-32F reports on.

For pWT, Trp-32F linewidth behaviour is comparable for the same residues that show comparable intensity profile to Trp-32F (Figure 4.11 and 4.13). Most of the residues show similar linewidth behaviour to the Trp-32F, which is a slight decrease in linewidth with temperature meaning there is peak sharpening. Since the residues corresponding to folded structures have similar intensity and linewidth profiles, it suggests the folded structures correspond to the temperature ramp comparably. However, for the loops, the linewidths show minimal changes during the intensity increase, but then the linewidths increase as the intensity decreases (Figure 4.13). The initial peak sharpening is similar to that observed for the more structured residues, but the peak broadening that occurs at the higher temperatures is unlike the structured residues. The signal decrease with peak broadening suggests the loops are more dynamic and access additional conformations at the higher temperatures. Overall, the peak shapes for the structured residues behave comparably to the Trp-32F peak whereas the loops deviate significantly.

Additionally, the apoSH mutants show comparable behaviour but several residues show extensive scatter. Unsurprisingly, the mutants that show similar behaviour to their Trp-32F peak, also show comparable behaviour in their linewidths (Figure 4.11 and 4.13). V148I residues show similar peak shapes to pWT in both linewidths and intensity. However, unlike V148I, the destabilized mutants have spectra collected in their unfolding temperature range, but the scatter in the linewidth data increases in the unfolding spectra, which limits the analysis. In general, the areas that show significant perturbation in the

intensity profiles for apoSH mutants relative to their respective Trp-32F, also show similar line broadening observed for the large loops in pWT.

Furthermore, apoSH mutants show more areas of intensity-based perturbation than observed for pWT, and the additional areas of perturbation also show similar line broadening. Specifically, G85R and pWT show the same residues in loops 4 and 7 being perturbed relative to their respective Trp-32F, but G85R also shows additional residues in loop 4 with perturbed intensity profiles which also show enhanced peak broadening (Figure 4.11 and 4.13). Additionally, V148G is the only mutant who does not show loop 7 being heavily perturbed relative to the Trp-32F intensity profile, and subsequently, only residues in loop 4 show peak broadening whereas loop 7 behaves more comparably to the other folded structures. Unfortunately, due to the extensive scatter, concluding on small, local dynamic differences based on linewidth behaviour is difficult. Thus, due to the limitations, areas with larger differences in intensity profiles relative to the Trp-32F peak show peak broadening, which aligns with increased dynamics with temperature that were not observed for the residues that are comparable to their Trp-32F peak.

## 4.5 Conclusion

From a single $^{1}H$-$^{15}N$-HSQC experiment, an ample amount of information is gained, which only increases with the addition of a changing condition. The true limitation comes from data interpretation; peak shape is often the less considered analysis due to the difficulties in correlating the peak shape changes to dynamics without additional data or complicated experiments. Here, we explore the information intensity profiles can provide on local structure and dynamics for apoSH pWT, and additionally how introducing

point mutations impact the structure. ApoSH SOD1 spectra show the presence of two tryptophan sidechain peaks, Trp-32F and Trp-32U. The intensity profiles for these peaks correlated strongly with the protein unfolding observed by DSC, suggesting these peaks correspond to the folded and unfolded conformers of SOD1, thus allowing the extent of structure throughout the protein to be assessed.

For pWT and mutants, the normalized intensity profiles and linewidths suggest the areas that show extensive perturbation relative to the Trp-32F peak, and show enhanced line broadening, are considered to be less structured. However, residues that do show similar peak shape to Trp-32F, both in terms of normalized intensity and linewidths throughout the temperature ramp, are reporting on structured regions of the protein. Additionally, the temperature that coincides with the maximum intensity, referred to as max temperature, can highlight local areas of protein unfolding throughout the temperature ramp. The lowest max temperature occurs at areas that are considered less structured, specifically the large loops, whereas the residues that form beta-strands in the mature enzyme have higher max temperatures. In conclusion, considering the different parameters gained from NMR experiments, like temperature coefficients and intensity profiles, although illuminating, is heavily dependent on the interpretation. By assessing the data with a control peak, such as two peaks that represent the folded and unfolded conformer, allow for a more robust analysis of peak shape changes due to a varying condition.

# Chapter 5: Insights into protein structure and dynamics from amide nitrogen temperature coefficients

Author Contributions:

Kyle Trainor: conceptualization, methodology, resources, validation, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft; Harmeen Kaur Deol: methodology, resources, validation, formal analysis, investigation, visualization, data curation, writing-original draft; Colleen Doyle: methodology, resources, validation, formal analysis; Elizabeth M. Meiering: conceptualization, supervision, project administration, funding acquisition.

## 5.1 Context

The NMR chemical shifts of proteins are extremely sensitive to the local magnetic environments of nuclei. The factors that determine chemical shifts may vary with temperature to different degrees; accordingly, temperature dependences offer insights into protein structure and dynamics beyond those revealed by isothermal chemical shifts. These temperature dependences are often well-approximated by a linear model, the slope of which is the ‘temperature coefficient’. Amide proton temperature coefficients may probe hydrogen bonding and temperature-dependent loss of structure; in contrast, no similarly straightforward interpretations of amide nitrogen temperature coefficients have been proposed - a glaring omission, since these data are often collected in parallel. Here, we facilitate interpretation by accounting for the effects of amino acid sequence, which

strongly impacts both amide nitrogen chemical shifts and their temperature dependencies. We present Calc$^N$, a sequence-based method of calculating the amide nitrogen temperature coefficients expected for random coil. Subtraction of these values from experimentally determined temperature coefficients yields a measure of difference from random coil, i.e., structure. We validate Calc$^N$ using experimentally determined temperature coefficients of an RNA-binding domain (dsRBD-1); though free dsRBD-1 is intrinsically disordered; our method proves sensitive to transient sampling of structural elements found in the RNA-bound form. In natively folded immunoglobulin-like domains, we observe substantial deviation from random coil throughout and intriguing differences between minor variants, e.g., those related by point mutation, multimerization, metal binding, and disulfide bond formation. Calc$^N$, accessible via the Shift-T server (http://meieringlab.uwaterloo.ca/shiftt), may unlock great value in data that are commonly acquired, yet rarely analyzed.

## 5.2 Introduction

The temperature dependences of both amide proton and amide nitrogen chemical shifts are frequently well approximated by a linear model.[202,203] Linear temperature coefficients (i.e., slopes determined by fitting chemical shift vs. temperature data to a linear model) of amide protons have proven to be useful, though imperfect, probes of hydrogen bonds[199,219], temperature-dependent loss of structure[66,240–242], and have also been used to distinguish between ordered and (intrinsically) disordered protein regions.[242–245] Although much is known about the

determinants of amide nitrogen chemical shifts[246–249], comparably straightforward interpretations of amide nitrogen temperature coefficients have yet to emerge.[66,202,250]

Here, after a brief introduction to factors known to influence amide nitrogen chemical shifts, we describe the Calc$^N$ (pronounced 'calculation') method of determining random coil temperature coefficients using published model peptide data[217]. Differences between these calculated values and experimentally measured amide nitrogen temperature coefficients are extremely sensitive to deviations from random coil and may therefore serve as a measure of structure. Crucially, referencing to random coil confers independence from amino acid sequence, enabling direct comparisons between the amide nitrogen temperature coefficients of different protein residues. We apply Calc$^N$ to the analysis of transient structure in an 'intrinsically disordered' domain and structural perturbations of natively folded immunoglobulin-like proteins.

### 5.2.1 Amide Nitrogen Chemical Shifts

The proteinogenic amino acids share (at a minimum) the same succession of 'backbone' atoms, but within a polypeptide the chemical shifts of the atoms in a particular amino acid may be strongly influenced by their local environment, including the characteristics of the attached and sequentially adjacent side chains.[247,249,251] Amide nitrogen chemical shifts are known to be influenced by diverse factors, some of which contribute to amino acid sequence dependence (Table 5.1).

Correlations between amide nitrogen chemical shifts and those of the $C_\alpha$/$C_\beta$ carbons from the preceding (N-1$^{th}$) residue have been reported.[247,249] As the $C_\alpha$

chemical shift is believed to be strongly influenced by the electron withdrawing or donating nature of the side chain to which it is directly bonded, it has been proposed that amide nitrogen chemical shifts may be subject to similar covalently mediated effects (i.e., attributable to the N-1$^{th}$ side chain).[247] Furthermore, amide nitrogen chemical shifts are also sensitive to $\psi_{N-1}$ and $\varphi_N$ backbone dihedral angles[248], which change the positions of the N-1$^{th}$ and N$^{th}$ side chains, respectively, relative to the N$^{th}$ amide nitrogen (Figure 5.1); thus, through-space effects may also contribute to sequence dependence.

Side chain dihedral angles may also play a role; in particular, the $\chi_1$ angle is often important[246,248], as some values position side chain $\gamma$ atoms very near the amide nitrogen.

Table 5.1: Factors influencing nitrogen chemical shifts

| Factor | Description |
|---|---|
| Inductive Effects | Covalently mediated, sequence-dependent electron withdrawal/donation |
| Backbone Dihedrals | Particularly $\varphi_N$ and $\psi_{N-1}$ |
| Side Chain Dihedrals | Particularly $\chi_{1,N}$ |
| Conformational Averaging | Sufficiently fast exchange between states results in observation of population-weighted average chemical shifts |
| Primary Hydrogen Bonding | Deshielding by hydrogen bonds in which the N$^{th}$ amide nitrogen acts as the donor |
| Secondary Hydrogen Bonding | Deshielding by hydrogen bonds in which the N-1$^{th}$ carbonyl oxygen acts as the acceptor |

Many side chains in proteins exhibit preferences for one of three non-eclipsed $\chi_1$ angles ($\simeq -60/60/180°$). However, more than one rotamer may be sampled; thus, experimentally measured chemical shifts may reflect conformational averaging of different rotamers.

Hydrogen bonding also deshields amide nitrogen atoms. Hydrogen bonds in which the $N^{th}$ amide nitrogen acts as the donor (primary) or, to a lesser degree, in which the N-$1^{th}$ carbonyl oxygen acts as the acceptor (secondary) may each have an impact.[246]

### 5.2.2 Random Coil Chemical Shifts and Temperature Coefficients

In general, the dependence of chemical shifts on amino acid sequence complicates comparisons between the atoms of different residues. Random coil chemical shifts (measured from peptide models expected to exhibit random coil-like behaviour in solution[217,251–254]) similarly depend on amino acid sequence; thus, the difference between a measured chemical shift ($\delta_{Obs}$) and that expected for a random coil having the same amino acid sequence ($\delta_{RC}$), known as the secondary chemical shift ($\delta_{Sec}$; Eq. 1), is approximately independent of sequence.

$$\delta_{Sec} = \delta_{Obs} - \delta_{RC} \qquad (1)$$

Structural considerations dominate secondary chemical shifts; they form the basis for chemical shift-based prediction of protein structure as well as the inverse problem of predicting chemical shifts from known structures. Here, by analogy with secondary chemical shifts, we

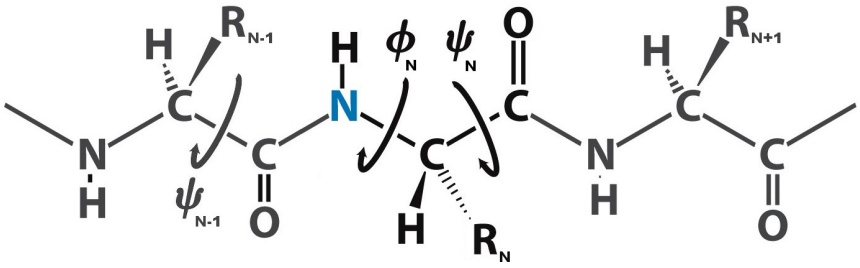



Figure 5.1: Protein backbone dihedral angles $\psi_{N-1}$ and $\varphi_N$ affect the positions of the N-1$^{th}$ and N$^{th}$ side chains relative to the N$^{th}$ amide nitrogen (blue).

define secondary temperature coefficients as the difference between observed and sequence-dependent random coil values (Eq. 2). Like secondary chemical shifts, secondary temperature coefficients are approximately independent of amino acid sequence.

$$\left(\frac{\Delta\delta}{\Delta T}\right)_{Sec} = \left(\frac{\Delta\delta}{\Delta T}\right)_{Obs} - \left(\frac{\Delta\delta}{\Delta T}\right)_{RC} \quad (2)$$

## 5.3 Results and Discussion

In order to translate experimental observations into the sequence-independent realm of secondary temperature coefficients, we require a method of determining the amide nitrogen temperature coefficients expected for random coil. It has been

demonstrated using peptide models that not only amide nitrogen random coil chemical shifts themselves, but also their rates of change with temperature strongly depend on local amino acid sequence (Table 5.2).[217] We note that several existing methods for calculating random coil amide nitrogen chemical shifts either neglect temperature dependences altogether or implicitly assume that they are sequence-independent.[254–256]

The number of random coil peptides that must be characterized in order to determine how interactions between sequentially adjacent side chains contribute to observed chemical shifts (and temperature coefficients) increases exponentially with the number of amino acids considered. Fortunately, consistent with their proximities to the amide nitrogen, it has been shown that the identities of the attached ($N^{th}$) and previous ($N\text{-}1^{th}$) side chains have the strongest influences[254,257], which may justify the simplification allowed by consideration of just two amino acids at a time.[217,252,257] The common practice of treating the contributions of each side chain as approximately independent also offers greatly simplifies the problem.[217,252,254,257]

From the data in Table 5.2, there are two subtly different ways to model the contributions of the $N^{th}$ and $N\text{-}1^{th}$ side chains to amide nitrogen temperature coefficients. We can take a value from the X3 column (i.e., the effect of the $N^{th}$ side chain) as our starting point; X3 is preceded by a glycine in all of the experimentally characterized peptide sequences, but we can model the effect of the $N\text{-}1^{th}$ side chain as the difference between two rows from the A4 column (Eq. 3, where $AA_{\#}$ is the amino acid type of the specified residue and from Table 5.2). Though more direct, this approach produces the same results as calculating sequence and temperature corrected amide

nitrogen chemical shifts at more than one temperature, then determining the slope of the line joining them.[217]

Table 5.2: Random coil amide $^{15}N$ temperature coefficients.

| | X3 (ppb/K)[1,2] | A4 (ppb/K)[2] | GX (ppb/K)[1,3] | XA (ppb/K)[1,4] |
|---|---|---|---|---|
| **Ala (A)** | -9.0 | -15.6 | -6.8 | -17.8 |
| **Cys (C)** | -12.2 | -14 | -10.0 | -16.2 |
| **Asp (D)** | -6.8 | -14.7 | -4.6 | -16.9 |
| **Glu (E)** | -8.1 | -16.2 | -5.9 | -18.4 |
| **Phe (F)** | -10.5 | -15.3 | -8.3 | -17.5 |
| **Gly (G)** | -7.6 | -6.8 | -5.4 | -9.0 |
| **His (H)** | -6.7 | -12.1 | -4.5 | -14.3 |
| **Ile (I)** | -17.1 | -26.5 | -14.9 | -28.7 |
| **Lys (K)** | -11.5 | -17.2 | -9.3 | -19.4 |
| **Leu (L)** | -5.1 | -17.7 | -2.9 | -19.9 |
| **Met (M)** | -8.3 | -17.3 | -6.1 | -19.5 |
| **Asn (N)** | -8.9 | -11.7 | -6.7 | -13.9 |
| **Pro (P)** | | -22.2 | | |
| **Gln (Q)** | -7.2 | -18.3 | -5 | -20.5 |
| **Arg (R)** | -12.2 | -16.5 | -10 | -18.7 |
| **Ser (S)** | -10.1 | -10.9 | -7.9 | -13.1 |
| **Thr (T)** | -15.4 | -17.3 | -13.2 | -19.5 |
| **Val (V)** | -19.0 | -25.5 | -16.8 | -27.7 |
| **Trp (W)** | -9.6 | -21.9 | -7.4 | -24.1 |
| **Tyr (Y)** | -12.4 | -22.0 | -10.2 | -24.2 |

[1] X represents any amino acid (determined by row)
[2] Measured from GGXAGG peptides[217]
[3] Calculated using Eq. 4 for comparison with X3 column
[4] Calculated using Eq. 3 for comparison with A4 column

$$\frac{\Delta\delta_{\mathrm{RC}}}{\Delta T} = \frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[AA_N] + \left(\frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[AA_{N-1}] - \frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[Gly]\right) \quad (3)$$

Alternatively, we can begin with a value from the A4 column (i.e., the effect of the N-1$^{th}$ side chain), then model the effect of the N$^{th}$ side chain as the difference between two rows from the X3 column (Eq. 4).

$$\frac{\Delta\delta_{\mathrm{RC}}}{\Delta T} = \frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[AA_{N-1}] + \left(\frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[AA_N] - \frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[Ala]\right) \quad (4)$$

Ideally, use of either Eq. 3 or Eq. 4 would give the same answer (for a given amino acid sequence), but we find that there are small discrepancies. Temperature coefficients for the A (alanine) in XA (where X is any amino acid except proline) can be calculated using Eq. 3 and directly compared to the A4 column of Table 5.2. Similarly, temperature coefficients for X in GX can be calculated using Eq. 4 for comparison to column X3 of Table 5.2. The temperature coefficients calculated using Eq. 3 are systematically more negative than the measured values (A4), while the temperature coefficients calculated using Eq. 4 are systematically less negative than the measured (X3) values, revealing modest errors (approximately 2 ppb/K) introduced by our imperfect assumptions.

The errors described above can be ascribed to the difference (-9.0 ppb/K vs. -6.8 ppb/K) in the temperature coefficients of A3 in the GGAAGG peptide and A4 in the GGGAGG peptide ($AA_{N-1}$=G and $AA_N$=A in both cases). We hypothesize that the less negative temperature coefficient in the latter case may be attributable to the presence of a glycine in the N+1$^{th}$ position, and that the temperature coefficients of residues

with non-glycine amino acids in the N+1[th] position may be better approximated by Eq. 3. Thus, we further attempt to limit discrepancies between calculated and experimental values by applying Eq. 4 when the N+1[th] amino acid is a glycine, and Eq. 3 otherwise, giving the composite Eq. 5.

$$\frac{\Delta\delta_{\mathrm{RC}}}{\Delta T} = \begin{cases} \frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[AA_{N-1}] + \left(\frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[AA_N] - \frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[Ala]\right), \text{if } AA_{N+1} = Gly \\ \frac{\Delta\delta_{\mathrm{X3}}}{\Delta T}[AA_N] + \left(\frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[AA_{N-1}] - \frac{\Delta\delta_{\mathrm{A4}}}{\Delta T}[Gly]\right), \text{otherwise} \end{cases} \quad (5)$$

We validate our method by calculating secondary amide nitrogen temperature coefficients for the intrinsically disordered dsRBD-1 domain from *Arabidopsis thaliana* DCL1. DCL1 is a ribonuclease, and the dsRBD-1 domain folds upon binding RNA; though generally disordered in the absence of RNA, chemical shift changes upon titration with urea suggest that the domain may transiently sample folded conformations that are destabilized by the denaturant[245]. In Figure 5.2 we compare these chemical shift changes with secondary temperature coefficients calculated using Eqs. 2 and 5. Because experimentally observed random coil $^{15}$N temperature coefficients vary over a ~20 ppb/K range (Table 5.2), we call extra attention to secondary temperature coefficients that fall outside of the ±20 ppb/K range. Consistent with random coil, many of the secondary temperature coefficients are near zero; however, in the H2 helix, where chemical shift changes upon titration with urea are most prominent, we observe large positive secondary temperature coefficients consistent with transient sampling of structure in the absence of denaturant.

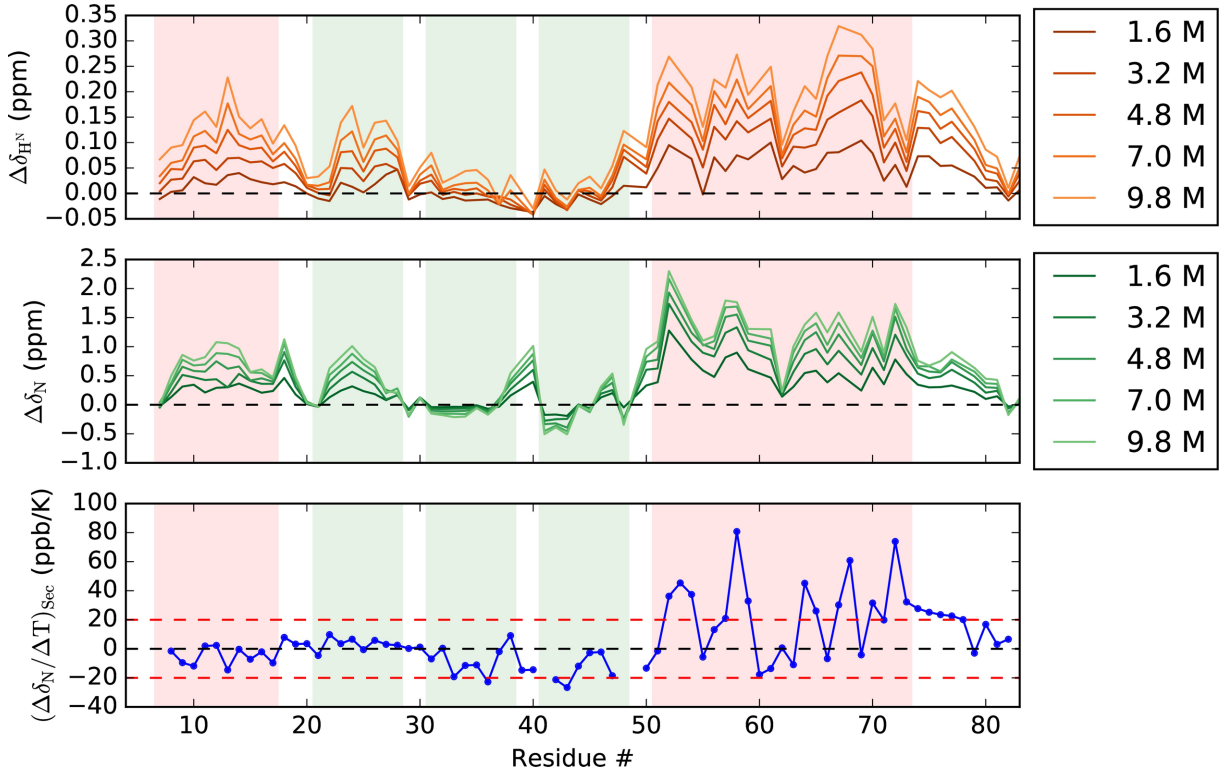



Figure 5.2: Chemical shift perturbations in the intrinsically disordered dsRBD-1 domain from DCL1 due to changes in urea concentration or temperature.[245] Structure induced in RNA-bound conformations is indicated: *α*-helix (light red); *β*-strand (light green). Amide proton (top) and amide nitrogen (middle) chemical shift changes due to titration with urea. Secondary amide nitrogen temperature coefficients (bottom) are random coil-like (near zero) over most of the domain, but large positive values are apparent in the H2 helix (residues 51-73). Dashed red lines demarcate a ±20 ppb/K range; secondary temperature coefficients outside of this range are particularly unlike random coil. Contiguous data points are joined by solid lines.

In contrast to intrinsically disordered dsRBD-1, we observe substantial deviation from random coil in secondary temperature coefficients throughout the immunoglobulin-like tenth human fibronectin type III domain D80E ($^{10}Fn3_{D80E}$; a mutant similar to wild type $^{10}Fn3$ in all respects except affinity for integrins[258,259]), including within many of the loops (Figure 5.3). This confirms our intuition that in folded proteins, even regions lacking secondary structure may be readily distinguishable from random coil.

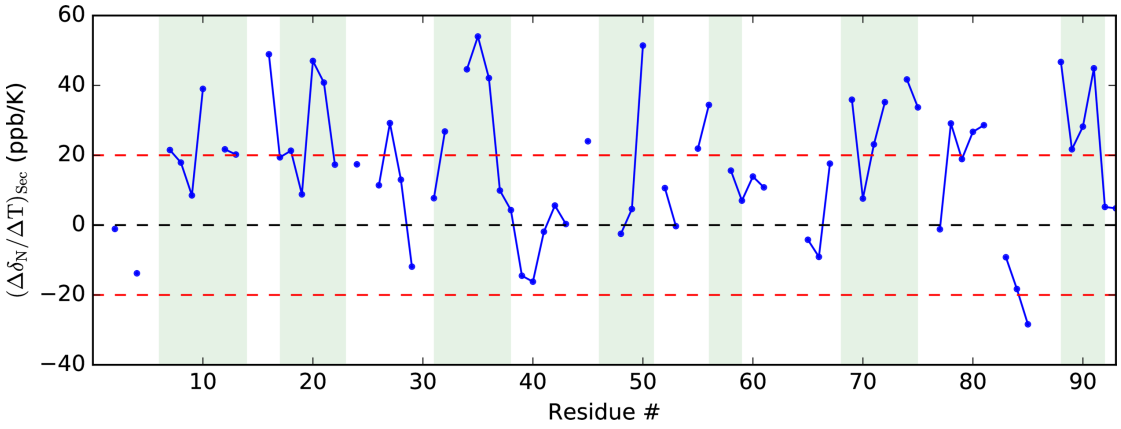



Figure 5.3: Secondary amide nitrogen temperature coefficients of $^{10}Fn3_{D80E}$. Values inconsistent with random coil are apparent throughout the protein (including loop regions), but especially in the $\beta$-strands (light blue background). Dashed red lines demarcate a ±20 ppb/K range; secondary temperature coefficients outside of this range are particularly unlike random coil. Contiguous data points are joined by solid lines.

Empirically, even though negative secondary temperature coefficients also suggest deviations from random coil, large positive (>20 ppb/K) secondary temperature coefficients appear to be more common in regions of secondary structure (Figures 5.2 and 5.3). In Figure 5.4, we observe that this trend continues for both dimeric human Cu, Zn superoxide dismutase 1 (holoSS SOD1) and its disulfide-free, unmetallated, and monomeric counterpart (apoSH SOD1). Mutations in SOD1 are linked to the fatal neurodegenerative disease familial amyotrophic lateral sclerosis (fALS). Here, we apply Calc$^N$ to detect intriguing differences between pseudo-wild type SOD1 (pWT; cysteine 6 and cysteine 111 are replaced by alanine and serine, respectively) and mutants thereof.

It has been hypothesized that mutant SOD1 may cause disease through the formation of toxic protein aggregates, and that immature states (such as reduced apo) play a key role in this process.[259] The relationships between stability, aggregation, and disease are not straightforward, but amide proton temperature coefficients suggest that

the structure of the apoSH state is more easily perturbed than that of the holoSS state (see Chapter 3), and calorimetric data show that the global stability of apoSH SOD1 is highly sensitive to mutations, e.g., E100G, that are well-tolerated by holoSS SOD1.[259,260]

The secondary temperature coefficients between holoSS SOD1 and apoSH SOD1 show different extents of structure between the two SOD1 states (Table 5.3). The most obvious difference between the two states is in the N-terminal and C-terminal ß-sheets. Similar to the amide proton temperature coefficients (see Chapter 3), the N-terminal face is comparable in both states, but the C-terminal face is more random coil-like in the apoSH SOD1 because of its missing maturation events. Additionally, the large loops, 4 and 7, are also involved in dimerization, metal-binding and formation of the disulfide bond in holoSS SOD1, and therefore their secondary temperature coefficient values are less random coil-like than in apoSH SOD1. Therefore, secondary temperature coefficients are reliable measures of extent of structure between different states of the same protein.

Table 5.3: Average secondary temperature coefficients for holoSS and apoSH variants.

| | HoloSS | | | ApoSH | | |
|---|---|---|---|---|---|---|
| | **hpWT** | **hE100G** | **hV148I** | **pWT** | **E100G** | **V148I** |
| **average** | -6.58 (95) | -6.72 (93) | -6.19 (91) | -9.25 (125) | -10.36 (107) | -9.22 (121) |
| **standard dev** | 10.37 | 10.56 | 10.58 | 13.80 | 16.48 | 13.42 |
| | | | | | | |
| **N-terminal B-sheet** | 14.21 (26) | 14.44 (21) | 13.34 (22) | 20.22 (31) | 17.60 (29) | 19.19 (31) |
| **C-terminal B sheet** | 16.18 (26) | 16.54 (27) | 16.48 (22) | 6.39 (24) | 5.83 (20) | 7.78 (23) |
| **N** | - | -1.41 (1) | - | 14.99 (1) | - | 12.52 (1) |
| **B1** | 18.03 (6) | 17.11 (4) | 9.75 (3) | 18.67 (6) | 20.84 (6) | 18.09 (6) |
| **loop I** | 21.07 (1) | - | - | 4.22 (4) | 5.56 (3) | 4.80 (3) |
| **B2** | 10.81 (6) | 15.13 (7) | 16.94 (6) | 23.07 (9) | 23.99 (9) | 21.25 (9) |
| **loop II** | -1.17 (1) | -1.14 (1) | -1.97 (1) | 5.25 (3) | 7.07 (3) | 6.11 (3) |
| **B3** | 19.13 (7) | 15.55 (6) | 17.98 (6) | 22.76 (8) | 18.48 (7) | 21.19 (8) |
| **loop III** | 2.67 (2) | 2.11 (2) | 1.40 (2) | 6.14 (3) | 0.00 (2) | -6.88 (3) |
| **B4** | 12.04 (7) | 13.00 (8) | 9.47 (6) | -9.89 (6) | -12.60 (6) | -13.30 (6) |
| **loop IV** | 8.55 (17) | 6.63 (18) | 10.32 (21) | 2.33 (19) | 3.92 (17) | 5.80 (19) |
| **B5** | 15.00 (7) | 13.61 (6) | 13.93 (6) | 19.70 (4) | 20.95 (4) | 20.99 (4) |
| **loop V** | -1.76 (3) | 7.32 (3) | 6.08 (4) | 5.91 (4) | 3.23 (4) | 3.96 (4) |
| **B6** | 8.87 (7) | 9.97 (4) | 8.68 (7) | 16.36 (8) | 7.09 (7) | 16.23 (8) |
| **loop VI** | 17.51 (7) | 18.20 (7) | 18.71 (8) | 14.89 (13) | 21.49 (8) | 15.61 (12) |
| **B7** | 20.56 (5) | 21.88 (6) | 21.45 (5) | 11.95 (5) | 9.21 (3) | 16.13 (4) |
| **loop VII** | 13.26 (11) | 13.37 (11) | 12.98 (9) | 8.80 (21) | 1.65 (19) | 6.85 (20) |
| **B8** | 17.12 (7) | 17.67 (7) | 21.08 (5) | 3.80 (9) | 5.76 (7) | 7.28 (9) |
| **C** | 30.40 (1) | 30.23 (2) | 32.36 (2) | 1.78 (2) | 2.23 (2) | 2.79 (2) |

The secondary temperature coefficients are averaged across all assignments, the beta-strands on N-terminal and C-terminal face and secondary structural elements, and each average also has the number of residues averaged in brackets.

Although the secondary temperature coefficients of holoSS pWT and E100G overlay almost perfectly (Figure 5.4A), differences between apoSH pWT and E100G are apparent (Figure 5.4B). Interestingly, although the global thermodynamic stability of apoSH pWT is far lower than that of holoSS, some regions of apoSH SOD1 appear to be more structured than holoSS pWT, perhaps as part of a mechanism for discouraging (toxic) aggregation in the absence of stabilizing disulfide bonds, metals, and dimerization. We observe evidence of this phenomenon near the C-terminal ends of β-strands 5 and 6 (residues 83-89 and 94-101, respectively), where the secondary temperature coefficients of apoSH pWT are notably higher than those of holoSS (Figure 5.4C). The E100G mutation prevents such structural changes in β-strand 6 by introducing a residue type (glycine) well known to have low propensity for β-sheet (Figure 5.4D).[54]

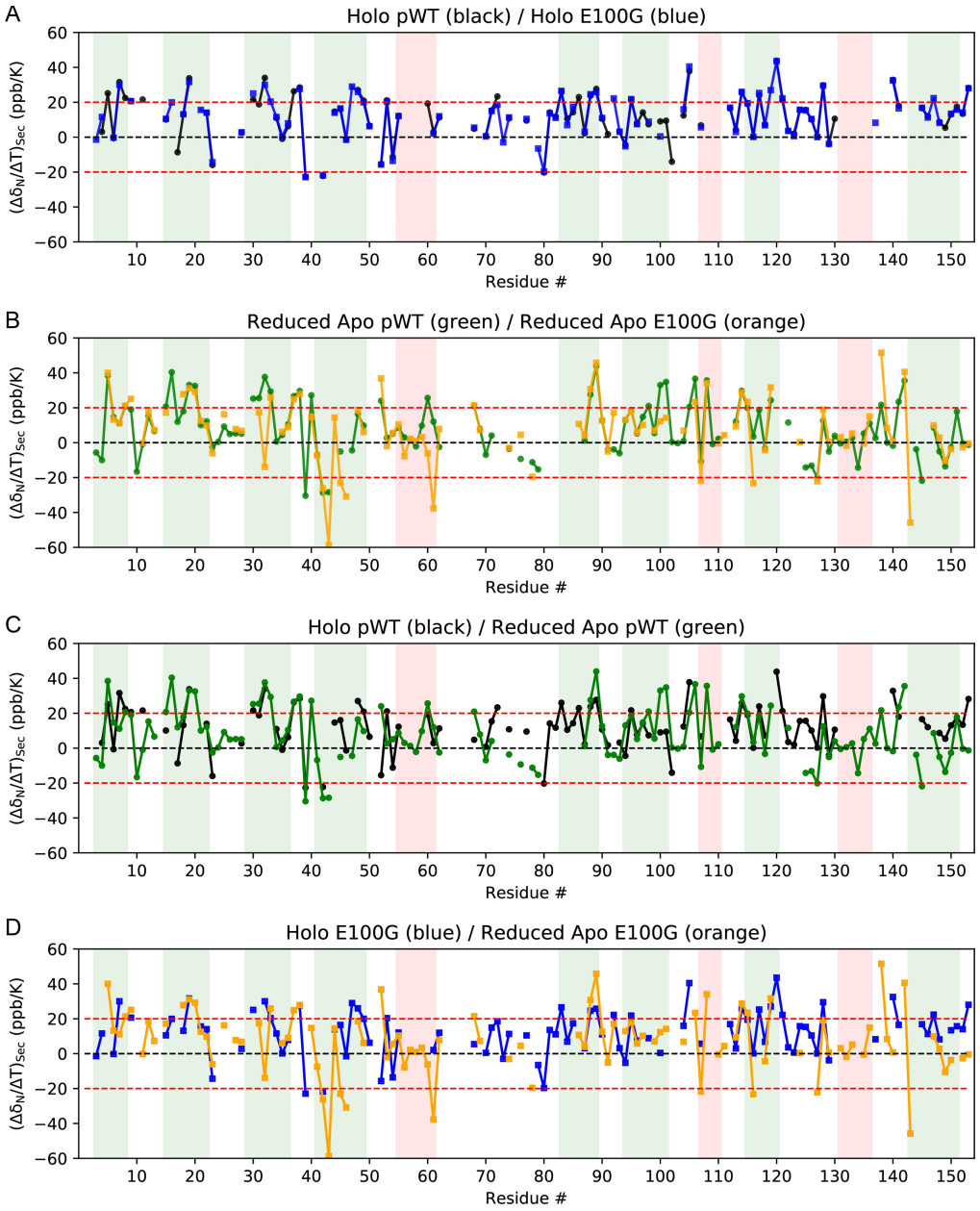



Figure 5.4: Secondary amide nitrogen temperature coefficients of SOD1 variants. A: holo (holoSS) pWT and E100G overlay near perfectly. B: differences between reduced apo (apoSH) pWT and E100G are readily apparent. C: parts of pWT *β*-sheets 5 and 6 (residues 83-89 and 94-101, respectively) may be more rigidly structured in reduced apo than in holo. D: the E100G mutation prevents rigidification of reduced apo *β*-strand 6. Secondary structure is indicated: *α*-helix (light red); *β*-strand (light blue). Dashed red lines demarcate a ±20 ppb/K range; secondary temperature coefficients outside of this range are particularly unlike random coil. Contiguous data points are joined by solid lines.

## 5.4 Conclusions

Variable-temperature amide proton and nitrogen NMR data are often collected in parallel, e.g., in a series of any 2D or higher $^{15}$N-edited/$^{1}$H-detected spectra. Although amide proton temperature coefficients are commonly calculated and interpreted, little use has been found for the corresponding amide nitrogen data. Calc$^{N}$ offers a new lens through which amide nitrogen temperature coefficients may be viewed, and we illustrate how subtleties such as structure transiently adopted by intrinsically disordered proteins, and perturbations induced by point mutation or protein maturation (e.g., disulfide bond formation, metal binding, and multimerization) may be detected.

## 5.5 Methods

### 5.5.1 NMR Sample Preparation

Purified $^{15}$N-labeled $^{10}$Fn3$_{D80E}$ in buffer containing 20 mM acetic acid/sodium acetate pH 4.0 and 90% $H_2O$/10% $D_2O$ was concentrated using an Amicon Ultra-4 centrifugal filter (EMD Millipore) to ~30 mg/mL.

Purified $^{15}$N-labeled holo SOD1 in buffer containing 20 mM hydroxyethyl piperazine N- 2-ethanesulfonic acid (HEPES) pH 7.8 and 90% $H_2O$/10% $D_2O$ was concentrated using an Ultracel 3K centrifugal filter (EMD Millipore) to ~30 mg/mL. The bound copper was reduced to $Cu^{1+}$ by addition of 5 mM sodium isoascorbate to avoid paramagnetic

line broadening by $Cu^{2+}$.

Purified $^{15}N$-labeled reduced apo SOD1 in buffer containing 20 mM HEPES pH 7.4, 1 mM tris(2-carboxyethyl)phosphine (TCEP), and 90% H2O/10% $D_2O$ was concentrated using an Amicon Ultra-4 centrifugal filter (EMD Millipore) to ~15 mg/mL.

### 5.5.2 Variable-Temperature NMR

Variable-temperature $^{1}H$-$^{15}N$ HSQC (Bruker pulse program 'hsqcetfpf3gpsi'[261–264]) $^{10}Fn3_{D80E}$ and SOD1 spectra were acquired using a Bruker AVANCE 600 MHz spectrometer equipped with a 5 mm triple resonance TXI probe. Amide proton chemical shifts are directly referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), while amide nitrogen chemical shifts are indirectly referenced to DSS using a $^{15}N/^{1}H$ Ξ ratio of 0.101329118.[265]

### 5.5.3 Amide Nitrogen Temperature Coefficients

$^{10}Fn3_{D80E}$ and SOD1 backbone resonance assignments were determined as previously described.[54,260] Assignments and variable-temperature NMR data for dsRBD-1 are as previously published.[245,266]

Amide nitrogen temperature coefficients of dsRBD-1, $^{10}Fn3_{D80E}$ and SOD1 variants were determined by fitting DSS-referenced amide nitrogen chemical shifts to a linear model. Secondary amide nitrogen temperature coefficients were determined by calculating the values expected for random coil peptides of the same amino acid sequences (Eq. 5) and subtracting these from experimentally measured temperature coefficients (Eq. 2).

# Chapter 6: Impact of salt concentration, protein concentration, and aggregate seeds on chemical shifts and peak shapes for immature Superoxide Dismutase-1 variants

Author Contributions:

Harmeen Kaur Deol: conceptualization, methodology, resources, validation, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft; Bao Khanh Nguyen: methodology, resources, validation, formal analysis; Elizabeth M. Meiering: conceptualization, supervision, project administration, funding acquisition.

## 6.1 Context

With all these different aggregate polymorphisms, it is clear point mutations and solutions conditions influence the aggregate structure. However, it is unclear how either influence the protein dynamics and in turn the aggregation pathway(s). Here, we explore how sensitive a 2D $^{15}$N-$^{1}$H HSQC spectrum is to protein concentration, salt concentration and protein aggregation, in addition to how these different factors influence apoSH SOD1 structure and dynamics. With this comparison, subtle differences in protein dynamics may be elucidated that would otherwise remain hidden in experiments that only capture a snapshot. Overall, by using NMR spectroscopy, we understand on the atomistic-level the impact different solutions conditions have on apoSH SOD1 protein dynamics, and ultimately parse the impact solutions conditions have on protein misfolding and aggregation.

## 6.2 Introduction

The application of high-resolution methods like nuclear magnetic resonance (NMR) spectroscopy to the study of proteins under varying solution conditions can offer powerful insight into protein structure, dynamics and interactions of molecules.[198,199] Through simplistic 2D NMR experiments, exploring both chemical shifts and peak shapes to gain insight into protein structure and dynamics. Previously, the temperature dependence of amide proton chemical shifts has been reported for various proteins.[66,203,211] The amide proton temperature coefficients for holoSS SOD1 specifically highlight the structured and unstructured parts of the native protein, and how the local structure changes upon the addition of point mutations or removal of maturation events. Additionally, in Chapter 3, the chemical shift perturbations of apoSH mutants relative to apoSH pWT SOD1 at folded and unfolding temperatures probe for local structural differences either due to point mutations or local unfolding, respectively. Moreover, by assessing the peak shapes for apoSH pWT and its mutants throughout the temperature ramp in Chapter 4, local areas of structural differences and initial protein unfolding were identified. Collectively these experiments further prove that simple 2D NMR experiments offer valuable insight on protein structure and dynamics.

However, due to their fast timescales, protein dynamics are difficult to capture and in addition to high-resolution methodology, they require identifying and exploring different solution conditions to help probe or amplify protein dynamics. In general, varying salt and protein concentration have been widely studied to explore protein dynamics related to folding and aggregation.[230,267–271] Additionally, SOD1 has been considered to have strain-

like behaviour[28], which can be probed by the presence of protein aggregates (in this case, an aggregate is anything larger than the monomer). More specifically, by varying salt concentration, protein concentration, or the presence of protein aggregates, apoSH SOD1 and its aggregates were characterized and showed the aggregate ensemble shifted structurally based on the mutation and solution conditions.[12] With these observed aggregate polymorphisms, the aggregation mechanism and more precisely, aggregate initiation remains elusive.

Here, we explore apoSH SOD1 protein structure and dynamics under variable salt concentration, protein concentration and aggregate seeding. In addition to these solution conditions, HSQC spectra are collected throughout a range of temperatures. The chemical shift and peak shape analyses are then compared to the values without the presence of the condition, for example the aggregate seeding experiment is compared to a sample without the presence of any aggregates. With this comparison, subtle protein dynamics may be elucidated that would otherwise remain hidden in static experiments. Overall, by using NMR spectroscopy, we understand on the atomistic-level the impact different solutions conditions have on apoSH SOD1 protein dynamics, and ultimately parse the impact solutions conditions have on protein misfolding and aggregation.

## 6.3 Methods

### 6.3.1 Protein preparation for NMR

Proteins for NMR samples were expressed using M9 minimal media under different growth conditions. Briefly, BL21 non DE3 *E. coli* were used, which do not confer any antibiotic resistance. Once transformed with the pHSOD1ASlacI$^{q}$, plasmid for

expressing SOD1, which confers ampicillin resistance, cells were grown overnight in LB broth with ampicillin. Next, 10 mL of the overnight culture was used to inoculate 1 L of M9 minimal media, which contains $^{15}NH_4Cl$ as the sole nitrogen source. Cells were grown to $OD_{600}$=~0.6, and then induced with IPTG. All protein samples underwent a copper charging and heat treatment before being purified by hydrophobic column chromatography to ensure proper metalation. Copper was added (final concentration of 1.4 mM) and then the sample was heated for 20 minutes at 70 °C. The solution had ammonium sulfate (final concentration of 3 M) added and then was filtered through a nitrocellulose filter (0.45 μm cutoff) prior to applying onto a hydrophobic-interaction column. The protein was eluted and placed in dialysis tubing (6-8 kDa) to be dialyzed against milliQ $H_2O$ for 4x8-hour exchanges. With the Amicon ultrafiltration membrane, the volume of the sample was reduced, and the concentration was checked with the use of the UV-vis spectrophotometer at 280 nm. Aliquots were then syringe-filtered into Eppendorf tubes and flash frozen in liquid nitrogen. Metals were removed via dialysis of holo SOD1 against ethylenediaminetetraacetic acid (EDTA) at pH 3.8.[208] Another set of exchanges were done with 50 mM sodium acetate, 100 mM NaCl at a pH of 3.8. The last set of exchanges were against milliQ water with one final exchange against 1 mM HEPES at pH 7.8. Metal contamination was assessed by PAR assay, initially developed by Crow *et al*.[209] and further modified by Colleen Doyle and Heather Primmer as described previously.[210]

Reduction of the disulfide bond was performed as described previously.[65] The method used the reducing agent tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl). Initially, oxidized apo SOD1 was denatured using 2 M guanidine hydrochloride

(GdnHCl) and 20 mM HEPES for a pH of 7.8. TCEP was then added to the sample, and it was placed in a desiccator for one hour before the sample was exchanged into the buffer (1 mM TCEP and 20 mM HEPES at pH 7.4). The success of the reduction was checked by running an SDS-PAGE gel. Preparation of the NMR sample required concentration of the sample to ~8-25 mg/mL SOD1, with a final sample volume of ~450 mL. 50 µL of $D_2O$ and sodium trimethylsilylpropanesulfonate (DSS) (final concentration of 1 mM) was added and the sample was transferred to an NMR tube (final sample volume of 500 mL). Upon final sample preparation, we also employed an initial $^{1}H$ 1D NMR experiment (p3919gp, Appendix H) to confirm both the TCEP reduction state and the metalation status (Appendix D and E).

6.3.2 NMR experiments

$^{1}H^{15}N$-HSQC spectra were acquired at varying temperatures from 1.2 °C to 38.6 °C using a Bruker Avance 600 MHz spectrometer, using the acquisition parameters described in Appendix J. The sample was equilibrated for 20 minutes at each temperature prior to acquisition, as described by Doyle *et al*.[66] All spectra were referenced to DSS. Spectra were processed in TopSpin 4.0.7, and analyzed in CCPN 2.4.2. Sequence specific resonances were previously determined for pWT, A4V, and V148I apoSH SOD1 at pH 7.4.[54,55]

Variable-temperature $^{1}H$-$^{15}N$ HSQC (Bruker pulse program `hsqcetfpf3gpsi'[261–264]) SOD1 spectra were acquired using a Bruker AVANCE 600 MHz spectrometer equipped with a 5 mm triple resonance TXI probe. Amide proton chemical shifts are directly referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), while amide nitrogen chemical shifts are indirectly referenced to DSS using a $^{15}N/^{1}H$ ratio of 0.101329118.[265]

Assignments and variable-temperature NMR data for dsRBD-1 are as previously published.[245,266] Amide nitrogen temperature coefficients of dsRBD-1 and SOD1 variants were determined by fitting DSS-referenced amide nitrogen chemical shifts to a linear model. Secondary amide nitrogen temperature coefficients were determined by calculating the values expected for random coil peptides of the same amino acid sequences (Eq. 5 in Chapter 5) and subtracting these from experimentally measured temperature coefficients (Eq. 2 in Chapter 5).

In CCPN 2.4.2, the intensity and linewidths were fit only in the direct ($^{1}$H) dimension using Gaussian peak shape with a boxwidth of 1. Due to significantly decreased dispersion in the indirect ($^{15}$N) dimension relative to the direct dimension, only the direct dimension was considered. Peak intensities were normalized to the TRP-32 folded side chain and slopes were fit to a line using Microsoft Excel. Maximum peak intensity was assessed using Microsoft Excel. The $^{1}$H linewidths were estimated from the width of the peak at half of the maximum height of the Gaussian peak fit.

## 6.4 Results and Discussion

Unlike the highly stable holoSS SOD1, apoSH SOD1 is marginally stable and cannot be heated past physiological temperatures without protein unfolding occurring. However, with the addition of variable solution conditions that may be destabilizing, we must limit additional processes while ensuring all other conditions are the same; the temperature range must be similar between the conditions and must encompass a range wherein the protein is considered folded. Additionally, based on the solution conditions, which includes increasing salt concentration, increasing protein concentration or

presence of aggregates, the mutant consideration is vital to ensure as few additional states are observed (further discussed below). Due to the diverse structural and thermal properties of apoSH SOD1 mutants, events such as local protein unfolding, native protein association, and non-native protein association are likely. Thus, to minimize the influence of some of these factors on the data, the proteins are compared on a comparable range of temperature from 9-18 °C for the temperature coefficient data (Table 6.1), similar to Chapter 3, and their whole folded temperature range based on their tryptophan 32 folded sidechain (TRP-32F) (Table 6.2), as previously discussed in Chapter 4.

Table 6.1: Temperature range for the temperature coefficient data.

| Mutants | Temperature range |
|---|---|
| A4V unagg[a] | 8.87 - 17.94 |
| A4V agg[a] | 7.48 - 17.86 |
| V148G LC[b] | 8.85 - 18.04 |
| V148G HC[b] | 8.61 - 17.92 |
| V148I no salt[c] | 9.30 - 19.08 |
| V148I salt[c] | 11.73 - 19.40 |

This temperature range data is based on the same temperature range using in Chapter 3.
[a]unagg and agg are short forms for the A4V sample without aggregates and with aggregates, respectively.
[b]LC and HC stand for low and high protein concentration, respectively.

[c]No salt and salt represent the V148I samples without NaCl and with 150 mM NaCl, respectively.

Table 6.2: Folded and unfolding temperature ranges for the apoSH mutants and their respective number of spectra acquired in brackets.

| Mutants | Folded[d] | Unfolded[e] |
|---|---|---|
| A4V unagg[a] | 5.65 - 17.94 (9) | 19.8 - 34.5 (9) |
| A4V agg[a] | 3.67 - 17.86 (9) | 19.49 - 36.32 (9) |
| V148G LC[b] | 1.04 - 19.88 (12) | 20.11 - 36.70 (9) |
| V148G HC[b] | 1.20 - 19.55 (12) | 21.05 - 36.03 (9) |
| V148I no salt[c] | 9.30 - 39.55 (13) | - |
| V148I salt[c] | 11.73 - 26.94 (7) | 26.97 - 38.64 (6) |

- for mutants not considered.

[a]unagg and agg are short forms for the A4V sample without aggregates and with aggregates, respectively.

[b]LC and HC stand for low and high protein concentration, respectively.

[c]No salt and salt represent the V148I samples without NaCl and with 150 mM NaCl, respectively.

[d]The folded temperature range for each sample based on the Trp-32F intensity and DSC data.

[e]The unfolding temperature range for each sample based on the Trp-32F intensity and DSC data.

### 6.4.1: The influence of salt concentration on apoSH V148I

The impacts of salt on proteins are complex, where both different cations and anions have a diverse impact across proteins. Previously, our group has shown the impact of NaCl and $Na_2SO_4$ on apoSH SOD1 aggregation is not limited to the protein but is highly dependent on the mutation.[12] For instance, A4V readily aggregates under any solution condition, both in the presence or absence of these salts, whereas V148I shows overall minimal aggregation under any solution condition but tends to form some population of smaller oligomers such as dimers under the presence of salt. However, this diversity in behaviour under the same solution condition further highlights the need to explore the impact salt has on the protein structure and dynamics that subsequently populate different states. Due to V148I showing such subtle aggregation with NaCl, and otherwise being considered pWT-like, it is an ideal mutant to assess the impact of salt on an apoSH mutant, without the concerns of significant protein association.

The reference sample is V148I without the addition of NaCl, whereas the salt sample is V148I with the addition of NaCl, which will henceforth be referred to as salt. Salt is titrated into the salt sample at a final salt concentration increase of 25 mM for each increment at 298 K, and at each increment a $^{1}H^{15}N$-HSQC spectrum is acquired. Throughout the salt titration, both pH and presence of higher order species are assessed

using a pH meter and dynamic light scattering (DLS) experiments, respectively. There are no observable changes in either. With a final salt concentration of 150 mM, to match both expected intracellular concentration[118] and previous apoSH aggregate studies[12], the sample is heated throughout a temperature ramp comparable to the reference sample (9.3 - 39.55 °C and 11.73 - 38.64 °C for the reference and salt sample, respectively) (Table 6.2). The reference and salt samples have HSQC spectra collected through the temperature ramp, and these data in addition to the salt titration spectra are analyzed below.

6.4.1.1 The chemical shift perturbations (CSP) for the salt titration highlight the non-specific interactions with NaCl are localized around the metal-binding sites

Salt is added to the salt sample of V148I at 25 mM increments until a final concentration of 150 mM NaCl at 298 K. With the addition of salt at a set temperature of 25 °C, the internal sample temperature did increase from 25.22 °C at 0 mM NaCl to 27.30 °C at 150 mM NaCl. Although minimal, it is important to consider this change when samples are assessed for peak shape changes. Notably, Sengupta *et al*. perform a salt titration with NaCl to observe oligomerization of wild-type mouse prion protein, and show CSP values larger than zero throughout the structure, suggesting they might be seeing the influence of temperature during their salt titration. However, since the temperature change is minimal and CSP values are considered significant only if they are more than average value + 1 standard deviation, the minimal temperature-influenced chemical shifts do not influence the analysis.

CSP values are useful for mapping changes in chemical shifts that can represent residues involved in events such as binding or structural changes due to a changing condition. Here, CSP values are calculated in two ways: one looks at only the difference between 0 mM (initial spectrum) and 150 mM (the last titration spectrum) (Figure 6.1A), and the second is using combined CSP values for all 25 mM NaCl titrant interval spectra from 0 mM to 150 mM (Figure 6.1B). The latter measurement is to assess if the chemical shift changes are gradual or based on a threshold of a certain salt concentration. The salt titration CSP values are calculated and mapped onto the 3D structure of holoSS SOD1 (Figure 6.1), to show which residues are most impacted by the addition of salt. For CSP values and combined CSP values, values larger than average + 1 standard deviation are considered perturbed from the reference (no salt sample), whereas values smaller than that are considered to be comparable to the reference.

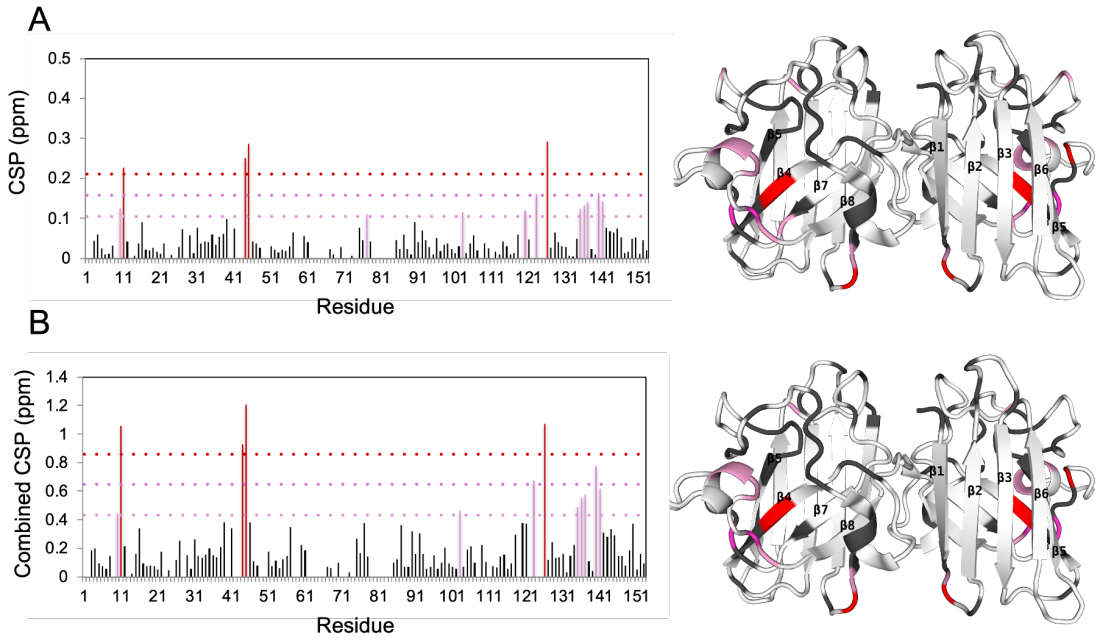



Figure 6.1: Chemical shift perturbations and combined chemical shift perturbations for V148I throughout the salt titration. CSP values were coloured based on the same cutoff for both samples. Residues are

coloured to aid in visualization according to the relative magnitude of CSP values. The pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1, (A) CSP values for V148I with 0 mM and 150 mM NaCl represented by the bars on the plot and mapped on to the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). (B) Combined CSP values for V148I throughout the salt titration represented as bars on the plots and mapped on to the 3D ribbon structure.

Both CSP values and combined CSP values show similar residues being perturbed with the addition of salt (Figure 6.1). Based on the CSP values between 0 mM and 150 mM NaCl spectra, the residues near metal-binding sites, more specifically localized in the electrostatic loop (loop 7), show significant perturbations. Additionally, these are the same residues that are perturbed for combined CSP values, and the combined CSP values ranged from 0.02 to 1.20. The highest combined CSP values are for residues D11, F45, and L126. At each of these residues, neighbouring residues also show significant CSP perturbations, supporting the impact of the salt propagating to the neighbouring residues.

Moreover, the impact of salt seems predominantly localized to the metal-binding region and the CSP values throughout the titration for perturbed residues seem to show the impact is additive as each salt titration interval (Figure 6.2). The residues most perturbed being at the site of metal-binding can be explained by the considerably positively charged pocket at the metal-binding sites, which is functionally used for guiding the superoxide anion to the active site.[272] Hence, the binding may be focused on the anion, $Cl^-$ interacting non-specifically with the positive charges around the metal-binding site.

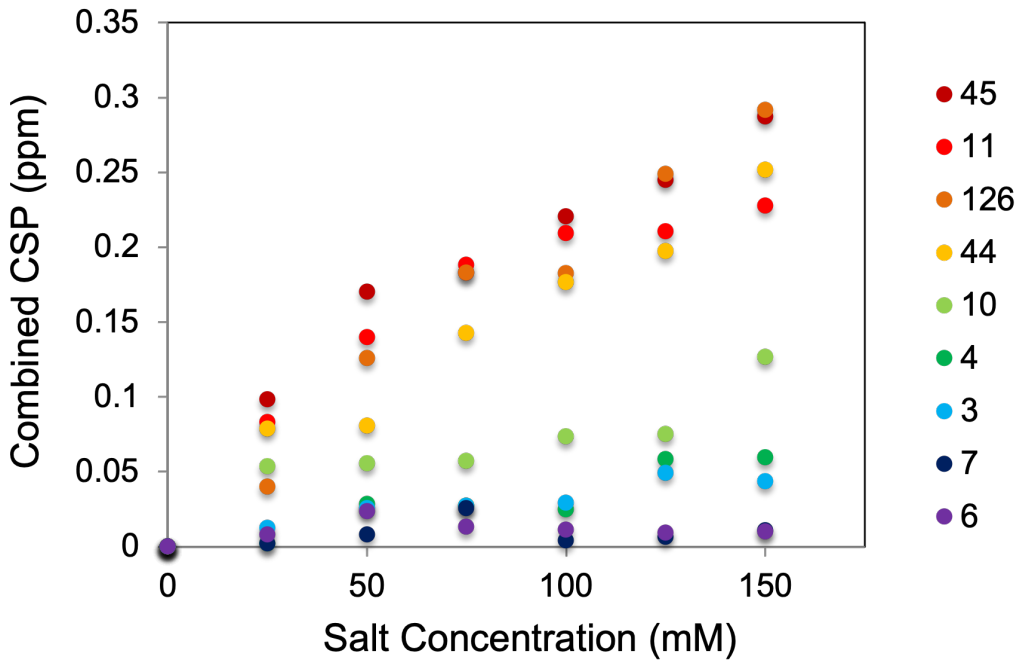



Figure 6.2: Combined CSP values for V148I throughout the salt titration for nine representative residues. Residues 45, 11, 126 and 44 show high CSP values throughout the NaCl titration, but residues 10, 4, 3, 7 and 6 do not.

Additionally, although the range of combined CSP values here is larger than observed for the salt titration on mouse prion protein (moPrP), the profiles of the CSP values against the salt concentration for the perturbed residues are remarkably similar.[230] From the CSP dependence on NaCl concentration for moPrP, they conclude the NaCl is weakly binding to the protein. This aligns with what is expected of NaCl interacting with SOD1, since there are minimal changes in the native monomer structure observed by FTIR upon the addition of NaCl.[12] Thus, the impact of salt is non-specific and focused around positively charged residues near the metal-binding sites.

6.4.1.2: Amide proton temperature coefficient values suggest the addition of salt promotes both local structuring and unfolding

Amide proton temperature coefficient values can report on loss of structure and intramolecular hydrogen bonding.[202,273] In Chapter 3, amide proton temperature coefficients are considered for several apoSH mutants, including V148I. Here, we assess the differences in temperature coefficient values for V148I with and without the addition of 150 mM NaCl in the apoSH V148I's folded temperature range (Table 6.1).[65] Also in Chapter 3, we discuss multiple ways of averaging temperature coefficients, and how each method shows different but important information. Similarly, the different averaging methods are used here to discuss the differences in temperature coefficient values for V148I with and without salt.

Table 6.3: Amide proton temperature coefficients for apoSH V148I without (no salt) and with salt.

| | V148I no salt | V148I salt |
|---|---|---|
| average | -4.82 (121) | -4.78 (115) |
| standard dev | 2.95 | 2.63 |
| average ddTC | -0.08 (119) | -0.11 (114) |
| | | |
| N-terminal B-sheet | -3.49 (31) | -3.59 (31) |
| C-terminal B sheet | -5.09 (23) | -4.73 (22) |
| N | -1.48 (1) | - |
| B1 | -4.41 (6) | -4.73 (6) |
| loop I | -4.90 (3) | -5.42 (3) |
| B2 | -2.28 (9) | -2.16 (9) |
| loop II | -5.28 (3) | -5.42 (3) |
| B3 | -2.26 (8) | -2.18 (8) |
| loop III | -6.24 (3) | -3.95 (2) |
| B4 | -7.01 (6) | -5.44 (6) |
| loop IV | -5.64 (19) | -5.66 (17) |
| B5 | -4.55 (4) | -4.83 (4) |
| loop V | -3.59 (4) | -3.69 (4) |
| B6 | -5.02 (8) | -5.30 (8) |
| loop VI | -4.67 (12) | -4.87 (11) |
| B7 | -3.93 (4) | -4.18 (5) |
| loop VII | -5.84 (20) | -5.98 (20) |
| B8 | -4.86 (9) | -4.46 (7) |
| C | -6.54 (2) | -6.82 (2) |

Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets. Average ddTC is the difference in average temperature coefficients relative to apoSH pWT.

Averaging all the temperature coefficient values for the assigned residues masks the local differences between the two samples (Table 6.3). The two average values, -4.82 ppb/K and - 4.87 ppb/K for the without and with salt samples, respectively, are not considered significantly different, however, the N-terminal and C-terminal ß-sheets values highlight there are more local differences (Table 6.3). More specifically, the apoSH V148I C-terminal ß-sheet is more stabilized in the presence of salt, while the N-terminal ß-sheet shows subtle increases in dynamics upon the addition of salt. Interestingly, according to the CSP values, the residues considered to be perturbed by the salt are predominantly located on the C-terminal ß-sheet (Figure 6.1). However, an in-depth assessment that probes more local areas of difference is necessary to understand the impact of salt on apoSH SOD1.

Furthermore, by considering the various averaging methods discussed in Chapter 3, local areas of differences are identified. By comparing the temperature coefficient profiles across the entire sequence for V148I with and without salt, it is clear the profile is predominately similar (Figure 6.3A), which agrees with the average temperature coefficient values (Table 6.3). Despite this, there are some points of localized difference such as residues responsible for forming the native helix in the electrostatic loop and ß8. Intriguingly, the residues in the electrostatic loop show increased dynamics upon the addition of salt, whereas ß8 becomes more stabilized upon the addition of salt. However, evaluating differences between the two samples based on individual residues without considering neighbouring residues can be misleading due to noise and/or anchoring data points.

Moreover, local areas are further probed by using rolling averages and averaging based on secondary structures. By using rolling averages over 5 residues, local structural differences can be easily discerned, without outliers having a significant impact. The rolling averages highlight that ß4 is more stabilized upon the addition of salt, while its neighbouring strand, ß7 is more destabilized (Figure 6.3B). Additionally, edge strands on the N-terminus ß-sheet, ß1 and ß6, are also more destabilized (Figure 6.3B), which explains the N-terminus ß-sheet being more dynamic in the salt sample (Table 6.3). Additionally, loop 7 also shows enhanced dynamics, specifically around the residues responsible for forming the native helix.

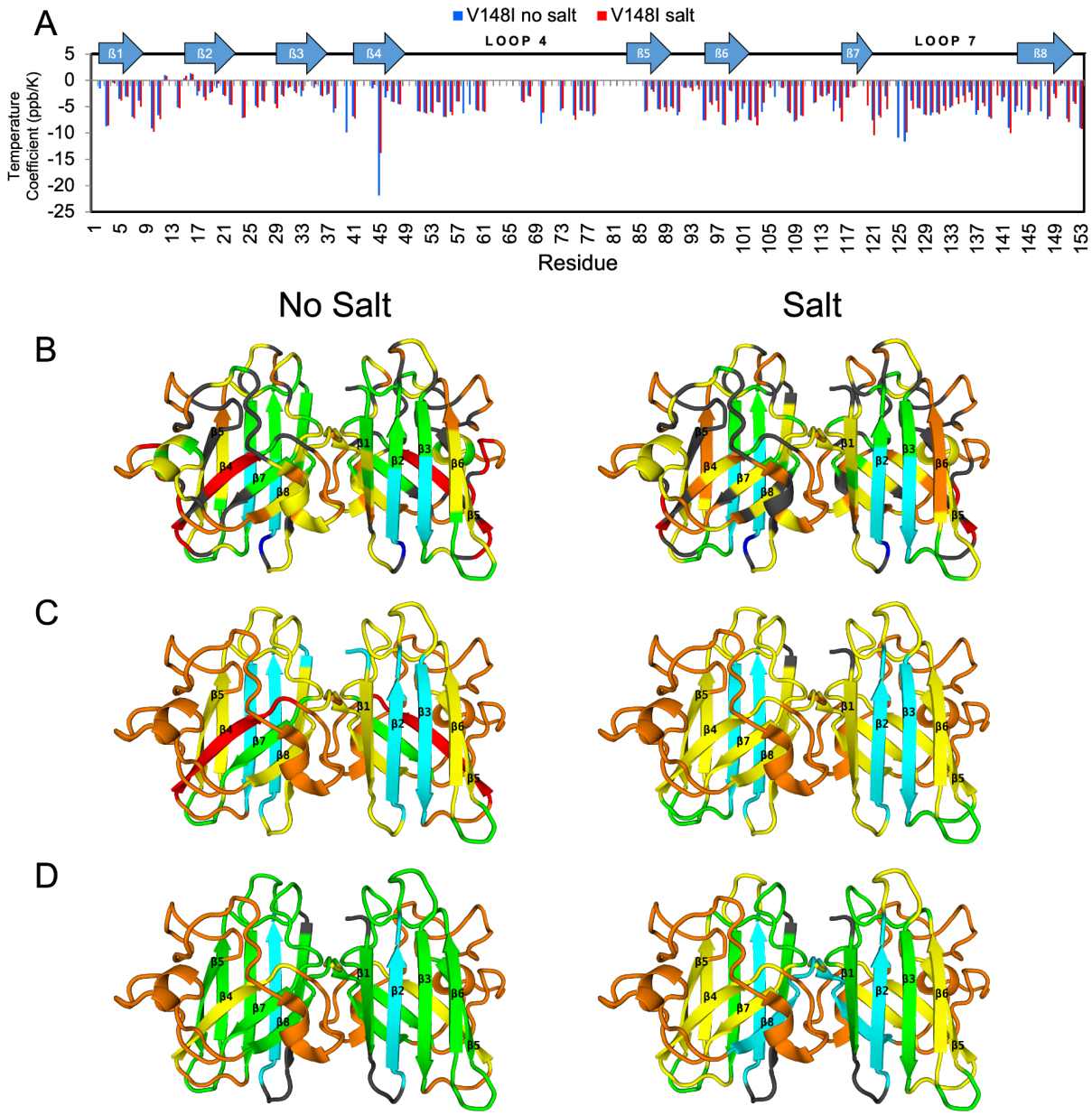



Figure 6.3: Amide proton temperature coefficients and their multiple averaging methods for apoSH V148I without salt (no salt) and with salt. (A) Temperature coefficients are represented by the blue (no salt) and red (with salt) bars in the plot. (B-D) show 3D ribbon structures for the no salt (left) and with salt (right) data. (B) Rolling average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). (C) Secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). (D) Averaging only hydrogen bonded residues, secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue for the more stable. Dark blue is for values larger than -1 ppb/K, cyan is for values

between -1 and -2.5 ppb/K, green is for values between -2.5 and -4 ppb/K, yellow represents values between -4 to 5.5 ppb/K, orange is for values between -5.5 to -7 ppb/K and values smaller than -7 ppb/K are coloured red. Grey indicates residues where no data are available.

Furthermore, the rolling averages highlight that the temperature coefficients for SOD1 residues have grouped behavior that is based on the secondary structure present in the mature SOD1 enzyme, aside from some outliers (Figure 6.3B). By averaging based on secondary structure, the changes across the secondary structures are considered. Here the key differences are in ß4 again showing stabilization and its neighbouring strand, ß7 becoming destabilized upon salt addition (Figure 6.3C), which is also observed in the rolling averages (Figure 6.3B). However, the other differences observed by rolling averages are no longer present, suggesting their differences are more local and impact fewer residues. Comparing rolling averages and averages based on secondary structure highlights that ß4 and ß7 are showing considerable differences upon the addition of salt.

Although averaging by secondary structure shows potential structural differences, averaging only the hydrogen bonded residues by their secondary structure may highlight changes in hydrogen bonding to explain the structural differences. The averaging of only the hydrogen bonded residues based on their secondary structure highlights significant changes in hydrogen bonding upon the addition of salt. While ß-strands 5-7 show more dynamics in the hydrogen bonded residues, ß4 and ß8 show there is gain of stability (Figure 6.3D). Aside from ß8, the number of residues averaged are comparable; ß8 has two residues not assigned in the salt sample, thus the gain of stability is not considered significant. Irrespective of ß8, the stabilization of ß4 and coincidentally the destabilization of its neighbouring strands, ß5, 6 and 7 suggest there are significant interruption to the hydrogen bonding upon the addition of salt.

The changes in temperature coefficients in apoSH V148I due to the addition of salt suggest there are more local changes throughout the structure. A significant change appears to be in ß4, and then propagates to the neighbouring strands and loop 7 (Figure 6.3). From the CSP values, V148I shows a pWT-like structure that is perturbed, mostly around ß4 with the addition of salt (Figure 6.1). The salt is thought to non-specifically interact with residues near ß4, and the temperature coefficient values suggest the residues become stabilized (Figure 6.2). However, this stabilization in ß4 from the salt interacting with its residues subsequently results in residues near it changing local structure and becoming more dynamic, as suggested by their CSP values and the temperature coefficient data, respectively (Figure 6.1 and 6.3). Although the perturbed strands are not considered to be involved in native or non-native dimerization as previously observed for apoSH mutants[54,55], we find that apoSH V148I can form dimer species in the presence of salt.[12] Thus, by probing protein dynamics in different solution conditions, additional dynamics that may play a role in native-like or non-native like processes may become prevalent that would otherwise be suppressed.

6.4.1.3 Intensity analysis

6.4.1.3.1 Trp-32F and Trp-32U peak intensities suggest NaCl is destabilizing and shifts the apoSH V148I toward unfolded structure

In Chapter 4, the two sidechain peak intensities for tryptophan-32 (Trp-32) correlated to protein unfolding. However, where one peak intensity decreased as the temperature entered the protein unfolding temperatures, the other increased. The decreasing intensity peak represents the Trp-32 sidechain in the folded state (Trp-32F) and the increasing intensity peak represents the Trp-32 sidechain in the unfolded state

(Trp-32U). Additionally, Trp-32F correlates well to protein unfolding based on their DSC thermograms for apoSH pWT and its mutants (discussed further in Chapter 4). For this reason and the fact that Trp-32U does not have an intensity for the more thermally stable SOD1s, Trp-32F is the ideal candidate for intensity normalization.

Here, we extend the Trp-32F intensity indicating protein folding and unfolding, discussed in Chapter 4, to assess changes to apoSH V148I in 150 mM NaCl relative to V148I without salt. Trp-32F correlates well with DSC endotherms[65], where increases in intensity represent the folded temperature range, and any decrease in intensity suggests the protein is starting to unfold. The V148I without salt spectra are only collected in the folded temperature range, and accordingly the Trp-32F peak only shows increasing peak intensity (Figure 6.4A). However, with the addition of salt, the Trp-32F peak now shows a decrease in peak intensity by 34 °C (Figure 6.4A), suggesting apoSH V148I unfolds at earlier temperatures in the presence of salt. Similar to Trp-32F intensity changes suggesting apoSH V148I in salt is unfolding earlier than without salt, the Trp-32U peak is also showing increases in intensity at 34 °C for the salt sample (Figure 6.4A). Thus, both Trp-32F and Trp-32U changes in the salt sample suggest salt promotes earlier protein unfolding.

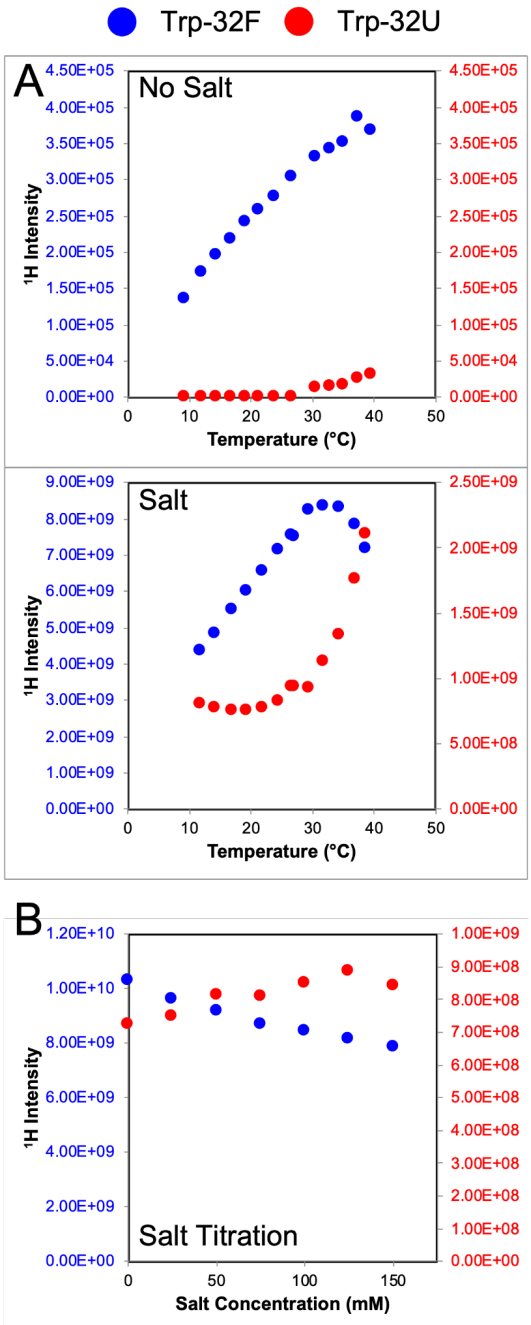



Figure 6.4: Trp-32F and Trp-32U peak intensities throughout the temperature ramp and salt titration. All peak intensities were normalized by receiver gain. The Trp-32F (blue) and Trp-32U (red) peaks are evaluated throughout the (A) temperature ramp and (B) salt titration. (A) The Trp-32F and Trp-32U peak intensity for the apoSH V148I samples without salt (top) and with salt (bottom). Throughout the folded temperature range, the TrpF-32F peak intensity increases, whereas the Trp-32U peak shows minimal changes. However, the Trp-32F and Trp-32U intensity begins to decrease and increase, respectively, when the protein is in the unfolding temperature ranges. (B) The Trp-32F and Trp-32U peak throughout the salt titration at 298 K. At 298K, raV148I experiences no protein unfolding according to its DSC endotherm. However, with each addition of salt the Trp-32F peak decreases in intensity while the Trp-32U peak increases in intensity.

Akin to Trp-32F, Trp-32U also behaves differently for V148I in the presence of salt relative to V148I with no salt. Without the addition of salt, V148I Trp-32U peak has negligible signal (Figure 6.4A). It is only with the addition of salt that V148I has a Trp-32U peak with considerable signal (Figure 6.4A). Interestingly, the thermally destabilized apoSH mutants consistently show a Trp-32U peak in their spectra, even in the folded range (Chapter 4). This may suggest that the equilibrium between the folded and unfolded states may shift more toward the unfolded state for the destabilized mutants relative to the stabilized ones. The presence of the Trp-32U peak in V148I with salt suggests the addition of the salt is shifting the folded-unfolded equilibrium toward more unfolded, whereas without salt, the equilibrium exclusively favours the folded state.

Moreover, the shift in the folded-unfolded equilibrium is also supported by the Trp-32F and Trp-32U intensity changes throughout the salt titration. With each subsequent salt titration, the Trp-32F intensity decreases and the Trp-32U intensity increases (Figure 6.4B). For the changes not being remarkably linear, it would have suggested the addition of salt is impacting the NMR parameters, such as the tumbling and relaxation, which subsequently impact the peak shape. However, with both the Trp-32F and Trp-32U peaks showing comparable changes in signal (Figure 6.4B), it is plausible the changes are due to the presence of salt shifting the relative folded and unfolded populations. Additionally, we have previously shown by attenuated total reflectance - Fourier transform infrared spectroscopy (ATR-FTIR) that the addition of NaCl shifts the apoSH mutants away from beta-structure and toward a more unfolded structure.[12] Therefore, in combination with the presence of Trp-32U peak intensity with salt and its increasing intensity with the addition

of salt relative to the decreasing intensity of the Trp-32F peak, the presence of salt, specifically NaCl, shifts the folded-unfolded equilibrium for apoSH SOD1 toward unfolded.

6.4.1.3.2 The large loops show significant points of perturbation for apoSH V148I in the present of salt relative to the absence of salt

Due to the large range of intensity profile behaviour, characterizing individual residues based on their intensity profiles becomes onerous. As discussed in Chapter 4, without a point of reference, intensity profiles are difficult to interpret, and adding additional solution conditions such as temperature and salt add to the difficulty. Thus, each residue intensity is normalized to its Trp-32F peak at all temperatures for apoSH V148I spectra (both with and without salt), which removes the nonlinear influence of temperature or changes in protein concentration on peak shape, without changing the temperature-dependent intensity profile (discussed further in Chapter 4). Since Trp-32F peak intensity correlates well with the folded conformer of the protein throughout the temperature ramp, and thus by normalizing all residues' intensity at each temperature to their Trp-32F peak, the initial analysis is for assessing the change in folded structure for V148I with the addition of salt.

Since V148I without salt does not have any spectra collected in the unfolding temperature range, only intensity profiles in the folded temperature range are considered (henceforth referred to as the folded intensity slopes). As discussed in Chapter 4 and above, all residue intensity profiles are normalized by their respective Trp-32F peak. The post-normalized folded intensity slopes that have values within the average ± 1 standard deviation mean their pre-normalized intensity slope is similar to the Trp-32F slope, and

thus these residues are considered to respond similarly to temperature and consequently share comparable structure (Figure 6.5). However, larger differences between pre-normalized slopes and Trp-32F slopes result in a post-normalization intensity slope that has an absolute value larger than zero, meaning the residue deviates from the folded structure.

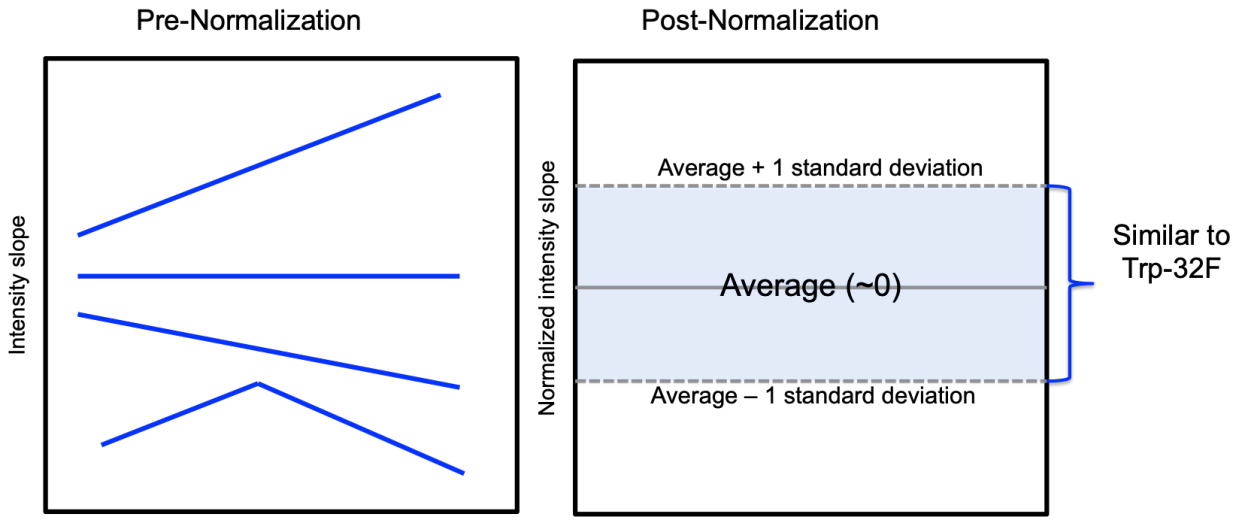



Figure 6.5: The various pre-normalization intensity profiles observed and the interpretation of the post-normalized intensity slopes. The plot on the right shows all the potential intensity profiles seen for the residues before normalization. Due to the diversity in profile behaviours, interpretation is difficult. However, upon normalization to the Trp-32F peak, the interpretation of slope of intensity is now in reference to folded/structured conformer. The average slope value is close to 0 (solid grey line), and slope values within the average ± 1 standard deviation (blue shades box) are considered as behaving comparably to that of the Trp-32F peak, and thus are considered structured. Outside this range, the slopes are reporting on less structured/differently structured relative to the Trp-32F.

For apoSH V148I with and without salt, the normalized slopes show considerable differences in values throughout the structure. Even though the post-normalized intensity slope profiles throughout the structure are comparable in the two samples (Figure 6.6), there are differences in the values between the samples suggesting some are moving toward structured or away from structured. The most notable differences are in the large loops, loops 4 and 7, where the intensity slopes for the salt sample are consistently more

negative than the sample without salt (Table 6.4). Although the large loops in both samples have intensity slopes significantly different from that of the Trp-32F slope (Figure 6.6), and thus are likely not in a folded conformation, the differences between the two samples suggest the loops behave differently between the two samples (Figure 6.6 and Table 6.4). The large loops are points of significant deviation both relative to Trp-32F and between the salt and no salt sample.

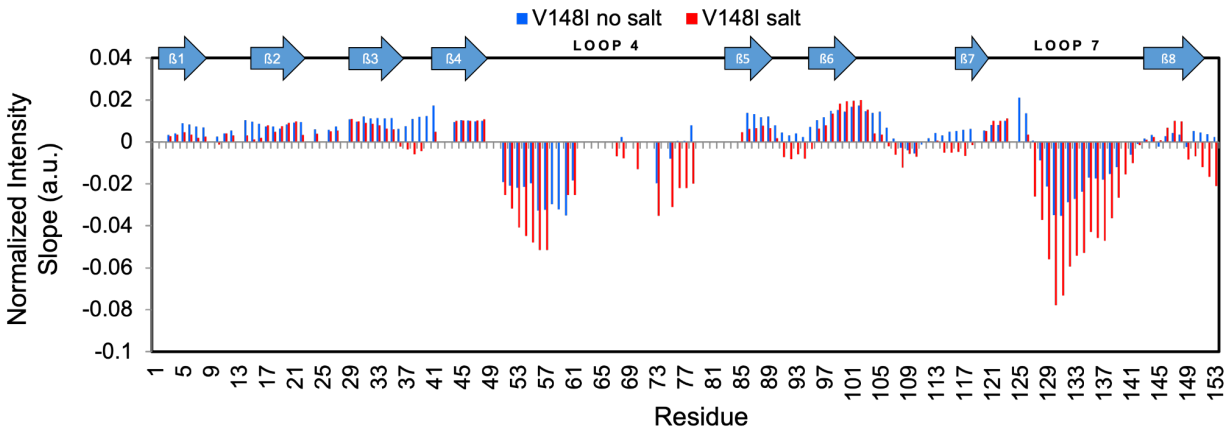



Figure 6.6: Post-normalized folded intensity slopes for apoSH V148I without salt (no salt) and with salt (salt). The rolling averages for normalized intensity slopes are considered for the folded temperature ranges. The rolling averages for normalized intensity slopes are for the sample without salt (blue) and with salt (red). The closer the values are to 0, the residue responds to temperature changes comparable to the Trp-32F peak that the data are normalized by.

Table 6.4: Average normalized intensity slopes for apoSH V148I without (no salt) and with salt.

| | V148I no salt | V148I salt |
|---|---|---|
| average | 0.0010 (124) | -0.0075 (120) |
| standard dev | 0.01339 | 0.02087 |
| | | |
| N-terminal B-sheet | 0.0095 (29) | 0.0068 (29) |
| C-terminal B sheet | 0.0079 (24) | 0.0034 (25) |
| N | 0.0002 (1) | - |
| B1 | 0.0065 (6) | 0.0032 (6) |
| loop I | 0.0040 (3) | 0.0020 (3) |
| B2 | 0.0085 (9) | 0.0054 (9) |
| loop II | 0.0063 (3) | 0.0048 (3) |
| B3 | 0.0105 (8) | 0.0071 (8) |
| loop III | 0.0107 (4) | -0.0046 (3) |
| B4 | 0.0112 (6) | 0.0094 (6) |
| loop IV | -0.0158 (19) | -0.0295 (17) |
| B5 | 0.0128 (4) | 0.0063 (5) |
| loop V | 0.0048 (4) | -0.0050 (4) |
| B6 | 0.0116 (8) | 0.0092 (8) |
| loop VI | 0.0049 (13) | 0.0004 (12) |
| B7 | 0.0055 (5) | -0.0025 (5) |
| loop VII | -0.0103 (20) | -0.0314 (20) |
| B8 | 0.0023 (9) | 0.0004 (9) |
| C | 0.0031 (2) | -0.0188 (2) |

Intensity slope values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

Aside from the loops, there are additional structures that show subtle differences between the two samples, but the differences are not considered significant relative to the Trp-32F peak. For instance, residues in ß4 and ß6-8 show positive intensity slopes in the no salt sample, but the values are negative in the salt sample (Figure 6.6 and Table 6.4). These strands are coincidentally also the strands that show differences in temperature coefficient values between the two samples (see section 6.4.1.2). It is important to note these residues' temperature coefficients in salt do not all show increased dynamics relative to the salt-free sample. Instead, the temperature coefficients show a mixture of behaviour between the strands, contrary to the similar shift from positive to negative folded intensity slopes for these strands. Thus, intensity slope changes

between the two samples may highlight the subtle differences in protein dynamics due to the addition of salt.

Overall, the areas that show perturbed behaviour in the salt sample relative to the salt-free sample, most notably the large loops and to some extent ß4 and ß6-8, show similar trends in their pre-normalized intensity profiles. In these perturbed areas, their intensity profiles with salt present show an initial increase in intensity followed by a decrease in intensity at higher temperatures. Similar to all the apoSH SOD1s discussed in Chapter 4, the loops for the salt sample show the temperature at which the decrease in intensity occurs, is much earlier than the rest of the structure. By extending the analysis for Trp-32F to the large loops, the initial increase in the signal can be representative of the residues in a more folded conformation, and the decrease in signal suggests the loops are undergoing structural changes that can be interpreted as local unfolding. Hence, it is interesting to consider which residues, and at what temperature, show hints of local unfolding.

6.4.1.3.3 Max temperature data suggest the salt sample collectively unfolds sooner than the salt-free sample

Since differences between the intensity profiles for V148I with salt and without salt manifest as differences in folded intensity slopes, the profiles need to be examined further to understand what these different profiles may be reporting on. Considering the loops for V148I without salt show a different intensity profile compared to the rest of the folded structure (see Chapter 4), it suggests these loops respond to temperature differently, and highlight protein unfolding may begin at the loops. The temperature that corresponds to

the maximum peak intensity for each residue will henceforth be referred to as the max temperature (Figure 6.7A). Since the salt-free V148I does not enter its protein unfolding according to their DSC endotherms[65] and its Trp-32F peak profile (Figure 6.4A), most of its residues only show signal increases (Chapter 4, Figure 4.2B), and thus their max temperature corresponds to the temperature of the last spectrum acquired (Chapter 4, Figure 4.9A).

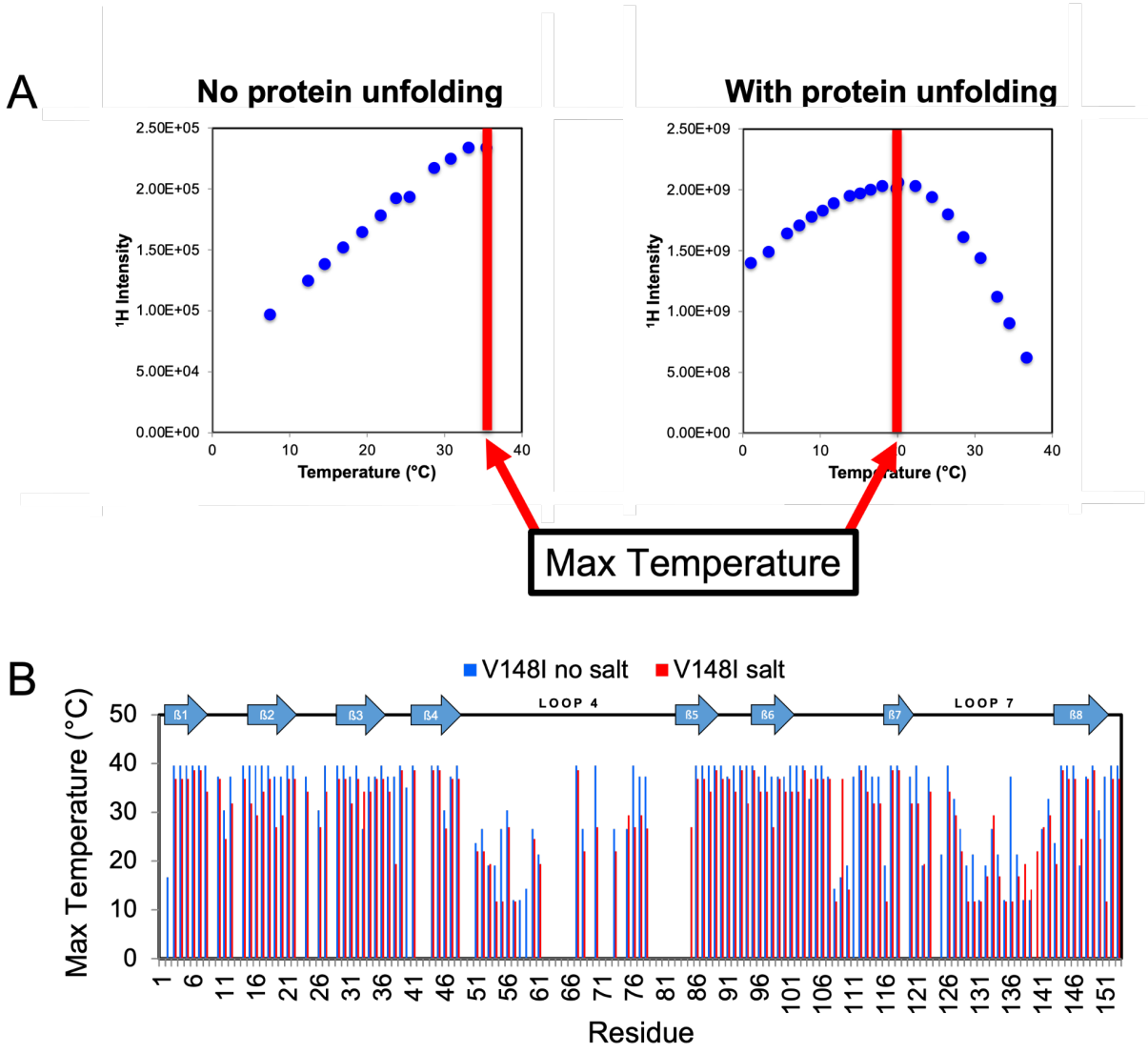


Figure 6.7: Maximum intensity temperature highlights local protein unfolding. (A) The temperature at which the highest peak intensity is observed is denoted as the max temperature (vertical red line). For V148I without salt, spectra are only acquired in the folded temperature range, thus the maximum intensity for most of the folded residues corresponds to the highest temperature spectrum. However, under destabilizing conditions, apoSH mutants' spectra are collected in both folded and unfolding temperature range, we extend the Trp-32F max temperature correlating to the start of protein unfolding to each residue, where the max temperature reports on local unfolding. (B) Plots of max temperature for apoSH V148I without salt (blue) and with salt (red).

However, the apoSH V148I sample with salt shows a decrease in intensity in its temperature range, which aligns with protein unfolding. Hence, V148I without salt has a max temperature at 39.55 °C, which is the highest temperature a spectrum is acquired, and V148I with salt shows a max temperature at 34.23 °C, even though its highest temperature is at 38.64 °C (Figure 6.7B). When comparing the max temperature between both samples, it is important to acknowledge the differences in temperature when the spectra are acquired; careful consideration is taken to ensure all spectra are collected at the same set temperature, therefore the differences are small (<1.05 °C). Therefore, comparing max temperatures between the two samples is valid since the values are considered over multiple residues and spectra are collected at similar temperatures.

When comparing the max temperature between the two samples, it is expected the sample with a Trp-32F peak that shows protein unfolding, as seen with the salt sample, will show local unfolding at lower max temperatures. The average max temperature between the salt and salt-free V148I samples is 29.90 °C and 33.02 °C, respectively (Table 6.5). As predicted, V148I with salt has a lower average max temperature, suggesting it is experiencing local unfolding earlier than V148I without salt. Additionally, the perturbed areas identified in the intensity slope data also show the largest decreases in max temperature for the salt sample relative to the salt-free one (Figure 6.7B, 6.8 and Table 6.5). In both samples, loops 4 and 7 show the lowest max

temperature relative to the rest of their structures, but these large loops in the salt sample (23.02 °C and 21.07 °C, respectively) have considerably lower max temperatures than the salt-free sample (26.02 °C and 21.07 °C, respectively).

Table 6.5: Average max temperatures for apoSH V148I without (no salt) and with salt.

| | V148I no salt | V148I salt |
|---|---|---|
| average<br>standard dev | 33.02 (124)<br>8.79 | 29.90 (120)<br>8.65 |
| | | |
| N-terminal B-sheet | 38.67 (29) | 35.05 (29) |
| C-terminal B sheet | 36.44 (24) | 32.76 (25) |
| N | 16.69 (1) | - |
| B1 | 39.55 (6) | 37.01 (6) |
| loop I | 35.01 (3) | 31.05 (3) |
| B2 | 39.05 (9) | 33.24 (9) |
| loop II | 35.76 (3) | 31.81 (3) |
| B3 | 37.08 (8) | 35.56 (8) |
| loop III | 37.30 (4) | 30.76 (3) |
| B4 | 37.66 (6) | 36.05 (6) |
| loop IV | 26.55 (19) | 23.02 (17) |
| B5 | 39.55 (4) | 34.70 (5) |
| loop V | 38.99 (4) | 36.64 (4) |
| B6 | 38.42 (8) | 33.90 (8) |
| loop VI | 33.23 (13) | 32.30 (12) |
| B7 | 34.55 (5) | 30.52 (5) |
| loop VII | 24.46 (20) | 21.07 (20) |
| B8 | 34.00 (9) | 29.77 (9) |
| C | 39.55 (2) | 36.84 (2) |

Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

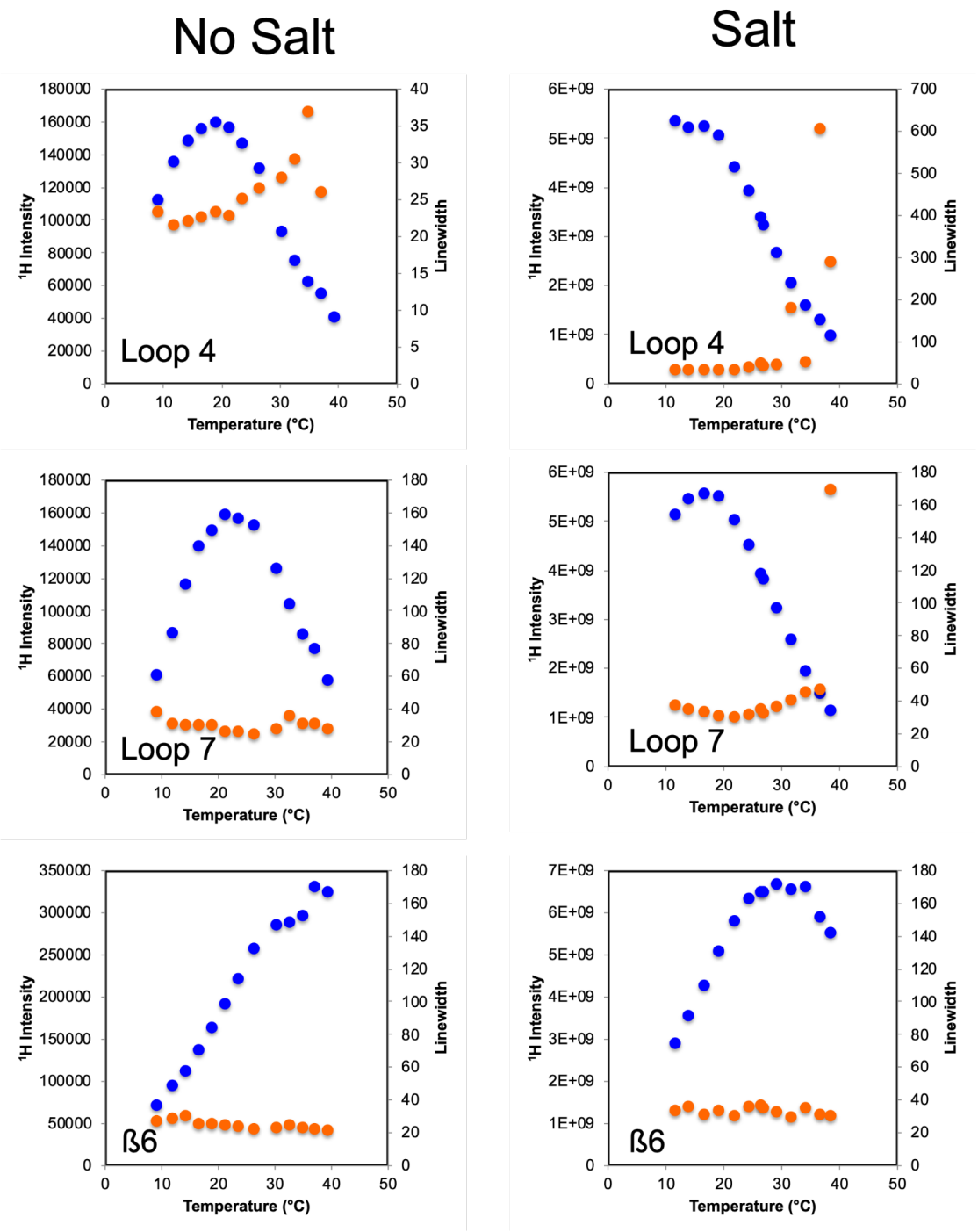



Figure 6.8: Peak profile for residues that behaved differently between the apoSH V148I without (left) and with salt (right). Representative residues are selected for structural regions that show differences in behaviour, and intensities (blue) and linewidths (orange) are overlaid to show differences between the two samples.

Moreover, ß4 and ß6-8 are also considered interesting points of differences between the salt-free and salt samples. All four beta-strands show decreases in max temperature, most considerably for ß6-8 for the salt sample relative to the salt-free sample (Figure 6.7B, 6.8 and Table 6.5). Since these strands have the same residues assigned, these differences are not due to differences in assignments. Both ß6 and ß7 temperature coefficients show increased dynamics in the presence of salt, and this decreased max temperature supports the salt is able to disrupt the structure in these strands and consequently leads to earlier local unfolding for these strands. However, ß4 does show a decrease in max temperature with the addition of salt, but its temperature coefficient data suggest it becomes less dynamic. Although initially perplexing, when assessing the ß4 max temperature relative to the rest of the secondary structures, it is clear in the salt sample, ß4 has one of the higher max temperatures, but in the salt-free sample, ß4 has one of the lowest max temperatures. Meanwhile, the relative max temperatures for ß6-8 did not change significantly between the two samples. This suggests ß4 becomes structured in the presence of salt, and therefore unfolds later relative to the rest of its structure. Overall, the salt samples start unfolding considerably earlier than the salt-free sample, with subtle shifts in local unfolding.

6.4.1.4 Secondary amide nitrogen temperature coefficients for perturbed ß-strands due to salt agree with proton amide temperature coefficient and intensity profile data

While proton amide chemical shifts are considered useful for probing hydrogen bonds and temperature-dependent loss of structure, nitrogen amide chemical shifts can serve as a measure of structure. In Chapter 5, we extensively discuss our approach to analyze and interpret nitrogen chemical shifts. By taking the difference between the

calculated nitrogen temperature coefficients for random coil SOD1 and the experimentally measured amide nitrogen temperature coefficients, denoted as the secondary amide nitrogen temperature coefficients, we can probe for deviations from random coil with independence from the amino acid sequence. In Chapter 5, apoSH pWT SOD1 and SOD1 mutants are probed for ordered and disordered protein regions, however, in this chapter, changes in ordered and disordered regions due to solution conditions are investigated.

The secondary nitrogen temperature coefficients for apoSH V148I with and without salt are compared to determine the impact salt has on apoSH SOD1 structure. Since random coil nitrogen temperature coefficients vary over a ~20 ppb/K range, values that are above 20 and below -20 ppb/K are considered different from random coil, or structured. Consistent with random coil, most of the secondary temperature coefficients for the loops and the C-terminus ß-sheet are close to zero, whereas the N-terminus ß-sheet values suggest it is structured for both salt and salt-free samples (Figure 6.9 and Table 6.6). Similar to the proton temperature coefficients, the secondary amide nitrogen temperature coefficient values throughout the protein are similar between the two samples.

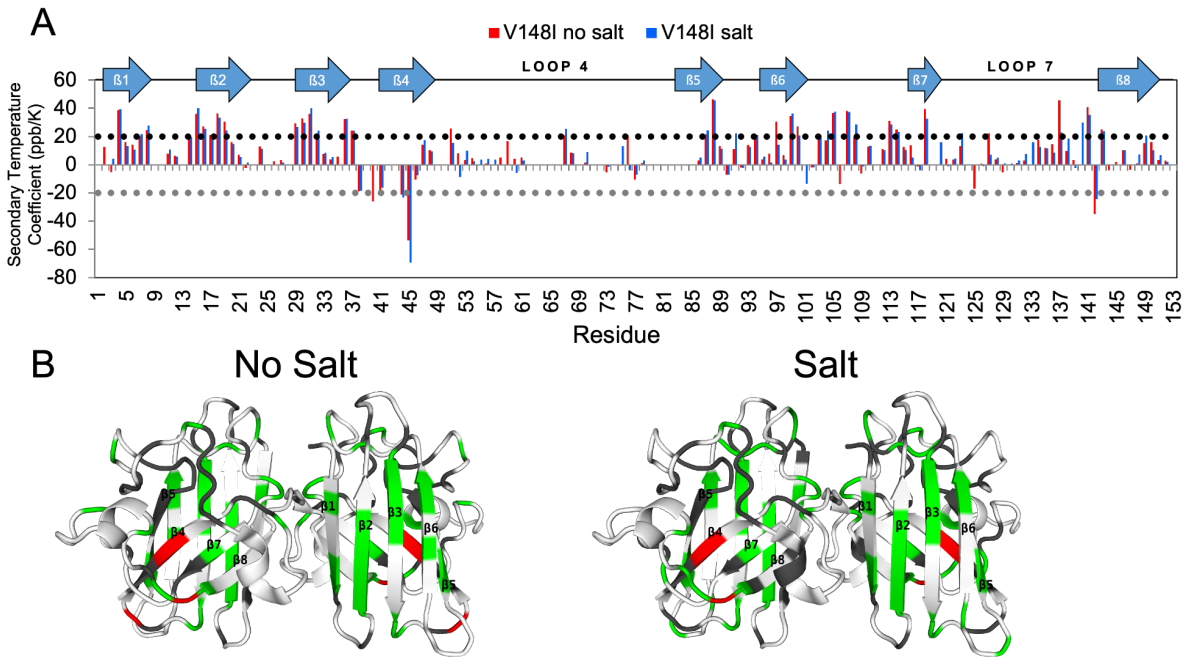



Figure 6.9: Amide nitrogen temperature coefficients for apoSH V148I without salt (no salt) and with salt. (A) Temperature coefficients are represented by the blue (no salt) and red (with salt) bars in the plot. (B) show 3D ribbon structures for the no salt (left) and with salt (right) data. Temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol) and coloured based on the +/-20 ppb/K cutoff. Residues that show values greater than 20 ppb/K and less than -20 ppb/K are coloured green and red, respectively. Residues within the -20 to 20 ppb/K range are coloured white. Grey indicates residues where no data are available.

Table 6.6: Amide nitrogen temperature coefficients for apoSH V148I without (no salt) and with salt.

| | V148I no salt | V148I salt |
|---|---|---|
| average<br>standard dev | -9.22 (121)<br>13.42 | -8.70 (115)<br>13.26 |
| | | |
| N-terminal B-sheet | 19.19 (31) | 18.00 (31) |
| C-terminal B sheet | 7.78 (23) | 7.49 (22) |
| N | 12.52 (1) | - |
| B1 | 18.09 (6) | 19.42 (6) |
| loop I | 4.80 (3) | 5.55 (3) |
| B2 | 21.25 (9) | 20.61 (9) |
| loop II | 6.11 (3) | 4.15 (3) |
| B3 | 21.19 (8) | 20.82 (8) |
| loop III | -6.88 (3) | 2.76 (2) |
| B4 | -13.30 (6) | -14.92 (6) |
| loop IV | 5.80 (19) | 4.32 (17) |
| B5 | 20.99 (4) | 21.62 (4) |
| loop V | 3.96 (4) | 6.33 (4) |
| B6 | 16.23 (8) | 11.17 (8) |
| loop VI | 15.61 (12) | 20.07 (11) |
| B7 | 16.13 (4) | 12.04 (5) |
| loop VII | 6.85 (20) | 8.83 (20) |
| B8 | 7.28 (9) | 11.21 (7) |
| C | 2.79 (2) | 2.07 (2) |

Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

Although the salt and salt-free samples share similar values for identical residues, there are subtle differences in averaging the secondary nitrogen temperature coefficients by SOD1s secondary structural elements. Interestingly the loop 7 that by proton temperature coefficients and intensity profile analysis show increased dynamics and decreased structure with the presence of salt, respectively, now by secondary amide temperature coefficients shows a small increase in value with the addition of salt (Figure 6.9 and Table 6.6). The salt-free and salt values for loop 7 are 6.85 and 8.83, respectively (Table 6.6). Even though the difference in the two values is small and does not consistently show large absolute values (>20 or <-20), it could suggest some residues are experiencing gradual structuring. Thus, the secondary nitrogen temperature

coefficient is useful for assessing overall ordered and disordered protein regions, but it may be less sensitive to smaller changes.

In contrast to the secondary nitrogen temperature coefficients of loop 7 deviating from the other data reported here, the perturbed ß-sheets do agree with their proton temperature coefficients and intensity profile data. Specifically, ß4 and ß6-7 show the largest perturbations according to both measures, and similarly, the secondary nitrogen temperature coefficients also show the largest differences in these strands between the two samples (Table 6.6). Amide proton temperature coefficients and the intensity profile data suggest ß4 is becoming more structured upon the addition of salt, but its neighboring strands ß6 and 7 become less structured. Secondary nitrogen temperature coefficients agree; the ß4 average secondary nitrogen temperature coefficient value moves further from zero, or further from its calculated random coil value, upon the addition of salt, and ß6 and 7 moves closer to zero, or closer to its calculated random coil value (Table 6.6).

Since the closer the value is to zero, the more consistent its value is to its random coil temperature coefficient, whereas the opposite suggests the value is less random coil-like. While ß4 and ß6-7 data agree with the other data reported above, loop 7 shows otherwise. A probable explanation may be due to the crowding of the residues along the indirect dimension, specifically for the unstructured loop (further discussed in Chapter 5). The crowded peaks increases assignment errors throughout the temperature range. Meanwhile, the ß-strands are more structured than loop 7, and subsequently their peaks are less crowded. Collectively, for more isolated peaks, the secondary amide nitrogen temperature coefficients for the perturbed ß-strands agree with the proton temperature

coefficient and intensity profile data, while more crowded peaks such as loop 7 residues show deviation from trends suggested by other data.

### 6.4.2: Protein concentration

Protein concentration is a key factor that impacts both protein folding and aggregation.[274–276] With increasing protein concentration, more consistently literature supports the proteins showing increased propensity to aggregate[276,277], and with some proteins showing contrary results.[12,278,279] The contrary results highlight the protein concentration range is important to understand; there is a range where protein folding is more favourable than protein aggregation, but past a certain threshold, protein aggregation is more favourable due to the enhanced chance of collision.[277,280] Additionally, our group has reported complex aggregation behaviour for apoSH SOD1 at high protein concentration.[12] Only some mutants show enhanced aggregation propensity at the high protein concentration, but others remained unchanged, suggesting that relative dynamics may play a critical role in protein folding and aggregation.

Previously the impact of protein concentration on apoSH SOD1 aggregation shows it does not exclusively result in increased aggregation[12] and rather, similar to the introduction of salt, the impact of protein concentration is dependent on mutation. With SOD1 having the ability to dimerize, aggregate and transiently unfold, increasing protein concentration may make some of these processes more favourable than others, which convolutes the interpretation further. Thus, by choosing a mutant that attenuates some of these processes, the impact of protein concentration can be better understood.

Fortunately, V148G is one of the few mutants that shows significantly perturbed dimerization in its immature states, consequently reducing the potential for contributions

from chances of dimerization in the high protein concentration sample. However, we cannot control the other two processes, structural changes and protein aggregation, as they are a likely consequence of changing protein concentration. Here, we investigate the differences in protein dynamics for V148G at two considerably different protein concentrations, 8.2 mg/mL (0.5 mM) and 24.0 mg/mL (1.5 mM), which will henceforth be referred to as the low or high protein concentration, respectively. Both samples show no presence of aggregates prior to heating which is assessed by DLS. Both the low concentration and high concentration samples have spectra collected throughout the same temperature ramp (1.0 - 36.7 °C and 1.2 - 36.0 °C, respectively) (Table 6.2). Only the folded temperature range is analyzed for the temperature coefficients, but the entire temperature range is considered for the peak intensity and linewidth analysis.

6.4.2.1 CSP values suggest the protein undergoes more native structuring at increased protein concentration

Although CSP are typically used for assessing chemical shift changes from a titration of a compound, as done for the V148I salt titration in section 6.4.1.1, here we extend this analysis to assess local structural changes arising from differences in protein concentration. In addition to assessing CSP values for both low and high concentration samples relative to pWT, the CSP values are also assessed relative to the two concentrations. Since V148G begins to unfold before 25 °C, the CSP values are calculated at 19 °C to ensure only the folded structures are considered. The CSP values relative to pWT indicate how native-like the samples are, and the CSP values relative to

each other show the structural variability between the two samples and in turn the impact protein concentration has on the local structure.

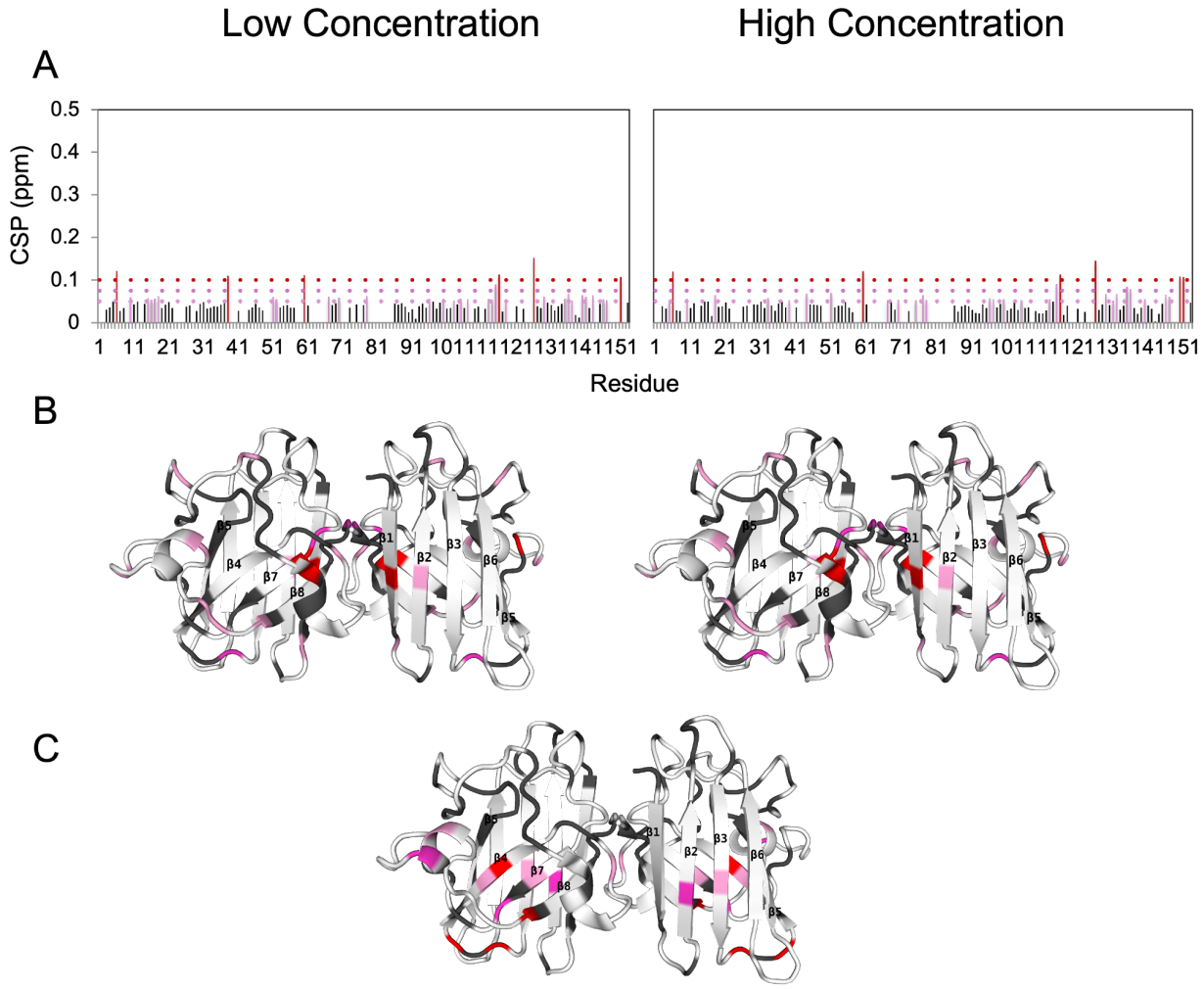



Figure 6.10: Chemical shift perturbations for V148G at low and high protein concentration. CSP values were coloured based on the same cutoff for both samples. Residues are coloured to aid in visualization according to the relative magnitude of CSP values. For (A) and (B) the pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1, respectively. Low concentration and high concentration refers to 0.5 mM and 1.5 mM protein concentration, respectively. (A) CSP values for V148G relative to pWT represented by the bars on the plot. (B) CSP values for V148G samples relative to pWT mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). (C) CSP values between the low and high protein concentration samples mapped onto a 3D SOD1 ribbon structure. White denotes CSP values smaller than average + 1 standard deviation, and grey indicates residues where no data are available. (B) CSP values are shown as vertical bars, where black bars are values that are smaller than average + 1 standard deviation.

CSP values show some differences relative to the native structure, and surprisingly the CSP values between the two samples also show subtle differences (Figure 6.10A, B).

For both the CSP values calculated relative to pWT and between the two concentrations, residues are considered perturbed if their CSP value is greater than the average + 1 standard deviation, otherwise are considered similar to the reference (pWT) or at both concentrations, respectively. For both concentrations, the perturbed CSP values relative to pWT are primarily located at the site of mutation and at the residues responsible for forming the native-helix in the electrostatic loop (loop 7) (Figure 6.10A, B). Although both concentrations show similar perturbed areas and standard deviation values, there are subtle differences in the residues perturbed. For instance, the dimer interface seems more perturbed, especially around the site of mutation, in the high concentration sample, whereas the low concentration sample shows more residues perturbed in the electrostatic loop.

When comparing the chemical shift values for the two concentrations against each other, the CSP values are smaller but highlight local structural differences (Figure 6.10C). Unsurprisingly the site of mutation is no longer perturbed, as both samples have the same mutation, and rather the significant sites of perturbation are concentrated around neighbouring residues in the beta-barrel and residues involved in forming the native-helix in the electrostatic loop. From the 11 residues that show significant differences, 6 of these residues are expected to be hydrogen bonded in the mature structure and 2 other residues are involved in forming the native-helix in the mature enzyme.[221] Since most of the perturbed residues located in/near the beta-strands are expected to be hydrogen bonded or involved in structuring upon maturation, the CSP value differences may be a result of differences in structuring.

Increasing the protein concentration for apoSH V148G may not exclusively be important for aggregation, but also protein folding. Most of the residues that show perturbed CSP values between the two concentrations are involved in hydrogen bonding, metal-binding and/or structuring upon maturation. However, Chapter 3 temperature coefficient data, Chapter 7 thermograms, and in vivo aggregation experiments[25] support apoSH V148G having perturbed metal-binding due to enhanced dynamics around the metal-binding region. Additionally, Sekhar *et al*. have shown that apoSH mutants can transiently access structured states such as the native-helix that forms upon maturation, and the transient states are more readily accessible at higher protein concentration.[55] Collectively, upon increasing protein concentration, apoSH V148G is structuring around the previously dynamic metal-binding region. Thus, apoSH V148G CSP values suggest increasing the protein concentration leads to local protein structuring.

6.4.2.2: Temperature coefficients highlight only minimal differences in samples at different protein concentration

Since the CSP values suggest the perturbations happen across beta strands but are limited to only one or two residues in each of the respective strands. Since the changes are more localized to a few residues, it is likely the differences in temperature coefficients between the low and high concentration samples are best probed by the rolling averages or individual values. However, it is also worth remembering the residues with perturbed CSP values do not necessarily coincide with differences in temperature coefficients as observed for apoSH V148I with and without salt. Again, to minimize the number of processes occurring, only the folded temperature range is considered for both

samples. Therefore, all the different averaging methods for the folded temperature coefficients are considered for apoSH V148G samples.

By comparing the temperature coefficient values across the protein for the two samples, smaller local regions show differences. Although there are individual residues that show considerable differences between the two concentrations, only differences across multiple residues are considered significant when assessing the temperature coefficient profile. With those considerations, loop 4 is interesting as it shows residues at the start of the loop are more dynamic in the low concentration sample but become less dynamic in the residues at the end of the loop relative to the high concentration sample (Figure 6.11A). Additionally, loop 7 and ß8 also show both increases and decreases in temperature coefficient values interspersed throughout these structural regions. Thus, as expected, the differences in temperature coefficients by varying the protein concentration are minimal.

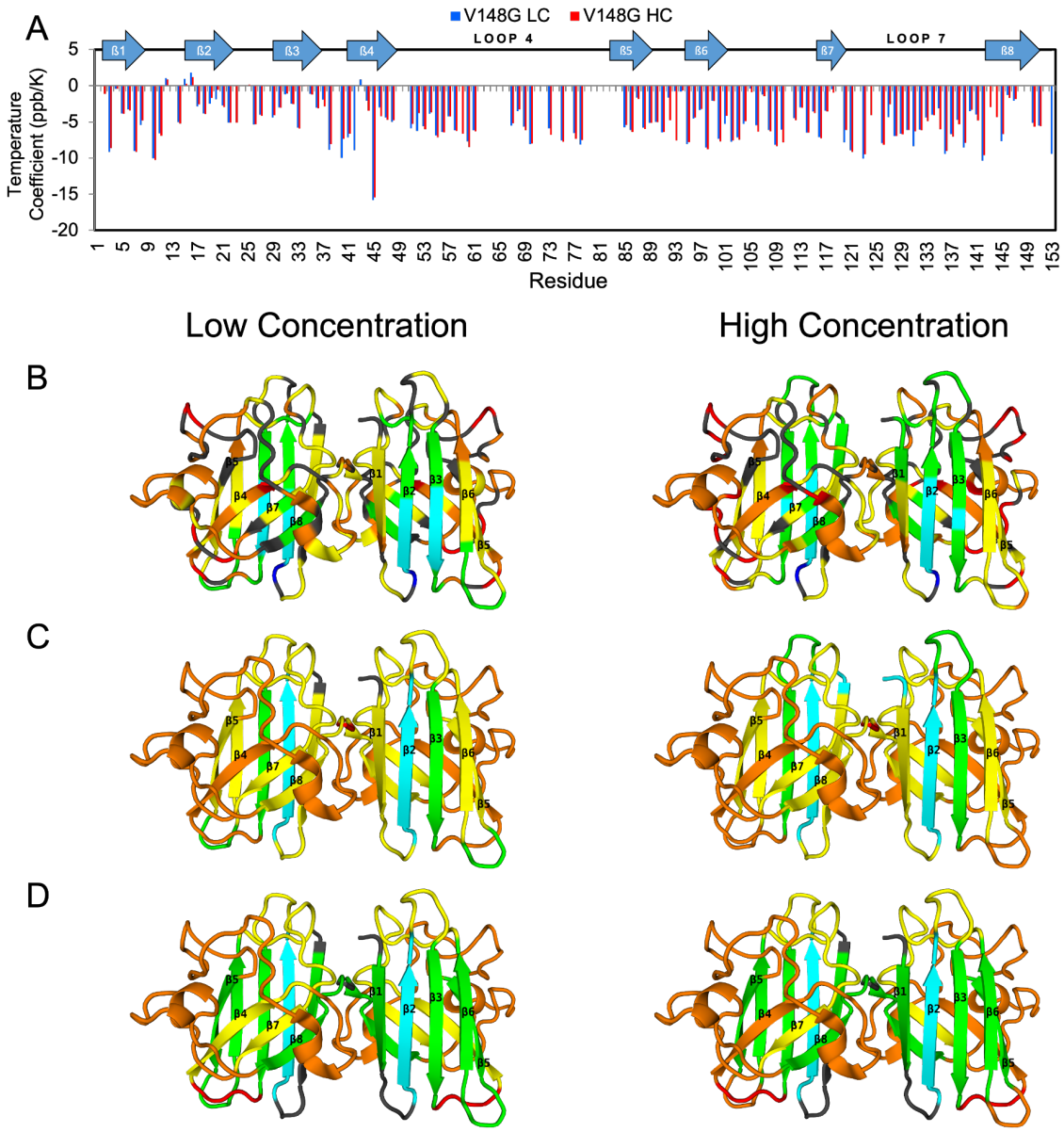



Figure 6.11: Amide proton temperature coefficients and their multiple averaging methods for apoSH V148G at low (LC) and high (HC) protein concentration. (A) Temperature coefficients are represented by the blue (low concentration) and red (high concentration) bars in the plot. (B-D) show 3D ribbon structures for the low protein concentration (left) and high protein concentration (right) data. (B) Rolling average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). (C) Secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5). (D) Averaging only hydrogen bonded residues, secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5). Residues are coloured by a red to blue gradient, with red representing less stable residues and blue for the more stable. Dark blue is for values larger than -1 ppb/K, cyan is for values between -1 and -2.5 ppb/K, green is for values between -2.5 and -4 ppb/K, yellow represents values between -4 to 5.5 ppb/K, orange is for values between -5.5 to -7 ppb/K and values smaller than -7 ppb/K are coloured red. Grey indicates residues where no data are available.

The different averaging methods further support only small, local differences in the structures from varying the protein concentration. The rolling averages highlight the differences are very small and mostly localized to parts of ß4 and the residues responsible for forming the native-helix in loop 7 (Figure 6.11B). Although other areas show differences in Figure 6.11B, the number of residues assigned is different, and thus not considered as truly different. Moreover, averaging the residues by secondary structure or only hydrogen bonding residues by their secondary reveal no significant differences at the two protein concentrations (Figure 6.11C, D, Table 6.7). This aligns with the averages for all assigned residues for the two samples showing negligible differences (-5.10 ppb/K and -5.06 ppb/K for the low and high concentration samples, respectively) (Table 6.7). Overall, similar to the chemical shift perturbations, the temperature coefficient values only show small local differences.

Table 6.7: Amide proton temperature coefficients for apoSH V148G at low protein concentration (LC) and high protein concentration (HC).

| | V148G LC | V148G HC |
|---|---|---|
| average | -5.10 (116) | -5.06 (122) |
| standard dev | 2.93 | 2.75 |
| average ddTC | -0.45 (111) | -0.54 (114) |
| | | |
| N-terminal B-sheet | -3.79 (31) | -3.70 (31) |
| C-terminal B sheet | -4.84 (23) | -4.75 (24) |
| N | - | -1.14 (1) |
| B1 | -5.18 (6) | -5.04 (6) |
| loop I | -5.17 (3) | -5.42 (3) |
| B2 | -2.35 (9) | -2.29 (9) |
| loop II | -4.70 (2) | -3.60 (4) |
| B3 | -2.62 (8) | -2.62 (8) |
| loop III | -6.91 (3) | -6.08 (3) |
| B4 | -5.68 (8) | -5.59 (7) |
| loop IV | -6.10 (18) | -6.12 (19) |
| B5 | -4.87 (5) | -4.96 (5) |
| loop V | -3.46 (4) | -5.94 (4) |
| B6 | -4.99 (8) | -4.85 (8) |
| loop VI | -5.13 (12) | -5.45 (12) |
| B7 | -4.48 (5) | -4.33 (5) |
| loop VII | -6.54 (19) | -6.38 (20) |
| B8 | -4.33 (5) | -4.10 (7) |
| C | -9.39 (1) | -9.34 (1) |

Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets. Average ddTC is the difference in average temperature coefficients relative to apoSH pWT.

#### 6.4.2.3 Intensity analysis

##### 6.4.2.3.1 Trp-32F and Trp-32U peaks show no discernible differences between the two protein concentrations

Since in section 6.4.1.3.1, we observe the presence of salt alters the folded-unfolded equilibrium, understanding how protein concentration impacts this equilibrium is also important to examine. For both apoSH V148G low and high concentration samples, all spectra have a peak for Trp-32F and Trp-32U, even in temperatures corresponding to folded temperatures for V148G.[65] The presence of Trp-32U is present in all spectra for destabilized mutants (A4V, G85R, E100G, and V148G), whereas the stabilized SOD1s

(pWT and V148I) do not have this peak in their folded temperature ranges, further discussed in Chapter 4. However, unlike V148I showing changes in Trp-32F and Trp-32U intensity upon solution condition change (Figure 6.4A), V148G's Trp-32F and Trp-32U peaks show negligible changes between the two concentrations (Figure 6.12). Therefore, the presence of any differences in intensity between the samples will exclusively report differences in dynamics due to the increased protein concentration, and not additional processes like protein unfolding.

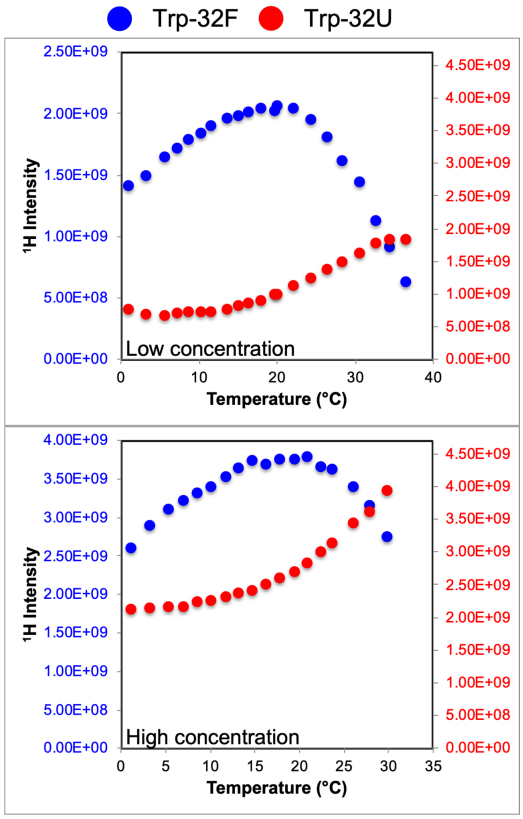



Figure 6.12: Trp-32F and Trp-32U peak intensities throughout the temperature ramp. All peak intensities were normalized by receiver gain. The Trp-32F (blue) and Trp-32U (red) peaks are evaluated throughout the temperature ramp. The Trp-32F and Trp-32U peak intensity for the apoSH V148G samples at low protein concentration (top) and high protein concentration (bottom). Throughout the folded temperature range, the TrpF-32F peak intensity increases, whereas the Trp-32U peak shows minimal changes. However, the Trp-32F and Trp-32U intensity begins to decrease and increase, respectively, when the protein is in the unfolding temperature ranges.

6.4.2.3.2 The loops intensity slopes show differences at high and low protein concentration

Since the temperature coefficients show minimal differences between apoSH V148G at low and high concentrations, it is expected the post-normalized folded intensity slopes are likely to show minimal differences as well. Although both samples have spectra collected in the folded and unfolding temperature ranges, only the folded temperature ranges are reported here. Akin to the temperature coefficient data, the folded intensity slopes values are comparable between the low and high protein concentration samples, with certain regions showing differences (Figure 6.13). Interestingly the regions of intensity slope differences are also areas that show differences in temperature coefficients, but the differences are more considerable for the intensity slopes.

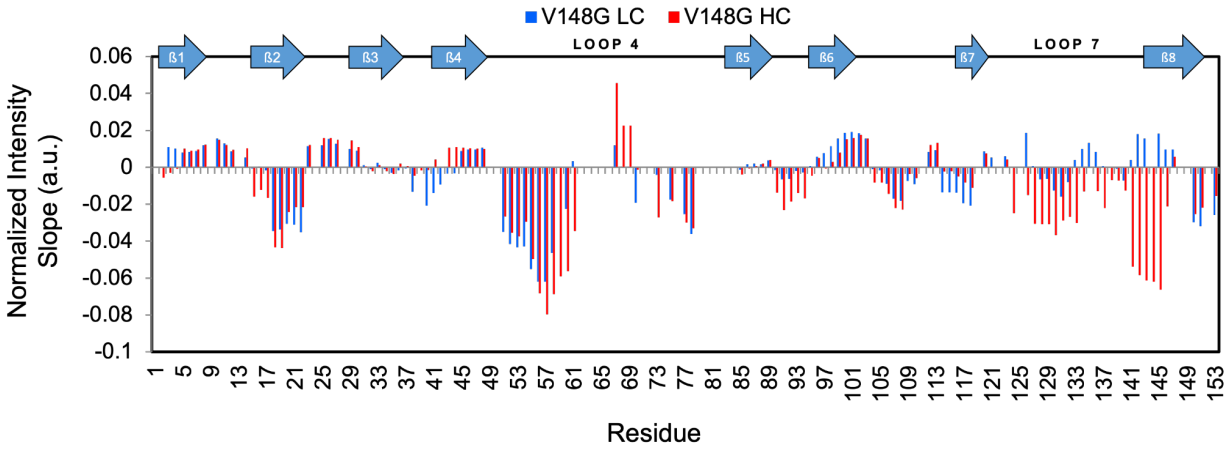



Figure 6.13: Post-normalized folded intensity slopes for apoSH V148G at low protein concentration (LC) and high protein concentration (HC). The rolling averages for normalized intensity slopes are considered for the folded temperature ranges. The rolling averages for normalized intensity slopes are for the sample at low protein concentration (blue) and high protein concentration (red). The closer the values are to 0, the residue responds to temperature changes comparable to the Trp-32F peak that the data are normalized by.

Table 6.8: Average normalized intensity slopes for apoSH V148G at low protein concentration (LC) and high protein concentration (HC).

| | V148G LC | V148G HC |
|---|---|---|
| average<br>standard dev | -0.0045 (119)<br>0.01807 | -0.0109 (123)<br>0.02297 |
| | | |
| N-terminal B-sheet | 0.0006 (31) | -0.0005 (30) |
| C-terminal B sheet | -0.0026 (24) | -0.0074 (24) |
| N | - | -0.0057 (1) |
| B1 | 0.0098 (6) | 0.0063 (6) |
| loop I | 0.0124 (3) | 0.0122 (3) |
| B2 | -0.0180 (9) | -0.0210 (9) |
| loop II | 0.0129 (4) | 0.0148 (4) |
| B3 | 0.0020 (8) | 0.0027 (8) |
| loop III | -0.0116 (3) | -0.0017 (4) |
| B4 | 0.0011 (8) | 0.0095 (7) |
| loop IV | -0.0276 (18) | -0.0297 (19) |
| B5 | 0.0016 (5) | 0.0005 (5) |
| loop V | -0.0040 (4) | -0.0173 (4) |
| B6 | 0.0096 (8) | 0.0031 (8) |
| loop VI | -0.0021 (12) | -0.0024 (12) |
| B7 | -0.0118 (5) | -0.0038 (5) |
| loop VII | 0.0017 (19) | -0.0219 (20) |
| B8 | -0.0014 (6) | -0.0360 (7) |
| C | -0.0259 (1) | -0.0155 (1) |

Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

Moreover, the intensity slope differences between the low and high concentration samples are predominantly in the large loops, 4 and 7 (Figure 6.13 and Table 6.8). For loop 4, residues 51-56 intensity slope values are larger in the low concentration sample, but in residues from 57-61, the values are larger in the high concentration sample. Interestingly, loop 4 also shows the earlier residues have an increase in structural dynamics and the later residues are more structured in the low concentration sample relative to the high concentration sample. Therefore, both temperature coefficient data and intensity slope data support the initial parts of loop 4 are less structured in the low

concentration sample, whereas the residues in the middle of loop 4 are more structured in the low concentration sample.

Additionally, loop 7 consistently shows large negative values for the high concentration sample, whereas the low concentration shows a mixture of small positive and negative values throughout the loop (Figure 6.13). Hence, loop 7 is less folded in the high concentration sample. Interestingly, loop 7 has large increases in values at the start and ends of the loop, but for the residues near the native-helix that forms in the mature enzyme, the values become smaller, suggesting there is structuring in this region. Interestingly, Sekhar *et al*. have reported apoSH A4V can transiently access the native-helix at high protein concentrations, that at low protein concentrations are not observable.[55] Therefore, at higher protein concentrations, apoSH V148G becomes more structured in regions that become structured upon maturation, while other areas of the large loops are more dynamic relative to the low concentration sample.

6.4.2.3.3 Due to apoSH V148G predominately existing as a monomer, its local unfolding remains unchanged at low and high protein concentrations

Even though apoSH V148G has minimal differences between the low and high concentration samples, the small differences may yet show differences in local unfolding. Similar to their temperature coefficient and intensity slope data, the differences in max temperature are minimal (Figure 6.14 and Table 6.9). The average max temperatures are similar, 19.21 °C for the low concentration sample and 18.75 °C for the high concentration sample (Table 6.9). Additionally, all the secondary structural elements also show comparable temperatures at the two concentrations (Table 6.9). With similar max

temperatures, apoSH V148G likely unfolds the same at both concentrations, and that is likely due to its perturbed dimerization. For a mutant that dimerizes readily, the unfolding, especially around the dimer interface, is likely to change. Overall, the minimal differences in max temperature between the high and low samples of apoSH V148G suggest the protein unfolds the same irrespective of protein concentration.

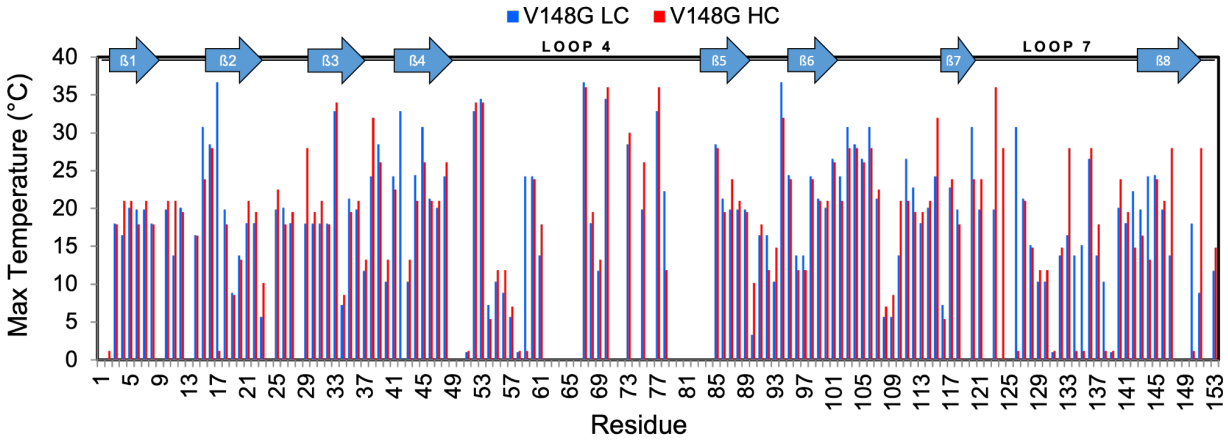



Figure 6.14: Plots of max temperature for apoSH V148G at low protein concentration (blue) and high protein concentration (red).

| | V148G LC | V148G HC |
|---|---|---|
| average | 19.21 (123) | 18.75 (124) |
| standard dev | 8.16 | 8.86 |
| | | |
| N-terminal B-sheet | 21.53 (31) | 19.79 (30) |
| C-terminal B sheet | 21.21 (25) | 20.85 (24) |
| N | - | 1.20 (1) |
| B1 | 18.73 (6) | 19.49 (6) |
| loop I | 17.93 (3) | 20.55 (3) |
| B2 | 21.23 (9) | 16.65 (9) |
| loop II | 15.93 (4) | 17.54 (4) |
| B3 | 19.17 (8) | 21.22 (8) |
| loop III | 18.70 (4) | 21.14 (4) |
| B4 | 23.53 (8) | 21.59 (7) |
| loop IV | 19.39 (19) | 18.87 (19) |
| B5 | 21.89 (5) | 22.40 (5) |
| loop V | 11.67 (4) | 13.71 (4) |
| B6 | 22.62 (8) | 21.45 (8) |
| loop VI | 21.14 (13) | 20.89 (13) |
| B7 | 20.98 (5) | 20.61 (5) |
| loop VII | 15.80 (19) | 15.09 (20) |
| B8 | 18.44 (7) | 18.83 (7) |
| C | 11.74 (1) | 14.84 (1) |

Table 6.9: Average max temperatures for apoSH V148G at low protein concentration (LC) and high protein concentration (HC). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

#### 6.4.2.4 ApoSH V148G at low and high protein concentrations have comparable structure based on their comparable secondary nitrogen temperature coefficients

Since there are minimal observable changes in the temperature coefficient and intensity profile data, it is unsurprising that the secondary nitrogen temperature coefficients show no significant differences (Figure 6.15 and Table 6.10). Instead, the high concentration apoSH V148G has similar values and therefore similar ordered and disordered regions to the low concentration apoSH V148G that was discussed in Chapter 5. Conclusions are only drawn when multiple residues are showing comparable behaviour; there are some residues that show outlier behaviour, but due to the significant crowding of peaks in the indirect dimensions, these are not considered meaningful.

Overall, the secondary nitrogen temperature coefficients suggest the structure is largely similar under both protein concentrations.

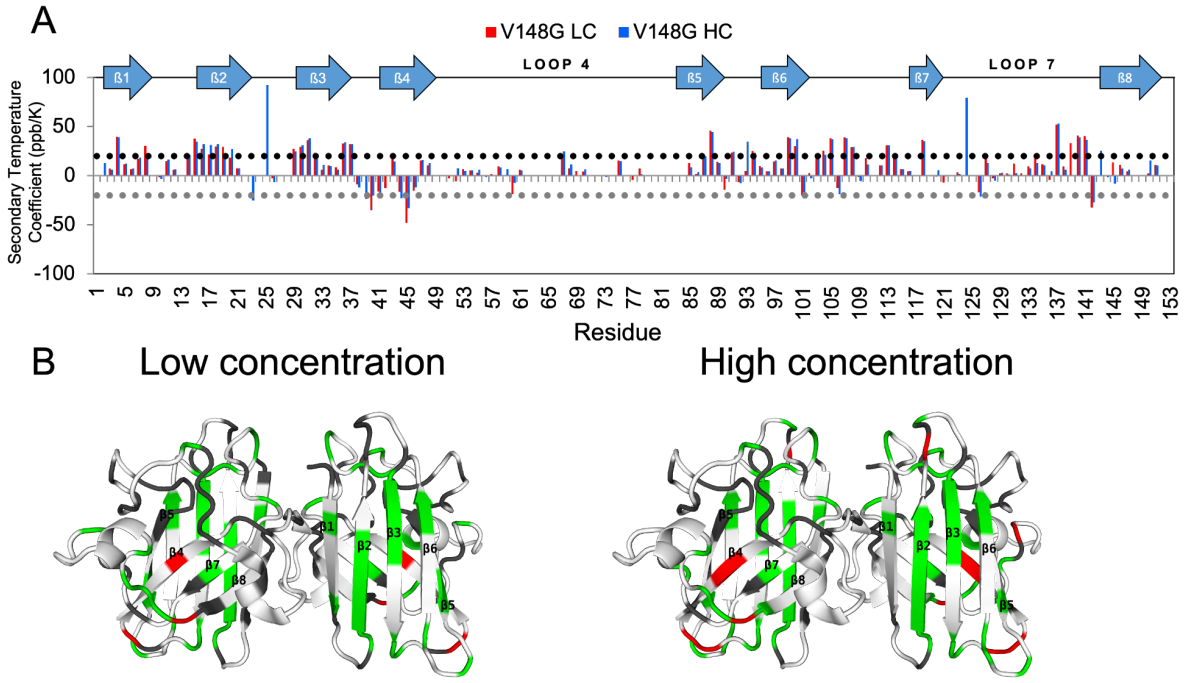



Figure 6.15: Amide nitrogen temperature coefficients for apoSH V148G at low (LC) and high (HC) protein concentration. (A) Temperature coefficients are represented by the blue (low concentration) and red (high concentration) bars in the plot. (B) show 3D ribbon structures for the no salt (left) and with salt (right) data. Temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol) and coloured based on the +/-20 ppb/K cutoff. Residues that show values greater than 20 ppb/K and less than -20 ppb/K are coloured green and red, respectively. Residues within the -20 to 20 ppb/K range are coloured white. Grey indicates residues where no data are available.

| | V148G LC | V148G HC |
|---|---|---|
| average<br>standard dev | -8.58 (116)<br>14.87 | -7.92 (123)<br>17.46 |
| | | |
| N-terminal B-sheet | 18.92 (31) | 19.09 (31) |
| C-terminal B sheet | 7.29 (23) | 7.56 (24) |
| N | - | 12.97 (1) |
| B1 | 18.81 (6) | 16.71 (6) |
| loop I | 6.48 (3) | 6.55 (3) |
| B2 | 21.33 (9) | 23.05 (9) |
| loop II | -0.96 (2) | 15.30 (4) |
| B3 | 21.73 (8) | 22.05 (8) |
| loop III | -4.06 (3) | -6.02 (4) |
| B4 | -7.96 (8) | -5.96 (7) |
| loop IV | 3.32 (18) | 4.91 (19) |
| B5 | 19.25 (5) | 17.89 (5) |
| loop V | 1.91 (4) | 12.11 (4) |
| B6 | 13.83 (8) | 14.56 (8) |
| loop VI | 17.98 (12) | 15.67 (12) |
| B7 | 9.48 (5) | 10.49 (5) |
| loop VII | 10.24 (19) | 10.81 (20) |
| B8 | 8.39 (5) | 7.81 (7) |
| C | - | - |

Table 6.10: Amide nitrogen temperature coefficients for apoSH V148G at low protein concentration (LC) and high protein concentration (HC). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

### 6.4.3: Protein aggregation

Protein aggregation and pre-aggregated samples are a likely consequence when studying proteins especially at concentrations necessary for protein NMR. Typically, scientists attempt to minimize the number of observable processes, especially for structural analysis. Clever scientists have developed complex NMR experiments to study larger proteins[76,82,281], and other equally clever scientists have used simple 2D HSQC experiments to systematically assess protein aggregation.[57,230,282] Although these experiments offer more insight into protein folding and misfolding, aggregation initiation remains predominantly elusive. Additionally, with the structural polymorphisms observed for apoSH SOD1, these different conformational “strains” may be analogous to the prion-

like aggregation reported previously.[28,29] Thus, understanding the protein aggregation and the role protein aggregates have on soluble protein is vital.

Here, we explore the influence the presence of aggregates has on both the NMR peak shape (chemical shifts, intensity, and linewidths) and soluble monomer structure throughout a temperature range. Since apoSH mutants show the ability to both natively and non-natively dimerize, unfold and/or aggregate, the mutant of choice and experimental parameters are chosen to minimize these additional processes while ensuring the sample readily aggregates. For these reasons, A4V is an ideal candidate since this mutant has significantly perturbed dimerization in immature states[87] (see Chapter 7), and its apoSH form readily aggregates under most conditions.[12]

This experiment allows for systematic assessment of the presence of aggregates through numerous ways. To ensure the protein is not experiencing any local unfolding, the spectra are acquired in a temperature range where the protein is folded (3.67 - 17.86 °C and 5.65 - 17.94 °C for the samples with and without aggregates, respectively) (Table 6.2).[65] Additionally, the sample with the aggregates (aggregated sample) present is compared to a sample of A4V without the presence of aggregates (unaggregated sample), which is diagnosed by DLS throughout the entire temperature range. Both samples are under identical conditions in terms of pH, buffer conditions and protein concentration (1.1 mM for the aggregated sample and 1.0 mM for the unaggregated sample). Although viewed as mostly a limitation, protein aggregates in the presence of soluble protein can be insightful in the influence of aggregates on NMR data as well as understanding protein aggregation mechanisms.

6.4.3.1 CSP values between the samples with and without the presence of aggregates show subtle differences on C-terminus face

In Chapter 3, CSP values for apoSH A4V show that it does not have a native-like structure. Similarly, in the presence of aggregates, A4V remains a less native-like fold, hence comparing the values to pWT does not offer interpretable data relative to the two samples (Figure 6.16B). Thus, CSP values are calculated between the two samples, the unaggregated and the aggregated sample (Figure 6.16C). Considering apoSH A4V begins to unfold at 25 °C, the CSP values for both samples are calculated at 19 °C. Residues are considered perturbed if their value is larger than the average of all the CSP values + 1 standard deviation, otherwise are not considered similar to each other (Figure 6.16).

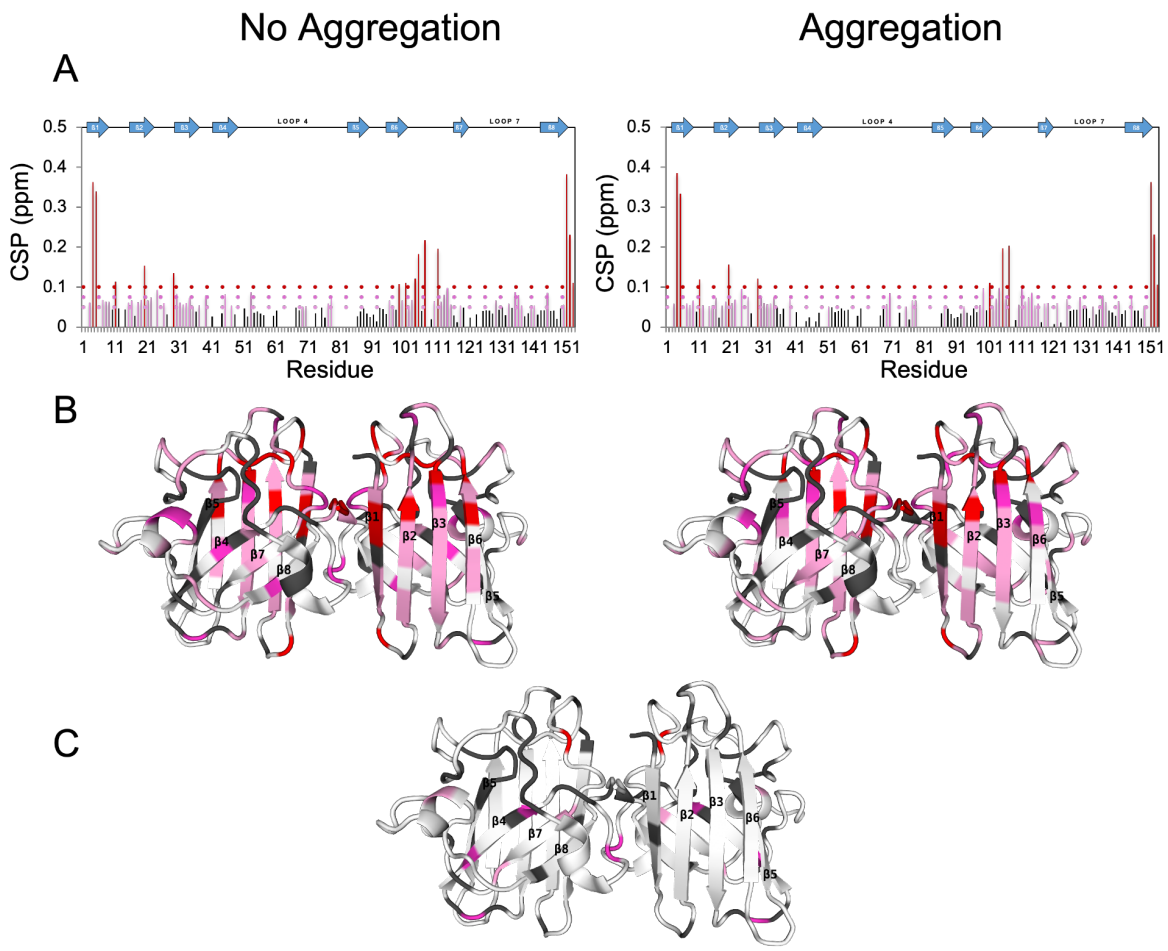



Figure 6.16 Chemical shift perturbations for A4V without aggregates and with aggregates. CSP values were coloured based on the same cutoff for both samples. Residues are coloured to aid in visualization according to the relative magnitude of CSP values. For (A) and (B) the pink, dark pink and red colours/lines represent 0.05–0.075, 0.075–0.1, >0.1, respectively. No aggregation and aggregation refer to A4V samples without and with presence of aggregates, respectively. (A) CSP values for A4V relative to pWT represented by the bars on the plot. (B) CSP values for A4V samples relative to pWT mapped onto the 3D ribbon structure for holoSS SOD1 (1HL5, rendered with PyMol). (C) CSP values between A4V samples with and without aggregates mapped onto a 3D SOD1 ribbon structure. White denotes CSP values smaller than average + 1 standard deviation, and grey indicates residues where no data are available. (B) CSP values are shown as vertical bars, where black bars are values that are smaller than average + 1 standard deviation.

By comparing the two samples, perturbed CSP values can reflect differences in chemical shifts due to changes in the local environment and/or the states being accessed by the residue. For instance, the aggregated sample may have some residues that in addition to the folded state(s) they access, they may also be accessing a transient state that interacts with the pre-formed aggregates. Interestingly, the 8 perturbed residues are

dispersed throughout the large loops (loops 4 and 7) and the ends of the beta-strands on the C-terminus beta-sheet (Figure 6.16A, C). Interestingly, 6 of the 8 residues are known to hydrogen bond in the mature enzyme[227], and previously in Chapter 3, the C-terminus beta-sheet for apoSH SOD1 is the least structured. The dynamic C-terminus and large loops suggest without stabilizing maturation events, these residues may be susceptible to protein association. Ultimately, the extent of perturbation is minimal relative to the other two solution conditions, but these changes are appear significant and suggest CSP values can be a reporter of residues impacted by the presence of aggregates.

6.4.3.2: The larger loops (4, 6 and 7) are more dynamic in presence of aggregates

Amide proton temperature coefficient values between the unaggregated and aggregated sample can highlight local areas that may be less structured due to transient self-association. Any observable differences between the aggregated and unaggregated apoSH A4V samples are complex to interpret. Simplistically, we can assume the folded soluble monomer in both samples can access comparable states, however, the folded monomer in the aggregated sample may now have additional states and shifts in equilibria for other states due to the presence of aggregates. With these changes in transient states, subsequent structural changes are likely to manifest in differences in temperature coefficients.

The temperature coefficients show differences between the aggregated and unaggregated apoSH A4V samples. By assessing the temperature coefficient values across the structure for both samples, residues in the earlier parts of loop 4, 6 and 7, the larger loops in the SOD1 structure, show increased dynamics in the aggregated samples (Figure 6.17A). Additionally, near the site of mutation, parts of ß1 and ß2 also show

increased dynamics in the aggregated sample. Otherwise, most of the temperature coefficient profile suggests the aggregated sample is more dynamic throughout the structure.

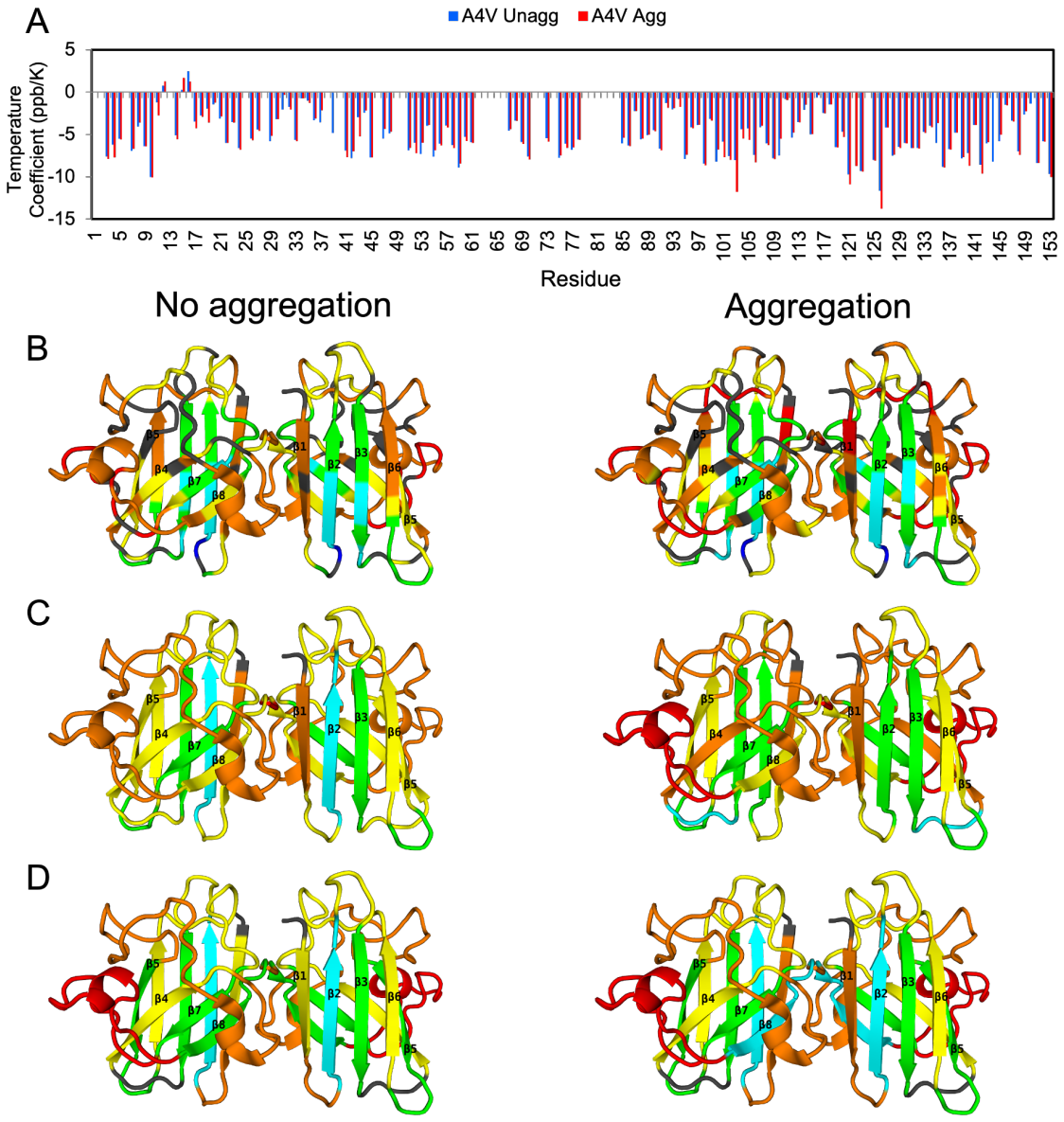



Figure 6.17: Amide proton temperature coefficients and their multiple averaging methods for apoSH A4V without aggregates (no aggregation) and with aggregates (aggregation). (A) Temperature coefficients are represented by the blue (without aggregates, unagg) and red (with aggregates, agg) bars in the plot. (B-D) show 3D ribbon structures for the without aggregates (left) and with aggregates (right) data. (B) Rolling average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5, rendered with PyMol). (C) Secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5). (D) Averaging only hydrogen bonded residues, secondary structural average temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5). Residues are coloured

by a red to blue gradient, with red representing less stable residues and blue for the more stable. Dark blue is for values larger than -1 ppb/K, cyan is for values between -1 and -2.5 ppb/K, green is for values between -2.5 and -4 ppb/K, yellow represents values between -4 to 5.5 ppb/K, orange is for values between -5.5 to -7 ppb/K and values smaller than -7 ppb/K are coloured red. Grey indicates residues where no data are available.

Interestingly, averaging temperature coefficients reveal the major differences appear to be in the large loops and not in the ß-strands. From averaging all the residues, the unaggregated sample has a larger value than the unaggregated sample, -5.31 ppb/K and -5.16 ppb/K, respectively (Table 6.11). Remarkably, when averaging the N-terminus and C-terminus ß-sheets, the C-terminus face is more dynamic than the N-terminus face, but both ß-sheets show negligible difference across the two samples (-4.16 and -4.71 ppb/K, for the unaggregated sample and -4.27 and -4.71 ppb/K for the aggregated sample, respectively) (Table 6.11). This suggests the increased average temperature coefficient for the aggregated sample is not from differences in the ß-strands, but from the loops.

| | A4V unagg | A4V agg |
|---|---|---|
| **average** | -5.16 (126) | -5.31 (124) |
| **standard dev** | 2.53 | 2.69 |
| **average ddTC** | -0.30 (120) | -0.50 (118) |
| | | |
| **N-terminal B-sheet** | -4.16 (30) | -4.27 (30) |
| **C-terminal B sheet** | -4.71 (27) | -4.71 (25) |
| **N** | - | - |
| **B1** | -6.06 (5) | -6.30 (5) |
| **loop I** | -4.21 (4) | -4.49 (4) |
| **B2** | -2.35 (9) | -2.62 (9) |
| **loop II** | -5.00 (4) | -5.18 (4) |
| **B3** | -2.93 (8) | -2.70 (8) |
| **loop III** | -4.22 (2) | -2.15 (1) |
| **B4** | -5.45 (7) | -5.55 (7) |
| **loop IV** | -6.07 (20) | -6.09 (20) |
| **B5** | -5.03 (5) | -4.91 (5) |
| **loop V** | -3.62 (4) | -3.80 (4) |
| **B6** | -5.30 (8) | -5.47 (8) |
| **loop VI** | -5.30 (13) | -5.66 (13) |
| **B7** | -3.44 (6) | -3.52 (6) |
| **loop VII** | -6.93 (20) | -7.21 (21) |
| **B8** | -4.93 (9) | -4.87 (7) |
| **C** | -7.74 (2) | -7.96 (2) |

Table 6.11: Amide proton temperature coefficients for apoSH A4V with the absence of aggregates (unagg) and the presence of aggregates (agg). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets. Average ddTC is the difference in average temperature coefficients relative to apoSH pWT.

Furthermore, the different local averaging methods reflect smaller areas of differences between the two samples. The rolling averages show that aside from the earlier parts of ß1, ß4, loop 4, 6 and 7 being more dynamic in the aggregated sample, ß8 is less dynamic in the aggregated sample (Figure 6.17B). Unsurprisingly, since some of these differences are localized to only a few residues relative to their respective secondary structural element, the differences are no longer present when averaging the temperature coefficients by their secondary structure. Instead, the secondary structure based averaging shows only ß2, ß4, loop 6 and loop 7 are dynamic in the aggregated sample (Figure 6.17C and Table 6.11). These changes do not seem to be a result of changes in hydrogen bonding since the secondary structural averages of only hydrogen bonded residues revealed on ß1 and ß8 show changes (Figure 6.17D).

The aggregated apoSH A4V sample shows more dynamics throughout the structure relative to the unaggregated sample. The increased dynamics are primarily localized to the larger loops, 4, 6 and 7, with some differences in ß1, ß4 and ß8. Since apoSH SOD1 exists predominantly as a monomer[65], and A4V shows perturbed dimerization in most immature states[87] (see Chapter 7), the periphery of the protein include the edge strands and the dynamic large loops. The periphery of the protein, especially the loops, may be more dynamic due to the influence of the aggregates present, as they may be readily interacting with the aggregates. Therefore, in the presence of the aggregates, the increased dynamics at the periphery of the A4V monomer may be suggesting priming of these residues prior to self-association.

6.4.3.3 Peak intensity analysis

6.4.3.3.1 Trp-32U intensity behaviour suggests the unfolded species actively aggregates in the presence of aggregates

Trp-32F and Trp-32U intensity behaviour have correlated well with their respective apoSH SOD1 folded and unfolding temperature ranges (see Chapter 4). Accordingly, the Trp-32F and Trp-32U peaks in the unaggregated sample show intensity decreases and increases, respectively, at the first unfolding temperature (~18 °C) (Figure 6.18). However, for the apoSH A4V aggregated sample, the Trp-32F and Trp-32U have differences relative to the unaggregated sample, and consequently these peaks likely report on other processes (Figure 6.18). For Trp-32F in the aggregated sample, the peak intensity starts decreasing at an earlier temperature (~16 °C) than expected for apoSH A4V (~18 °C) and declines more in intensity in the unfolding temperature range relative to the unaggregated sample. Therefore, Trp-32F intensity increases and decreases are largely consistent with the mutants folded and unfolding temperature range, respectively.

Moreover, this difference in intensity behaviour between the aggregated and unaggregated apoSH A4V sample is subtle in Trp-32F but more obvious in Trp-32U. In the unaggregated sample, Trp-32U intensity remains mostly unperturbed in the folded temperature range, and then, although noisy, gradually increases in signal upon entering the unfolding temperatures (Figure 6.18). In contrast, Trp-32U in the aggregated sample exclusively shows a decline in intensity, but similar to the unaggregated sample, only shows changes in the linewidths at unfolding temperatures (Figure 6.18). The Trp-32U in the aggregated sample is decreasing in population throughout the temperature range,

and most likely adding to the aggregates. With any unfolded states, represented by Trp-32U, feeding into the aggregate, the folded state (indicated by the Trp-32F) must reestablish its equilibrium with the unfolded state and thus its intensity changes are hinting at unfolding sooner than expected. Hence, the intensity decrease for the Trp-32F in the aggregated sample occurs at an earlier temperature than observed for the unaggregated sample. Collectively, in the presence of aggregates, the unfolded species appears to add to the aggregate population, and consequently, the folded species must reestablish the folded-unfolded equilibrium.

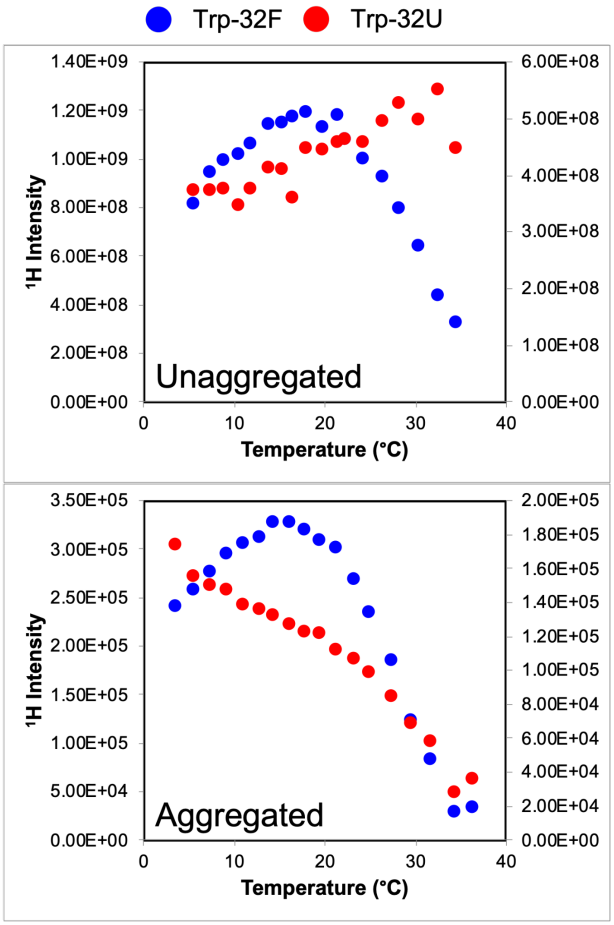



Figure 6.18: Trp-32F and Trp-32U peak intensities throughout the temperature ramp. All peak intensities were normalized by receiver gain. The Trp-32F (blue) and Trp-32U (red) peaks are evaluated throughout the temperature ramp. The Trp-32F and Trp-32U peak intensity for the apoSH A4V samples without (top) and with (bottom) the presence of aggregates. For the unaggregated sample throughout the folded temperature range, the TrpF-32F peak intensity increases, whereas the Trp-32U peak shows minimal changes. However, the Trp-32F and Trp-32U intensity begins to decrease and increase, respectively, when

the protein is in the unfolding temperature ranges. Meanwhile, the Trp-32U for the aggregated sample consistently decreases in intensity, even in the folded temperature range.

#### 6.4.3.3.2 Temperature-dependent peak intensity analysis

For the folded intensity slopes, there are considerable differences between the aggregated and unaggregated samples. Intriguingly, both samples show comparable profiles, however the values in some structural regions of the unaggregated sample are smaller than the aggregated sample (Figure 6.19). The most notable areas of differences are ß2, loop 6, and loop 7; each of these areas have larger absolute values in the unaggregated sample which suggests they are less structured (Figure 6.19 and Table 6.12). Aside from ß2, the loops (6 and 7) show increased dynamics by temperature coefficients for the aggregated sample, not the unaggregated sample. This likely means Trp-32F for the unaggregated sample is not reflective of other folded structures throughout the monomer, which is discussed further in Chapter 4. Therefore, with a much smaller range in values, the Trp-32F peak for the aggregated sample seems to more closely reflect the folded structures in its monomer than for the unaggregated sample.

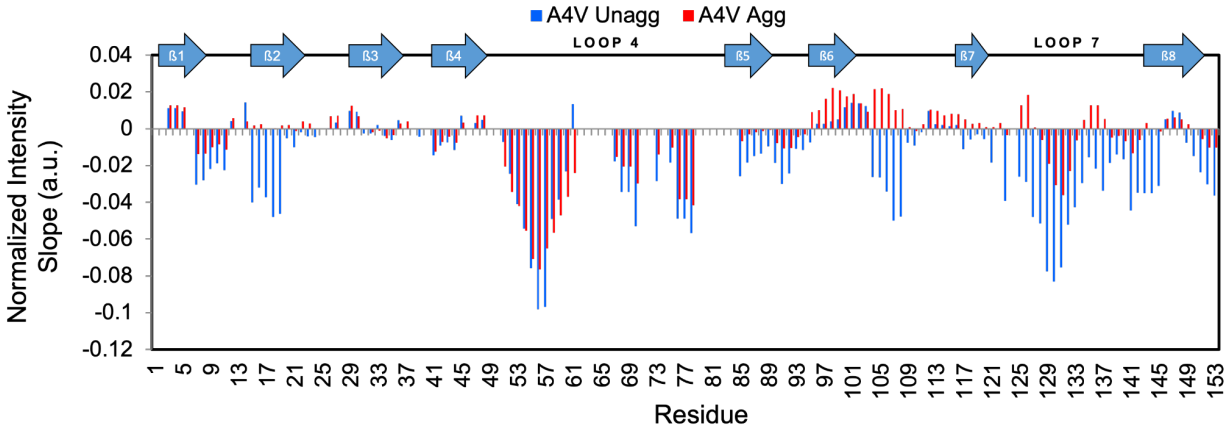



Figure 6.19: Post-normalized folded intensity slopes for apoSH A4V without aggregates (unagg) and with aggregates (agg). The rolling averages for normalized intensity slopes are considered for the folded temperature ranges. The rolling averages for normalized intensity slopes are for the sample without aggregates (blue) and with aggregates (red). The closer the values are to 0, the residue responds to temperature changes comparable to the Trp-32F peak that the data are normalized by.

| | A4V unagg | A4V agg |
|---|---|---|
| **average** | -0.0194 (126) | -0.0049 (124) |
| **standard dev** | 0.02399 | 0.01898 |
| | | |
| **N-terminal B-sheet** | -0.0074 (29) | 0.0039 (29) |
| **C-terminal B sheet** | -0.0095 (27) | 0.0005 (25) |
| **N** | - | - |
| **B1** | -0.0053 (5) | 0.0020 (5) |
| **loop I** | -0.0147 (4) | -0.0060 (4) |
| **B2** | -0.0229 (9) | 0.0016 (9) |
| **loop II** | -0.0013 (4) | 0.0042 (4) |
| **B3** | 0.0013 (8) | 0.0013 (8) |
| **loop III** | -0.0019 (2) | 0.0041 (1) |
| **B4** | -0.0039 (7) | -0.0020 (7) |
| **loop IV** | -0.0418 (20) | -0.0379 (20) |
| **B5** | -0.0164 (5) | -0.0027 (5) |
| **loop V** | -0.0209 (4) | -0.0084 (4) |
| **B6** | 0.0026 (8) | 0.0140 (8) |
| **loop VI** | -0.0126 (13) | 0.0105 (13) |
| **B7** | -0.0038 (6) | 0.0046 (6) |
| **loop VII** | -0.0386 (20) | -0.0042 (21) |
| **B8** | -0.0138 (9) | 0.0022 (7) |
| **C** | -0.0333 (2) | -0.0103 (2) |

Table 6.12: Average normalized intensity slopes for apoSH A4V with the absence of aggregates (unagg) and the presence of aggregates (agg). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

The aggregated samples intensity slopes and temperature coefficient data not aligning suggest there may be more complexities at play. In the aggregated sample, the Trp-32F and Trp-32U peaks reveal the sample is actively aggregating. Therefore, at minimum, each folded residue reflects the processes of protein folding, potential protein unfolding and subsequently any aggregation states the residue may access. However, the unaggregated samples will predominately only show protein folding processes until the protein begins to unfold. These additional processes are a probable culprit for the convoluted intensity data observed for the aggregated sample. These data further highlight the complexities associated with actively aggregating samples.

6.4.3.3.3 Max temperature for the aggregated sample may not exclusively reflect protein unfolding

With knowledge the sample is actively aggregating, the max temperature for the aggregate sample can now represent both local unfolding and/or association with the aggregate. Interestingly, akin to its Trp-32F behaviour, the unaggregated sample on average unfolds later than the aggregated sample (Table 6.13). With additional processes like the A4V monomer reestablishing its folded and unfolded equilibrium, on top of active aggregation, the aggregated sample shows a decrease in its folded population sooner. Although in section 6.4.3.3.1, it is clear the unfolded state is actively declining, and therefore assumed to be aggregating, it is unclear if the Trp-32F (folded state) itself is involved in aggregation. Therefore, this shift to lower max temperature for the aggregated sample could be a result of the monomer unfolding more and/or aggregating.

Moreover, max temperature analysis shows there are different sites of perturbation relative to the other analysis done above. Specifically, ß4, loop 4 and ß6 show decreases and ß7 shows increases in max temperature for the unaggregated sample relative to the aggregated sample (Figure 6.20 and Table 6.13). Now with the complexities of active aggregation, these differences could suggest the two monomers unfold differently, but likely some of these structures are self-associating in one sample and therefore showing an earlier max temperature. Therefore, peak shape analysis is complicated with added processes such as active protein aggregation.

| | A4V unagg | A4V agg |
|---|---|---|
| average<br>standard dev | 17.92 (126)<br>7.07 | 16.67 (125)<br>6.22 |
| | | |
| N-terminal B-sheet | 18.68 (29) | 16.99 (29) |
| C-terminal B sheet | 19.00 (27) | 18.56 (25) |
| N | - | - |
| B1 | 16.61 (5) | 17.48 (5) |
| loop I | 23.09 (4) | 19.20 (4) |
| B2 | 17.18 (9) | 16.29 (9) |
| loop II | 16.56 (4) | 13.96 (4) |
| B3 | 19.73 (8) | 18.63 (8) |
| loop III | 14.24 (2) | 20.59 (2) |
| B4 | 20.30 (7) | 15.57 (7) |
| loop IV | 19.04 (20) | 15.44 (20) |
| B5 | 19.25 (5) | 19.96 (5) |
| loop V | 9.76 (4) | 10.49 (4) |
| B6 | 20.64 (8) | 18.62 (8) |
| loop VI | 17.41 (13) | 17.88 (13) |
| B7 | 19.77 (6) | 23.14 (6) |
| loop VII | 16.25 (20) | 14.72 (21) |
| B8 | 16.68 (9) | 15.58 (7) |
| C | 16.00 (2) | 14.49 (2) |

Table 6.13: Average max temperatures for apoSH A4V with the absence of aggregates (unagg) and the presence of aggregates (agg). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

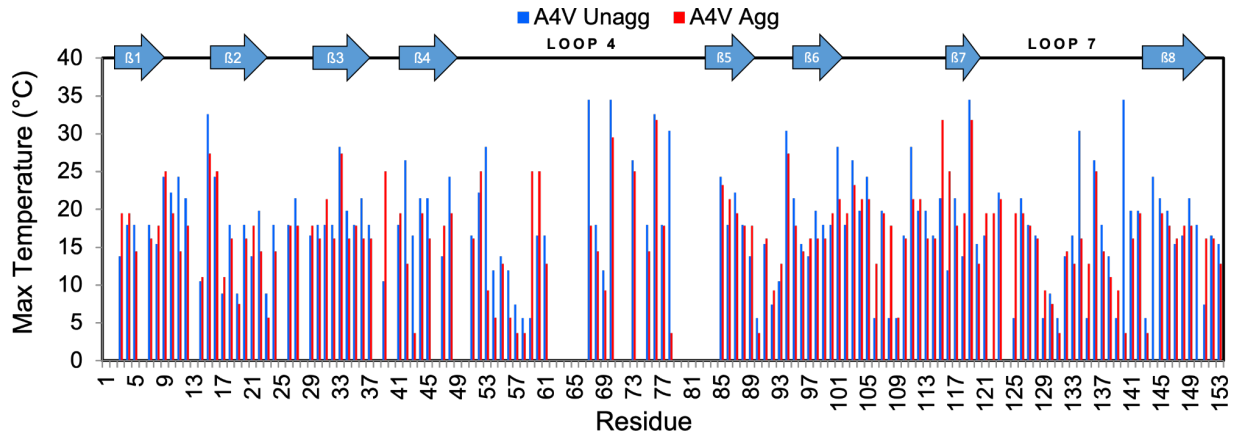



Figure 6.20: Plots of max temperature for apoSH A4V without aggregates (blue) and with aggregates (red).

### 6.4.3.4 Secondary amide temperature coefficients show minimal structural differences between the aggregated and unaggregated sample

From the secondary temperature coefficients, we find relative changes in ordered and disordered regions between the aggregated and unaggregated apoSH A4V samples.

Similar to their amide proton temperature coefficient and intensity profile data, the secondary temperature coefficient values are similar between the two samples (Figure 6.21). The unaggregated and aggregated N-terminus ß-sheets share similar values (18.91 and 18.35, respectively), but the C-terminus ß-sheet for the aggregated sample is not as random coil-like as the unaggregated sample (Table 6.14). Again, contrary to the amide proton temperature coefficient data, the aggregated sample appears less random coil-like, which agrees with the conclusions from the intensity profile data. However, collectively the differences in secondary temperature coefficients are not significant, and do not suggest there are considerable structural changes.

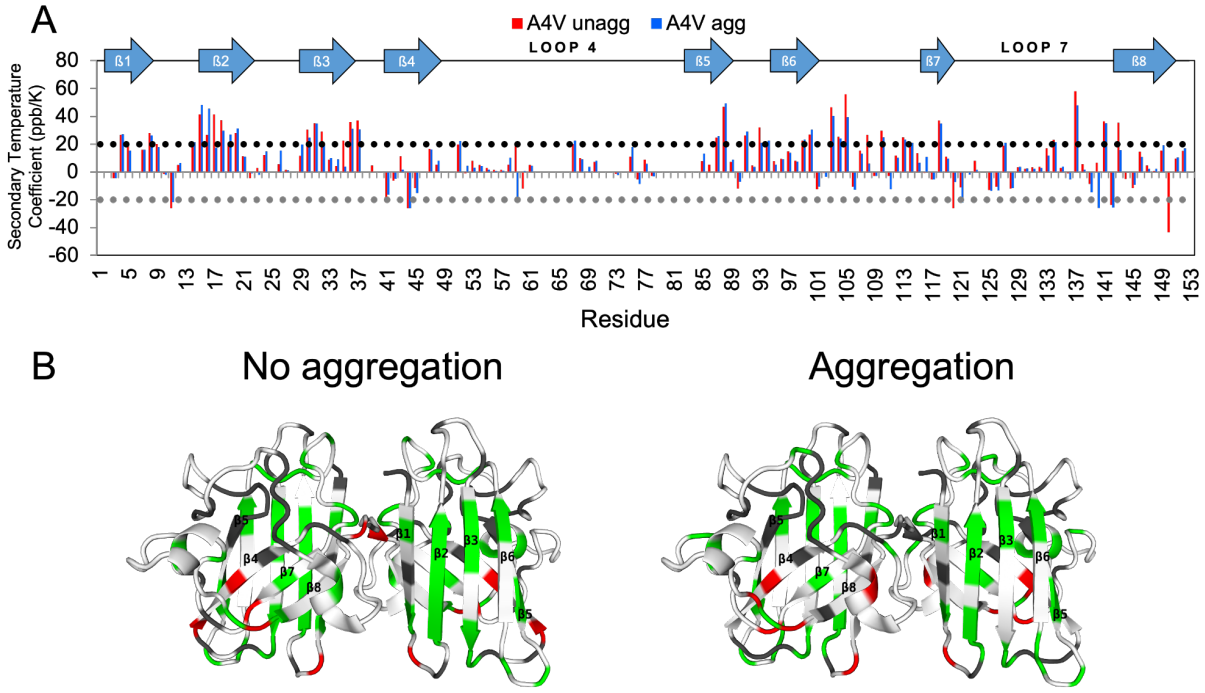



Figure 6.21: Amide nitrogen temperature coefficients for apoSH A4V without aggregates (no aggregation) and with aggregates (aggregation). (A) Temperature coefficients are represented by the blue (without aggregates, unagg) and red (with aggregates, agg) bars in the plot. (B) show 3D ribbon structures for the no salt (left) and with salt (right) data. Temperature coefficients are mapped onto the 3D structure of holoSS SOD1 (1HL5) and coloured based on the +/-20 ppb/K cutoff. Residues that show values greater than 20 ppb/K and less than -20 ppb/K are coloured green and red, respectively. Residues within the -20 to 20 ppb/K range are coloured white. Grey indicates residues where no data are available.

| | A4V unagg | A4V agg |
|---|---|---|
| average<br>standard dev | -9.25 (125)<br>15.44 | -10.17 (123)<br>13.56 |
| | | |
| N-terminal B-sheet | 18.91 (30) | 18.35 (30) |
| C-terminal B sheet | 5.21 (27) | 7.46 (25) |
| N | - | - |
| B1 | 16.89 (5) | 16.06 (5) |
| loop I | -0.61 (4) | 0.10 (4) |
| B2 | 24.59 (9) | 25.58 (9) |
| loop II | 5.52 (4) | 7.39 (4) |
| B3 | 22.20 (8) | 19.09 (8) |
| loop III | 20.83 (2) | 30.69 (1) |
| B4 | -4.70 (7) | -5.27 (7) |
| loop IV | 5.21 (20) | 4.19 (20) |
| B5 | 18.36 (5) | 19.51 (5) |
| loop V | 12.80 (4) | 11.61 (4) |
| B6 | 11.96 (8) | 12.69 (8) |
| loop VI | 18.45 (13) | 13.17 (13) |
| B7 | 4.91 (6) | 8.23 (6) |
| loop VII | 5.47 (20) | 1.02 (21) |
| B8 | 2.28 (9) | 7.37 (7) |
| C | 15.08 (1) | 17.03 (1) |

Table 6.14: Amide nitrogen temperature coefficients for apoSH A4V with the absence of aggregates (unagg) and the presence of aggregates (agg). Temperature coefficient values are averaged across all assignments (average), the beta-strands on N-terminal (ß1-3, 6) and C-terminal face (ß4, 5, 7, 8) and secondary structural elements, and each average also has the number of residues averaged in brackets.

### 6.4.4 Salt concentration, protein concentration and presence of aggregates: impacts on protein and NMR experiments

In this Chapter, the influence of salt concentration, protein concentration and presence of aggregates on apoSH SOD1s and their NMR data are explored. Salt shows the most dramatic impact on the protein and subsequently its NMR data. Both varying protein concentration and absence/presence of aggregate data show more subtle differences in their data based on their chemical shift, and more noticeable differences in their intensity profile analysis. Overall, these data highlight how solution conditions can impact both the protein dynamics and NMR data.

The impact of salt on protein has been extensively studied and found to be complex. In addition to screening charges, Hofmeister salt effects can alter protein stability and self-association. Moreover, salts can alter the balance of electrostatic and hydrophobic interactions between proteins.[267,269,283,284] For apoSH V148I, the binding of NaCl appears to be non-specific and centralized to the metal-binding region. More specifically there is a positively charged pocket that is critical to guiding the superoxide anion[272], and it is likely the anion ($Cl^-$) is interacting with the residues in this pocket. Additionally, the salt causes stabilization of ß4 while destabilizing its neighbouring strands, ß6 and ß7, and shifts the folded-unfolded equilibrium toward unfolded at folded temperatures. These points of disruption may explain how NaCl promotes apoSH SOD1 unfolding.[12] Overall, salt impacts apoSH V148I structure and subsequently perturbs its dynamics which can be measured by temperature-dependent chemical shift and intensity peak analysis.

Generally, increasing protein concentration increases crowding, which may minimize unfolding and/or favour intermolecular interactions.[285,286] Although the cellular concentration of SOD1 is reported to be 10-40 µM[12,287–289], and we are studying apoSH V148G at the mM range, it is important to note the cellular environment is generally very crowded. Since apoSH V148G does not dimerize readily, and shows at high protein concentration, it does not self-associate readily[12], it serves as an ideal reporter for the impacts of protein concentration on the apoSH monomer. Interestingly, at the low and high protein concentrations, the samples show negligible differences throughout all HSQC analyses. However, if a protein is susceptible to transient dimerization or self-association, it is likely the HSQC spectra will reflect that. Therefore, aggregation observed

at higher protein concentration for apoSH SOD1s is likely a result of enhanced collisions[12], and potentially amplified dynamics.[55] In general, the apoSH SOD1 concentration data suggest protein concentration does not significantly alter monomer dynamics.

Moreover, the presence of aggregates may be important in understanding the role aggregates have on the soluble monomer, and which states may be involved. The presence/absence of aggregate data show remarkable differences only by the intensity profiles and not by temperature-dependent chemical shifts. This could be explained by the fact that the intensity profiles can report shifts in population to states that may be in slow exchange, where the peak shape changes but the chemical shift may not. Moreover, our data suggest the presence of aggregates leads to continuous aggregation, which shifts the folded-unfolded equilibrium toward the unfolded state. Additionally, it is clear the unfolded population is declining throughout the temperature ramp, suggesting it is actively aggregating, whereas the folded species suggests earlier unfolding but it is unclear if it is aggregating.

Furthermore, the role each state has on aggregation appears complex. Although, extent of aggregation *in vivo* has correlated with extent of unfolding based on the DSC endotherms[25], it does not exclusively explain the *in vivo* and *in vitro* aggregation.[12,25,65] Previously, scientists reported partially folded or natively folded SOD1s being critical in the disease-relevant aggregation pathways.[290] However, it is impossible to exclude the extent of unfolding as a contributor to aggregation, and instead, these unfolded/unfolding species may be available to form an intermediate along the aggregation pathway. In addition, unfolded species have been available to template in conformational strain-based

aggregation. Prior research supports that SOD1 aggregation can show strain-like behaviour[28,37,291–293] and perhaps more unfolded species are critical to aggregate growth rather than initiation. In general, our work highlights how the presence of aggregates can impact NMR data, but also alter protein behaviour.

## 6.5 Conclusion

NMR is a powerful tool that can help identify structural perturbations between point mutations, solution conditions and protein states. Here, we systematically explored the impact of salt concentration, protein concentration and the presence of aggregates on apoSH SOD1s and their NMR data. The mutants chosen were to ensure additional processes were minimized, and the impact of the solution condition could be evaluated. In general, increasing salt has the most noticeable impact on the protein structure, predominantly centred around the pocket of positive charge near the metal-binding sites, and subsequently its NMR data. However, the protein concentration and presence of aggregates reveal more subtle differences. In general, our work highlights how each of these solution conditions can alter protein dynamics and to what extent the NMR data is impacted. Without these systematic studies exploring the impact different solution conditions have on protein structure and dynamics, and subsequently experimental data, the goal to understand protein folding, misfolding and aggregation will remain elusive.

# Chapter 7: Distinct Impacts of Zinc Binding on Dimerization and Thermal Unfolding of Disulfide-Reduced ALS-Associated Mutant SOD1s

Author Contributions:

Harmeen Kaur Deol: conceptualization, methodology, resources, validation, formal analysis, investigation, visualization; overall conceptualization, data curation, software, visualization, writing-original draft; Shanise Reddekopp: methodology, resources, validation; Elizabeth M. Meiering: conceptualization, supervision, project administration, funding acquisition.

## 7.1 Context

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease characterized by motor neuron degeneration. Mutations in superoxide dismutase 1 (SOD1), an antioxidant enzyme, are linked to ALS; the prevalent hypothesis for how mutant SOD1 causes disease is the formation of toxic protein aggregates. SOD1 undergoes post-translational modifications *in vivo*, including metal binding, disulfide bond formation and dimerization, to reach its final maturation state, a stable homodimer. Recent studies support the hypothesis that immature states play key roles in the disease pathology, with monomeric forms being particularly toxic. The zinc-bound form of SOD1 with a reduced disulfide bond (EZnSH SOD1) is considered a critical initial state involved in the maturation pathways. Our previous work shows EZnSH SOD1 forms a weak dimer relative to its more mature states, but minimal studies have explored the effects a diverse

set of ALS-associated mutants have on the stability of the Zn-bound state. Here we analyze EZnSH SOD1 using a combination of isothermal titration calorimetry (ITC), to measure dimer dissociation as well as metal binding, with complementary differential scanning calorimetry (DSC), to measure global protein unfolding; this approach allows for the dissection of diverse impacts of mutations on SOD1 stability. Interestingly the mutations in EZnSH SOD1 had either a destabilizing effect or no effect on the dimer interface. These results are vastly different from studies using the same variants in the metal free, disulfide intact form, where all mutations significantly destabilized the interface. These findings may result in a direct consequence in further maturation of SOD1, as the zinc bound form is key in binding to the copper chaperone (CCS), which catalyzes formation of mature holo SOD1. These findings help rationalize maturation of SOD1 and potential roles for monomer forms of SOD1 in the neurotoxic aggregation linked to ALS.

## 7.2 Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease involving cytotoxic conformations of Cu, Zn superoxide dismutase (SOD1). It has been widely hypothesized that immature states of this enzyme are involved in the disease pathology. In addition to the multiple SOD1 gene mutations associated with ALS, maturing SOD1 involves multiple post-translational modifications, hence multiple states can be considered as immature. Mature SOD1 is a dimer with both metal cofactors, copper and zinc, and an intrasubunit disulfide bond. Considering no specific lack of modifications is associated with disease, characterizing each state becomes crucial to expand our

understanding of the disease mechanism. Aside from disease, SOD1 also serves as a useful model to decipher the impact of post-translational modifications on protein folding and misfolding.

Zinc binding to SOD1 is considered a critical initial step in the SOD1 maturation pathway. The zinc ion serves in enhancing both SOD1 stability and enzyme activity.[67] The binding of the zinc ions aids in structurally stabilizing the zinc-binding loop and the electrostatic loop in SOD1. Since residues of the zinc-binding loop are involved in the dimer interface, stabilizing the zinc-binding loop plays an important role in the dimerization of SOD1. However, the electrostatic loop is responsible for guiding the substrate to the active site, where the copper ion cycles between Cu(I) and Cu(II) redox states to complete the enzyme function. Aside from the zinc ion structurally stabilizing the electrostatic loop, it is also involved in helping the copper ion transition between redox states. Thus, potential perturbations in zinc binding will have detrimental effects to SOD1 structure and function.

The immature zinc bound SOD1 states serves an essential role in maturation, but its dynamic nature in conjunction with ALS-associated mutation could also lend way to aberrant association. Metal-free and disulfide reduced SOD1 (apoSH SOD1) is thought to exist predominantly as a monomer to which a zinc ion binds (EZnSH SOD1) to form a weak homodimer.[44,46,56] The formation of this dimer interface is important because it is the same interface the copper chaperone for SOD1 (CCS) recognizes to form a heterodimer with SOD1 to finish the final steps of maturation.[43,46,294] Thus, perturbations in this dimer interface may have detrimental impacts on SOD1 maturation, resulting in a pool of immature SOD1, which can give rise to folding intermediates involved in toxic aggregation.[34,221,295,296]

With multiple studies on this EZnSH SOD1 state, few have been able to systematically quantitate the dimer interface or global stability of the Zn-bound state using a diverse set of ALS-associated mutations. Here we analyze zinc bound, disulfide reduced SOD1 (EZnSH SOD1) using a combination of isothermal titration calorimetry (ITC), to measure dimer dissociation as well as metal binding, with complementary differential scanning calorimetry (DSC), to measure global protein unfolding. The combination of techniques allows for the dissection of diverse impacts of mutations on SOD1 stability. Interestingly, the location of the mutation in EZnSH SOD1 relative to a metal-binding site or the dimer interface impacts dimerization. These results are vastly different from studies using the same variants in the metal free, disulfide intact form, where all mutations significantly destabilized the interface, regardless of position. The global stability of EZnSH SOD1 was not significantly perturbed by mutations, and suggests the zinc bound state stability is from the monomer, not dimer. These findings may result in a direct consequence in further maturation of SOD1, as the zinc bound form is key in binding to the CCS, which catalyzes formation of mature holo SOD1. These findings help rationalize impact of metal binding and dimerization in protein folding, and potential roles for monomer forms of SOD1 in the neurotoxic aggregation linked to ALS.

## 7.3 Methods

### 7.3.1 Recombinant expression and purification of holo human SOD1

Expressed human pWT SOD1 (varies from wild-type by substituting Cys 6 and 111 with Ala and Ser, respectively) [65,297] was expressed in Escherichia coli (strain QC799) cells with a pHSOD1ASlacI1 vector as described previously.[52,65,298] Protein was purified

using osmotic shock to release SOD1 from the periplasm and hydrophobic interaction chromatography to separate hSOD1 from other proteins present. Purity was determined using a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

7.3.2 Demetallation and reduction of holo SOD1

All chemicals are from BioShop Canada, unless otherwise stated. Holo SOD1 was demetallated through pH-induced unfolding and metal chelation using ethylenediaminetetraacetic acid (EDTA) as described previously.[52,208] Complete removal of metals was confirmed by differential scanning calorimetry, providing a characteristic $T_{0.5}$. The disulfide bond formed by Cys 57 and 146 was reduced using chemical unfolding with 2 M guanidine hydrochloride followed by reduction with TCEP-HCl.[65] Protein concentration was measured by the absorbance and the molar extinction coefficient of 5,400 $M^{-1}$ $cm^{-1}$ for SOD1 monomer at 280 nm.[299] Reduction of the sample was confirmed by iodoacetamide modification of the free thiols and subsequently run on an SDS-PAGE.[300]

7.3.3 Titration with zinc into apoSH SOD1s

A range of stoichiometric equivalents (0.5-2) of zinc sulfate (Fisher Scientific) was added to a reduced apo pWT and mutant samples, followed by an anaerobic incubation at ambient temperature. Correct metal status was verified using a chelator, 4-(2-pyridylazo)rescorcinol (PAR, Sigma-Aldrich). The assay was adopted from Crow *et al*.[209] and Mulligan *et al*.[301] as modified by the Meiering lab[210] (further explained in S2.0).

7.3.4 Using isothermal titration calorimetry (ITC) to measure dimer dissociation

ITC experiments were performed to characterize the dimer interface by measuring the heats of dissociation as described previously[53,302,303] using a Microcal Isothermal

Titration 200 instrument (Microcal Inc., Northampton, MA). A concentrated sample of EZnSH SOD1 (0.934-1.14 mM dimer in 20 mM HEPES and 1 mM TCEP (pH 7.4)) was prepared using ultrafiltration concentration. Small volumes (0.2-0.5 µL) of the sample were injected into the reaction cell containing an identical buffer, to measure the heat associated with dissociating the dimer. The dissociation was measured at a range of temperatures (10-37 °C). Integrating the power versus time trace quantified the heats for each injection. The heat plots were fit using Microcal Origin 7.0 (Microcal Inc, Northampton, MA, USA) to a dimer dissociation model[302,303] using (Eq. 2.1):

$$q_i = V\Delta H_d\left([M_i]-[M_{i-1}]\left(1-\frac{v}{V}\right)-f_m[M_o]\frac{v}{V}\right)+q_{dil} \quad \text{(Eq. 2.1)}$$

where $q_i$ is the heat of dissociation normalized per mol of total monomer injected, $i$ is the injection number, $V$ is the reaction cell volume, $\Delta H_d$ is the enthalpy of dimer dissociation (per mole of total protein as monomer), $v$ is injection volume, $[M_o]$ is the total monomer concentration in the syringe, $[M_{i-1}]$ and $[M_i]$ is the total protein concentration in the cell before and after injection, respectively, and $q_{dil}$ is the heat of titrant dilution not related to dissociation. The fraction of monomer, $f_m$ present in the syringe is described by (Eq. 2.2):

$$f_m = \frac{1}{4[M_o]}\left(-K_d+\sqrt{K_d^2+8K_d[M_o]}\right) \quad \text{(Eq. 2.2)}$$

where $K_d$ is the dissociation constant from driving reduced EZnSH-pWT dimer ($N_2$) to folded monomers (M), $N_2 \rightleftharpoons 2M$. When fitting $\Delta H_d$, $q_{dil}$, and $K_d$ were set as floating parameters and were used to calculate $\Delta G_d(T_{exp})$ using (Eq. 2.3):

$$\Delta G_d(T_{exp}) = -RT_{exp}\ln K_d(T_{exp}) \quad \text{(Eq. 2.3)}$$

where $R$ is the universal gas constant.

7.3.5 Zinc titration using ITC

Zinc titrations were performed using ITC (Microcal Isothermal Titration 200 instrument (Microcal Inc., Northampton, MA)) in a more commonly used manner to assess zinc affinity for both metal-binding sites in apoSH SOD1. Both protein sample and titrant were degassed thoroughly prior to experiments, and the reference cell was filled with MilliQ water. Volumes of 0.35 µL of $ZnSO_4$ (0.193 mM and 0.394 mM, Fisher Scientific) were injected into samples of pWT (0.0042 mM and 0.0079 mM, respectively) in 20 mM HEPES, 1 mM TCEP at pH 7.4. The data were fit to a 2-site sequential binding model built into Microcal Origin 7.0.

7.3.6 Differential scanning calorimetry (DSC)

DSC experiments were performed as described previously[52,65], using an LLP cap DSC (MicroCal Inc., Malvern Instruments Ltd.). Varying concentrations of samples (0.68-4.143 mg $mL^{-1}$) in 20 mM HEPES, 1 mM TCEP at pH 7.4 were scanned at a rate of 1 °C $minute^{-1}$. Prior to data fitting, each protein scan has the reference scan, buffer against buffer, subtracted from it and normalized to the protein concentration. Each scan was fit to both 2-state and 3-state with monomer intermediate models using Microcal Origin 7.0 (Microcal Inc., Northampton, MA, USA), represented by (Eq. 2.4) and (Eq. 2.5) respectively[52,65,304]:

$$C_p = (A+Bt)(1-a)+(E+Ft)a+\left(\frac{b\Delta h_{cal}^2(t)}{RT^2}\right)\left(\frac{a(1-a)}{2-a}\right) \quad \text{(Eq. 2.4)}$$

$$C_p = (1-a_1)(A+Bt)+a_1(1-a_2)(C+Dt)+a_1a_2(E+Ft)+$$
$$\left[\frac{b_1\Delta h_{cal,N_2\leftrightarrow 2M}(t)+a_2b_2\Delta h_{cal,M\leftrightarrow U}(t)}{RT^2}\right]\Delta h_{cal,N_2\leftrightarrow 2M}(t)\frac{a_1(1-a_1)}{2-a_1}+$$
$$\left[\left[\frac{b_1\Delta h_{cal,N_2\leftrightarrow 2M}(t)+a_2b_2\Delta h_{cal,M\leftrightarrow U}(t)}{RT^2}\right]\frac{a_1a_2(1-a_1)}{2-a_1}+\left[\frac{b_2\Delta h_{cal,M\leftrightarrow U}(t)}{RT^2}\right]a_1a_2(1-a_2)\right]$$
$$\Delta h_{cal,M\leftrightarrow U}(t) \quad \text{(Eq. 2.5)}$$

where $C_p$ is the total specific heat absorption at temperature, $t$ (in Celsius); $A$ ($E$) represents the intercepts of the folded (unfolded) baseline; $B$ ($F$) is the slope of the folded (unfolded) baseline; $\alpha$ is the extent of protein unfolding; $\beta$ is the molecular weight of the dimer multiplied by the ratio of van't Hoff to calorimetric enthalpy; $\Delta h_{cal}$ is the specific calorimetric enthalpy of unfolding at unfolding; $R$ is the universal gas constant; $T$ is the temperature in Kelvin. The subscripts *1* and *2* are notation for dimer dissociation ($N_2 \leftrightarrow 2M$) and monomer unfolding ($M \leftrightarrow U$), respectively, and derivations are defined in A1.0.

## 7.4 Results and Discussion

### 7.4.1 Nonmetal-binding and dimer-interface mutants have strong affinity for the first equivalent of zinc

Prior to assessing dimer dissociation strength, the zinc affinity for the zinc site is essential to characterize for both pWT and all mutants. The mutants chosen varied in location and types of mutation (Figure 7.1). To ensure dissociation of the dimer is not followed by subsequent metal loss, zinc titration is performed at a protein concentration mimicking that of the anticipated first injection during the dimer dissociation experiments. From the $K_d$ data in section 7.4.2, the protein concentration for the zinc titration experiments ensures the population remains predominantly monomeric (> 80 %) for the first zinc equivalent. However, the titration is carried until multiple (> 4) zinc equivalents are added to the apoSH SOD1 sample to understand the strength of zinc affinity for both metal sites.

Zinc has an extremely high affinity for the zinc site of apoSH pWT (75 nM)[305], but also can bind to the copper site of apoSH pWT with a relatively lower affinity (100 μM).[56] Although zinc has a high affinity for the zinc site, the affinity may be weakened for the monomer relative to the dimer, such that the initial ITC injection into the calorimeter cell might drive zinc release upon dimer dissociation. If the affinity is strong, the binding heats will be large and detectable, but weakened affinity will result in low heats as a result of minimal binding.

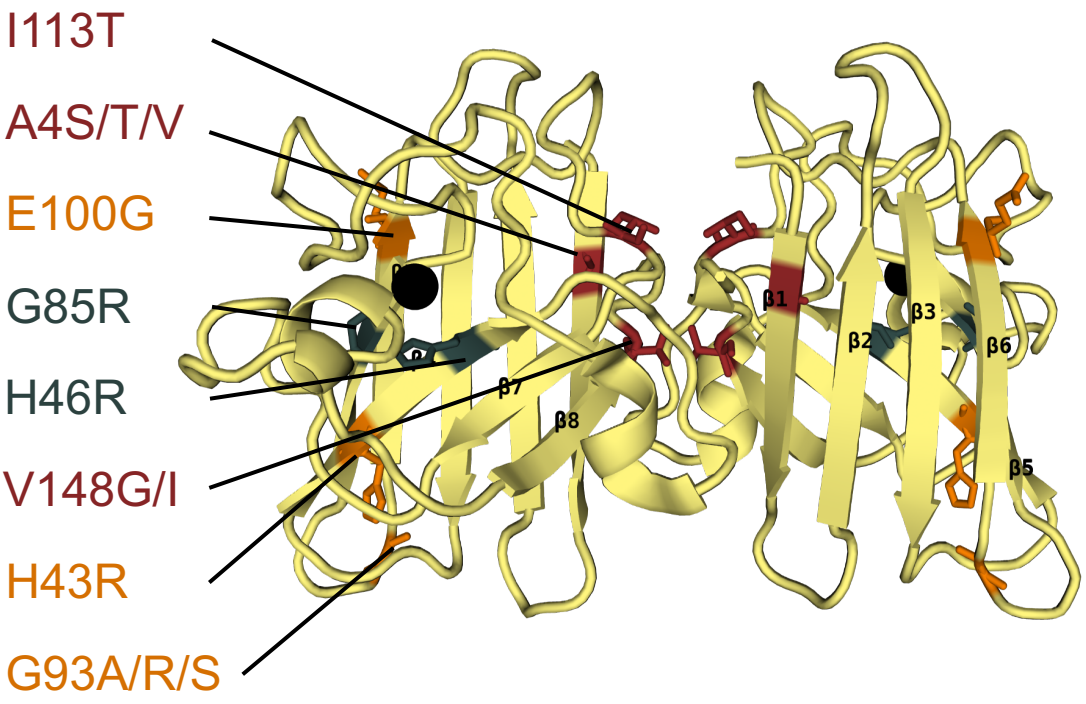



Figure 7.1: Homodimeric E,Zn SOD1 with mutations. Mutation sites and the native side chains are shown as sticks. Dimer interface mutations are in red, metal binding mutations are in blue, and non-metal binding mutations are in orange (PDB:1HL5, rendered with PyMol).

The zinc affinity for the first metal site is heavily dependent on the location of the mutation (Figure 7.2A). For the nonmetal-binding and dimer-interface mutants, the zinc titration heats, and their profile matches that of pWT (Figure 7.2A). The affinity for the zinc

is outside the fittable range of ITC and does not offer reliable values, but the measurable heats are qualitative indicators of the binding affinity. As expected, the heats until the first equivalence of zinc are large (henceforth referred to as Phase 1) and binding is saturated by the first equivalence of zinc, which suggests the first binding event is strong (Figure 7.2A). However, the subsequent equivalent of zinc (Phase 2) has lower heats than the heats for the first binding event, and never reach saturation, which agrees with literature findings that the second binding event is weaker.[56]

Where the nonmetal and dimer-interface mutants have pWT-like affinity for the zinc, the metal-binding mutants have heats that do not agree with pWT-like behaviour (Figure 7.2A). Both G85R and H46R are known to have issues binding metals, and unsurprisingly their heats are significantly dampened for Phase 1 and throughout the titration, further supporting their inability to effectively bind the metals at these low monomer concentrations. Therefore, as observed in section 7.4.2, the dimer dissociation heats for both metal-binding mutants are accompanied by additional heats which we now understand as metal dissociation. Collectively, only metal-binding mutants have weakened metal binding to the monomer, whereas the rest of the mutants have a strong affinity for zinc as monomers which means the zinc will not dissociate upon dimer dissociation.

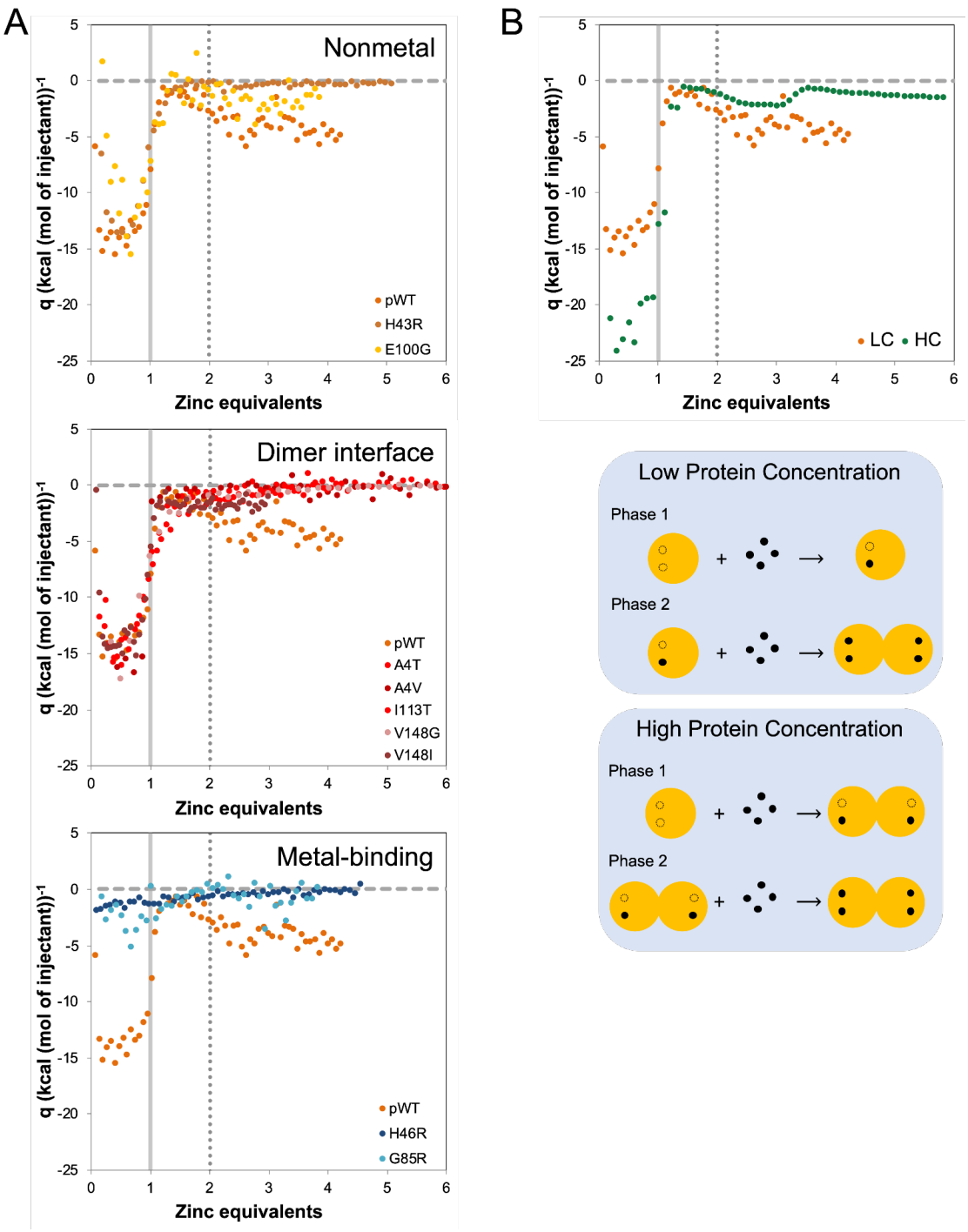



Figure 7.2: Zinc titration ITC isotherm with apoSH SOD1 and at varying protein concentration. (A) ApoSH SOD1 binding zinc in the first site (Phase 1, grey solid vertical line) results in large exothermic peaks. The smaller exothermic peaks for the zinc binding to the copper site (Phase 2, grey dotted vertical line) are due to its decreased binding affinity and potential dimerization. (B) Zinc titration data for pWT and V148I with high (1.46 mM, HC) and low (0.008 mM, LC) concentrations, respectively. In the first phase (Phase 1), the HC sample has heats of both zinc binding and dimerization. In a second phase (Phase 2, around 2 equivalents of added zinc, grey dotted vertical line), zinc may bind to the copper site. Associated heats of

subunit association depend on the protein concentration (see main text). Horizontal gray dashed line is the reference for zero heats.

Similar to the dampened heats for the metal-binding mutants for Phase 1, the zinc binding to the second site (Phase 2) also has small heats. The heats in both Phase 1 and 2 are dependent on protein concentration (Figure 7.2B). At low protein concentration where the protein remains predominantly monomeric with one zinc equivalence, Phase 1 heats exclusively correspond to zinc binding and no dimerization since the protein remains monomeric. In Phase 2, the heats at low monomer concentration include the binding of zinc to the copper site and the subsequent dimerization. In contrast, the high protein concentration sample has heats in Phase 1 that correspond to both zinc binding and dimerization. Then Phase 2 has heats for only the second zinc binding since the dimer formation occurred in Phase 1 at this concentration. This difference in expected events at Phase 2 explains the higher heats observed at low concentration relative to the high concentration sample (Figure 7.2B). In the low concentration samples, there is both dimerization and metal binding, whereas in the high concentration there is exclusively metal binding, respectively. In summary, the zinc binding is strong for the apoSH monomer, suggesting it will not dissociate upon dimer dissociation.

### 7.4.2 Metal-binding mutants and dimer-interface mutants impact the ability to dimerize

Isothermal titration calorimetry (ITC) is typically used for measuring binding events, but it can also be useful for assessing oligomerization events. Here we measured the equilibrium dimer dissociation ($K_d$) for the EZnSH SOD1 throughout a temperature range by titrating concentrated (>0.75 mM) samples into a thermally-regulated calorimeter cell. The small endothermic heats from dimer dissociation at all temperatures were fit both

individually and globally to a dimer dissociation model (Figure 7.3A). In this section, $K_d$ values are reported at 25°C, unless otherwise stated. We use ITC to quantitate dimer dissociation for a diverse set of mutants which we classify based on their location relative to either the dimer interface or a metal-binding site (Figure 7.1). Mutants that are either near or involved in the dimer interface are classified as dimer-interface mutants. Mutants that are not associated with the dimer interface were grouped as either nonmetal-binding mutants or metal-binding mutants based on whether the mutant has no impact on metal-binding or is known to perturb the metal-binding sites, respectively. Quantitating the dimer interface allows us to assess how fALS-mutations impact the dimer interface which may have a cascading effect on maturation.

Due to their position relative to the dimer interface and metal-binding regions, the EZnSH nonmetal-binding mutants should not perturb the dimer interface (Figure 7.3B). Our control, pWT, has a $K_d$ of 51±41 μM at 25°C and 15±3 μM from the global fits (Table 7.1). All the other nonmetal-binding mutants are as follows: H43R, G93A, G93R, G93S, and E100G. They have an average value of 41±16 μM at 25°C, and individually they all have values comparable to pWT within error. Notably, we used multiple G93 mutants to systematically assess the impact of hydrophobics and charge in nonconsequential regions and found negligible impact on the dimer strength. These data suggest mutants that are neither near the dimer interface or a metal-binding site can form a dimer of comparable strength to pWT in the zinc bound state. As expected, all the EZnSH nonmetal-binding mutants show dimerization upon zinc binding, and they have comparable dimer dissociation values to each other.

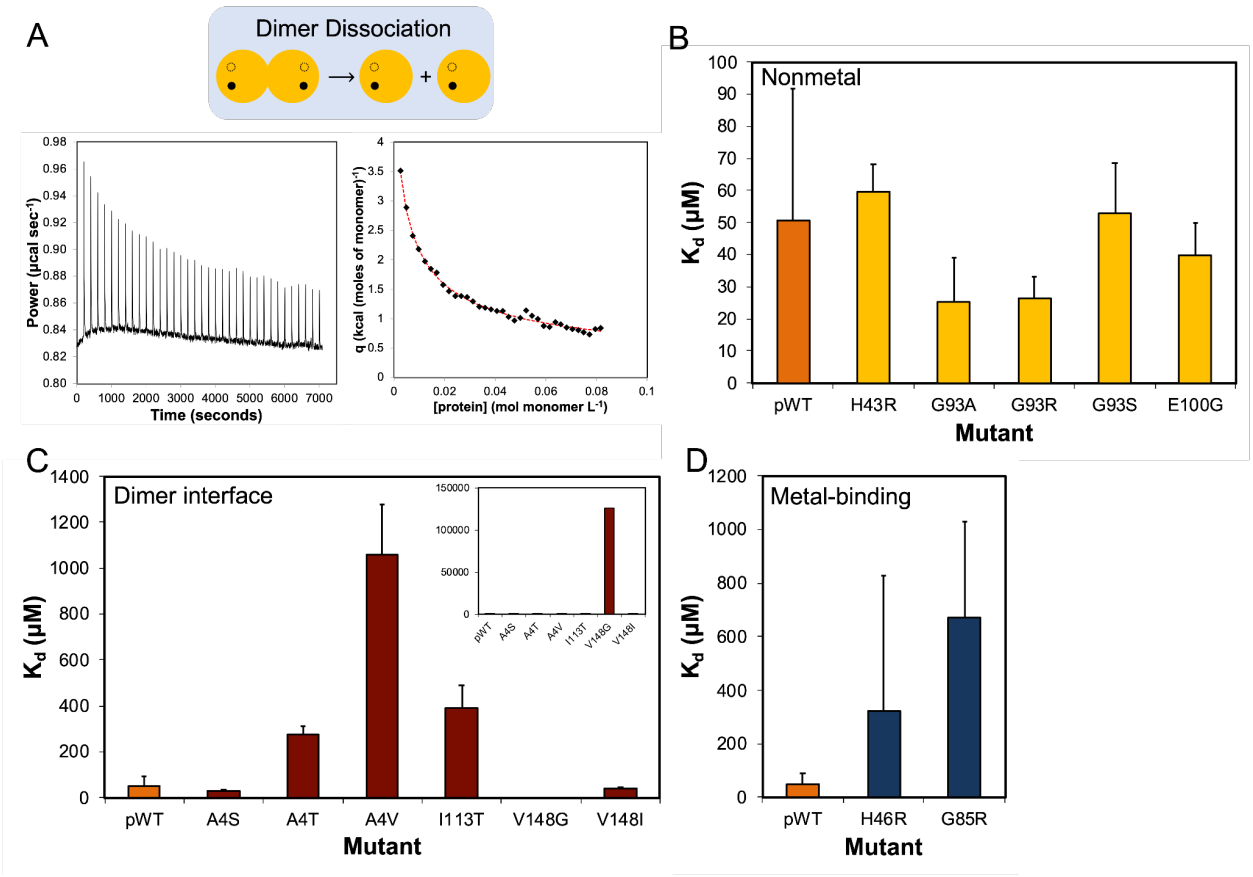



Figure 7.3: Dimer dissociation heats and strength for EZnSH SOD1 mutants relative to pWT. (A) Representative heats, power, and the integrated heats, q, from the raw data are plotted against SOD1 monomer concentration (diamonds) which is fit to a dimer dissociation model (red line). (B), (C), and (D) are the $K_d$ values at 25 °C for nonmetal binding (yellow), dimer interface (red) and metal-binding (blue) mutants relative to pWT (orange), respectively. Error bars are from multiple experiments. (C) Due V148G's high value, it is plotted in the offset plot.

Unlike the nonmetal-binding mutants, dimer-interface mutants are expected to affect dimerization based on their location and mutation. We classified the following mutants as dimer-interface mutants: A4S, A4T, A4V, I113T, V148G, and V148I. Unsurprisingly, mutations in this group had both stabilizing and destabilizing effects on the dimer interface (Figure 7.3C). Mutations in A4 show significant diversity based on the mutation, however considering the interface is formed by predominantly hydrophobic-based association[306], perturbing the hydrophobicity is likely destabilizing. Interestingly, although serine and threonine are both polar uncharged amino acids in the place of a

hydrophobic residue, the serine substitution was stabilizing (28±5 μM) whereas the threonine was destabilizing (273±41 μM) (Table 7.1). However, most surprising was A4V being the most destabilizing A4 mutation (1056±223 μM) considering the SOD1 dimer interface is predominantly hydrophobic-based association, one would think increasing the hydrophobicity would only serve to increase dimer strength, not decrease. It is clear there are more factors impacting dimerization other than hydrophobicity, and perhaps here the bulkiness of the amino acid substitution should be considered.

Since A4 mutations impact on dimer strength did not correlate with expected hydrophobicity trends, it is worth considering the impact decreasing the hydrophobicity may have on dimerization in other positions. In both A4T and I113T, the substitution of a hydrophobic amino acid for a polar uncharged one, the interface was destabilized to a comparable extent (273±41 μM and 390±100 μM, respectively) (Figure 7.3C and Table 7.1). Similarly, V148G also has a mutation that results in decreased hydrophobicity and has the most destabilized interface (126045±218200 μM). It is worth noting due to the error in this value, and the minimal heats of dissociation, EZnSH V148G is assumed to exist as a monomer, and does not dimerize in the zinc bound state. However, replacing V148 with a conserved mutation like isoleucine did not perturb the interface and rather formed an interface of comparable stability to pWT (39±5 μM). It is interesting that both conserved mutations, V148I and A4V, had such a contrasting impact on the interface. Overall, dimerization is sensitive to the dimer-interface mutants, and shifts in hydrophobicity from the mutations do not exclusively explain the impact on the interface.

Comparable to the dimer-interface mutants, mutants that have aberrant metal binding may not result in the interface-stabilizing expected upon zinc binding. Two

mutants with known metal issues are H46R and G85R; H46R is a copper and zinc binding residue, and G85R is a zinc binding residue.[225,307] Due to the nature of these mutations, understanding their zinc affinity is vital. As expected, they were the only two mutants that showed perturbed zinc affinity (section 7.4.1), which may explain their significantly weakened dimer interfaces, 323±507 μM for H46R and 670±360 μM for G85R (Figure 7.3D and Table 7.1). The destabilized interface here is perhaps not a result of the same interface perturbation observed for dimer-interface mutants. Instead, the metal-binding mutants experience weakened metal binding which results in less structural stabilization. With limited stabilization, the interface is less structured compared to mutants that feel the complete stabilizing impact of the zinc binding.[307] Interestingly for both mutants, the dimer dissociation endotherm is characteristically noisier than the data collected for the other mutants. In lieu of the typical sharp endothermic peak, the metal-binding mutants have increased signal and peak tailing, most noticeably in G85R. Both the increased signal and tailing observed suggest there are additional processes other than dimer dissociation occurring, such as zinc dissociation and/or significant structural rearrangement, that might result in slow equilibration. To summarize the metal-binding mutants had weakened zinc affinity that subsequently impacted dimerization. Therefore, EZnSH SOD1 dimerization is dependent on the mutation and its position relative to the dimer interface.

### 7.4.3 Disulfide bond formation stabilizes the dimer interface more than zinc binding

Both metal binding and disulfide bond formation are key stabilizing post-translational modifications. Both are critical in structurally stabilizing SOD1 but do so

through different means. The disulfide bond is critical in forming the interface by anchoring the zinc-binding loop to the β-barrel, which favours dimerization.[308,309] In contrast to the direct structural role the disulfide bond has on the interface, from the data above, zinc binding has an indirect impact on stabilizing the interface. Zinc binding is locally stabilizing by constraining the dynamics of the zinc binding loop, which allows for a more native-like conformation that may favour dimerization.[310] However, due to their different methods of interface stabilization, mutations in either state likely impact the interface to differing extents. Here we compare the interface stability of the oxidized apo (apoSS) to the reduced zinc bound (EZnSH) to understand the individual contributions to SOD1 dimer stability.

Due to the direct structuring of the interface upon disulfide bond formation, we hypothesize apoSS SOD1 forms a tighter dimer than EZnSH SOD1. Between the pWT control for each state, it is clear the apoSS pWT has a ~750 times tighter interface than the EZnSH pWT (Figure 7.3 and Table 7.2).[53] It is worth noting, the dimer dissociation was only consistently reported at 37°C for all apoSS variants. Only EZnSH pWT dissociation was measured up to 37°C, 66±8 μM, which is only slightly higher than the value at 25˚C (51±41 μM). However, due to the decreased thermal stability of the EZnSH mutants (section 7.4.5), EZnSH mutants are reported up to 25°C to ensure the heats report only on dimer dissociation, and no other processes such as monomer unfolding. Nevertheless, the disulfide bond formation has more stabilizing effects on the dimer interface than zinc binding.

Although the pWT control shows significant differences in interface strength between the two states, introducing certain mutations can destabilize the strongest of

interfaces. Overall, we notice all the apoSS mutants have destabilized interfaces relative to apoSS pWT. However, as noted above, EZnSH mutants have more grouped behaviour based on their location and mutation. Interestingly, relative to their apoSS form, all EZnSH mutants have destabilized interfaces, except for G93R (Figure 7.4 and Table 7.2). G93R is the only mutant with comparable values in both apoSS and EZnSH, 20±3 μM and 26±7 μM, respectively. All the other mutants show no trends in dimer destabilization based on mutation or location between the two states, meaning the rank of mutants that are most destabilized to least destabilized were different between the two forms (Table 7.2). The weakened EZnSH dimer interface relative to the apoSS dimer might be important for transient dimerization to allow binding to the CCS for further maturation.

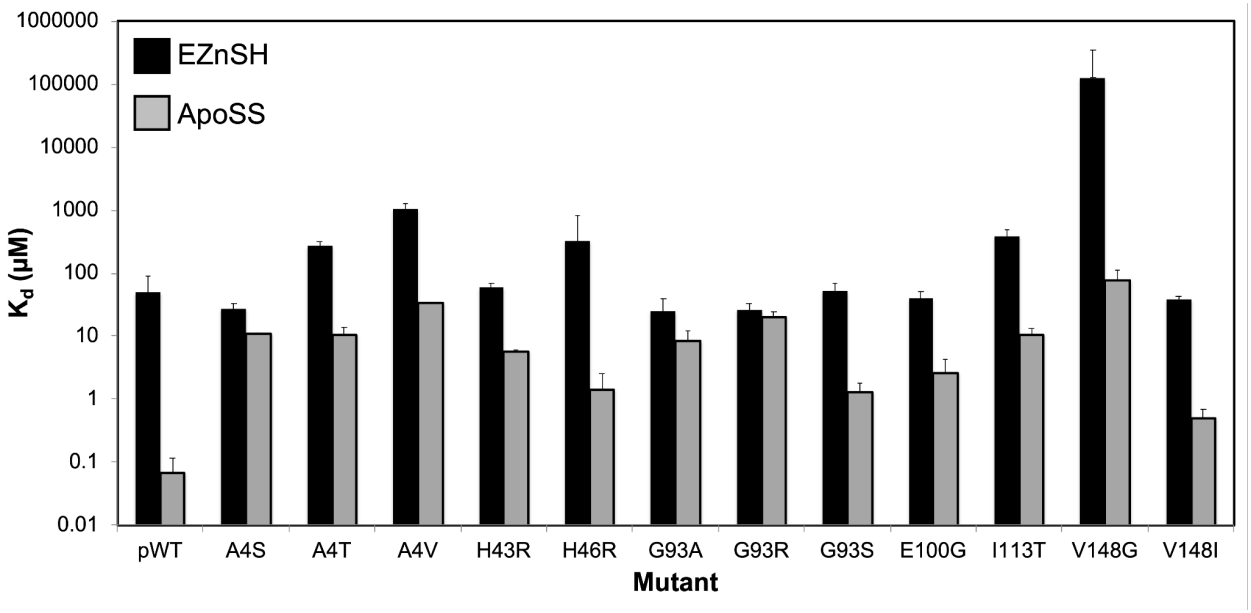



Figure 7.4: Dimer interface strength for EZnSH (black) and apoSS (grey) SOD1 mutants relative to pWT. Error bars are from averaging across multiple experiments. Due to the large range of values, the data are plotted on a logarithmic scale.

Although the individual fits for EZnSH SOD1 at 25˚C did not correlate with apoSS SOD1 at 37˚C, comparing the global fits of EZnSH SOD1 to the apoSS state shows an

interesting trend. Considering the dimer dissociation values of EZnSH have smaller ranges in values than apoSS SOD1 with temperature, we perform global fits on EZnSH data for a temperature range that coincides with minimal temperature-dependent dissociation. Correspondingly, when we compare the global fit values of EZnSH to the apoSS SOD1 values above, no trends are notable when we compare all mutants. However, when the mutants are grouped and ranked based on the earlier classification, nonmetal-binding, metal-binding and dimer-interface mutants, the later classification show a strong correlation. Upon ranking the nonmetal and metal-binding mutants in each state based on their dimer dissociation values, there was no significant trend ($R^2$= 0.10) (Figure 7.5). However, the dimer-interface mutants show a trend against each state ($R^2$=0.67) (Figure 7.5), with a notable outlier being A4S, and removal of this mutant increases the $R^2$ to 0.82. This suggests the mutation might have comparable destabilizing effects relative to the other dimer-interface mutations, but to varying extents in different SOD1 forms. This would explain why the mutants ranking relative to each other is correlated, but the $K_d$ values are remarkably different.

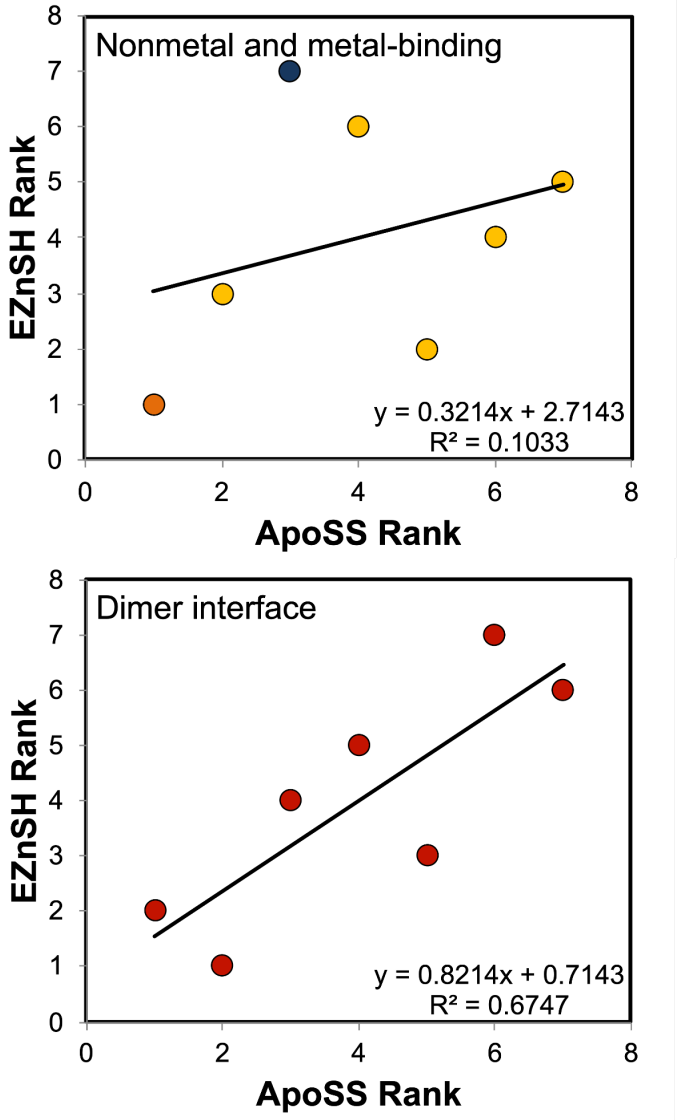



Figure 7.5: EZnSH and ApoSS Kd values at 25 °C and 37 °C, respectively, are ranked for nonmetal and metal-binding mutants and dimer interface mutants. The data are fit to a linear model (black line), and data are coloured according to their respective location. The nonmetal binding (yellow), dimer interface (red) and metal-binding (blue) mutants are shown relative to pWT (orange).

### 7.4.4 ApoSS dimer dissociation causes significant structural rearrangement relative to the EZnSH dimer

By measuring the dimer dissociation equilibrium ($K_d$) at different temperatures, other thermodynamic parameters can be evaluated, such as enthalpy of dissociation

($\Delta H_d$) and the specific heat capacity upon dimer dissociation ($\Delta C_{p,d}$). Understanding whether dimer dissociation is enthalpically favourable (exothermic) or unfavourable (endothermic), offers insight on the types of interactions that drive dissociation. Exothermic or endothermic dissociation is driven by hydrogen bonding, salt bridges and van der Waal forces, or hydrophobic events, respectively.[311] In addition, the temperature dependence of the dissociation enthalpy gives the specific heat capacity of dissociation, which indirectly describes the structural rearrangement that occurs upon dissociation.[87] Together, these values offer insight on the interactions driving dissociation.

By comparing enthalpy of apoSS SOD1 to the EZnSH SOD1, we parse out the extent to which different forces influence the two dimer interfaces. Both $\Delta H_d$ are positive values, which indicates dissociation is endothermic (Figure 7.6). However, the average $\Delta H_d$ values for apoSS are significantly larger than EZnSH, 12±7 kcal (mol monomer)$^{-1}$ at 37°C and 4.3±2.3 kcal (mol monomer)$^{-1}$ at 25°C, respectively (Table 7.2). It is known that enthalpy is temperature dependent, but the temperature range the apoSS SOD1 dimer dissociation is measurable is different from the EZnSH SOD1. However, due to the minimal temperature dependence of EZnSH SOD1, the comparison between enthalpy values highlights that the apoSS $\Delta H_d$ is consistently higher than the EZnSH $\Delta H_d$ aside from A4S (Table 7.2). This suggests there is more energy associated with dissociation for the apoSS state than the EZnSH state. The exception, A4S, has values of 4.5±1.3 kcal (mol monomer)$^{-1}$ at 37 °C for the apoSS form and 5.1±1.3 kcal (mol monomer)$^{-1}$ at

25 °C for the EZnSH form. While the $\Delta H_d$ value for EZnSH A4S is within standard deviation of average value for this state, the $\Delta H_d$ apoSS A4S is considerably lower than the expected average (Table 7.2), which suggests apoSS heats could be reporting on processes other than dimer dissociation. These other processes can include dissociation of aggregates, monomer unfolding, and aggregation in the cell, which can alter the observed heat. Overall, the $\Delta H_d$ values suggest dimer dissociation is endothermic, with the apoSS form having much larger endothermic character than the zinc bound SOD1.

Table 7.1: Summary of ITC measurements of the dimer interface of EZnSH and ApoSS SOD1 mutants at 25 °C and 37°C.

| | EZnSH | | | | | ApoSH | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mutants** | **$K_d$ (μM, 25 °C)** | **Error** | **$K_d$ Rank** | **ΔH (kcal (mol monomer)$^{-1}$)** | **Error** | **$K_d$ (μM, 37 °C)** | **Error** | **$K_d$ Rank** | **ΔH (kcal (mol monomer)$^{-1}$)** | **Error** |
| **pWT** | 50.77 | 40.76 | 8 | 4.01 | 0.92 | 0.07 | 0.05 | 1 | 15.40 | 4.40 |
| **A4S** | 27.74 | 4.74 | 2 | 5.06 | 1.34 | 10.90 | 0.20 | 10 | 4.50 | 1.30 |
| **A4T** | 272.80 | 40.99 | 9 | 3.01 | 0.10 | 10.60 | 3.30 | 9 | 19.60 | 1.90 |
| **A4V** | 1056.00 | 223.49 | 11 | 2.07 | 0.06 | 34.20 | 0.40 | 12 | 18.60 | 1.90 |
| **H43R** | 59.68 | 8.53 | 4 | 3.71 | 0.12 | 5.70 | 0.40 | 6 | 11.50 | 0.70 |
| **H46R** | 322.51 | 506.91 | 12 | 3.40 | 1.07 | 1.40 | 1.10 | 4 | 8.10 | 2.20 |
| **G93A** | 25.21 | 13.96 | 6 | 2.48 | 0.11 | 8.40 | 3.50 | 7 | 7.00 | 1.00 |
| **G93R** | 26.24 | 6.84 | 3 | 2.88 | 0.21 | 20.00 | 4.10 | 11 | 22.80 | 0.90 |
| **G93S** | 52.86 | 15.61 | 7 | 5.11 | 0.92 | 1.30 | 0.50 | 3 | 8.80 | 2.30 |
| **E100G** | 40.00 | 10.00 | 5 | 2.88 | 0.34 | 2.60 | 1.70 | 5 | 8.00 | 2.40 |
| **I113T** | 390.00 | 100.00 | 10 | 4.31 | 0.04 | 10.20 | 3.00 | 8 | 15.10 | 1.25 |
| **V148G** | 126044.54 | 218199.84 | 13 | 11.31 | 9.66 | 76.80 | 34.00 | 13 | 25.30 | 0.70 |
| **V148I** | 39.00 | 4.60 | 1 | 4.12 | 0.11 | 0.50 | 0.20 | 2 | 5.70 | 1.10 |

All data are fit to a dimer dissociation model (Equation 7.1), and data fit values and its error averaged over multiple (>2) experiments. EZnSH V148G is not considered reliable as the first data point included process other than dimer dissociation.

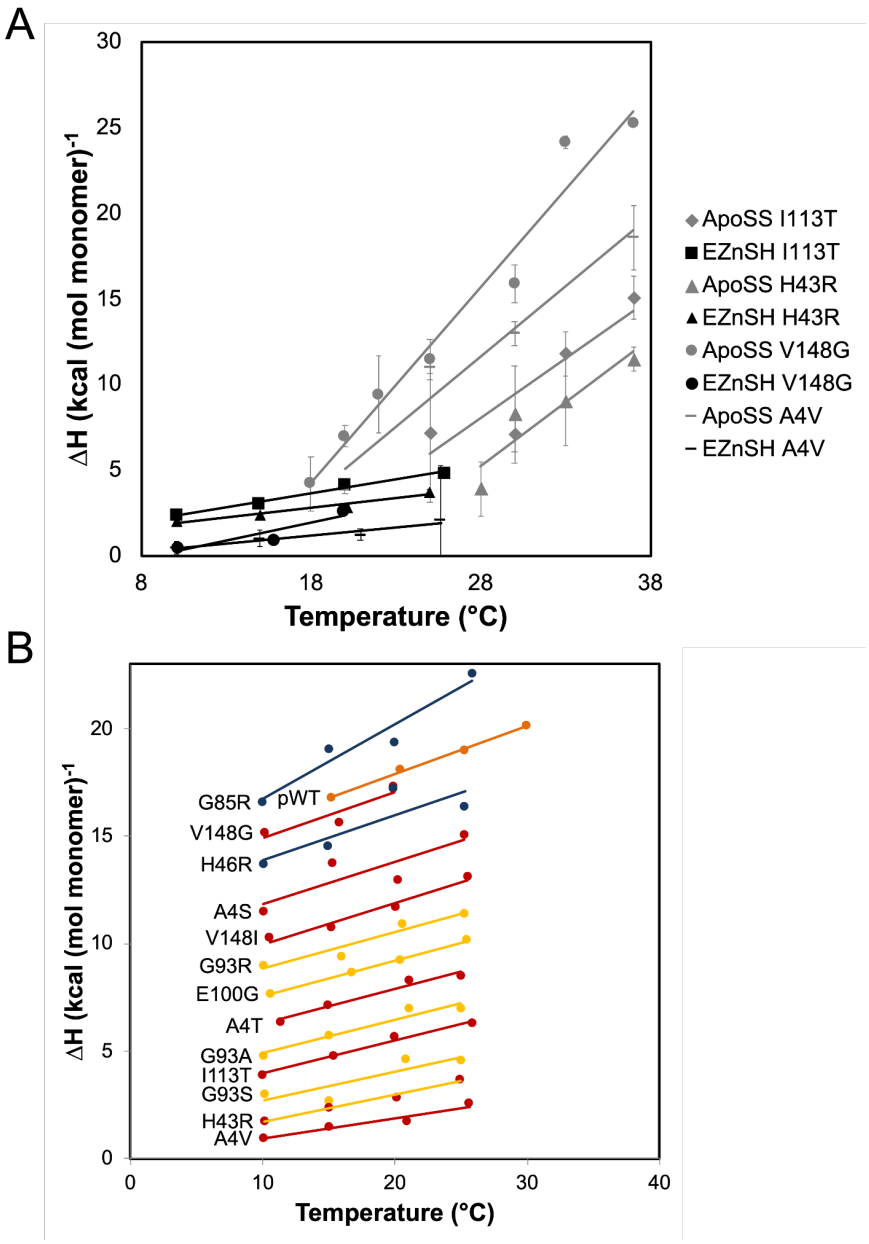



Figure 7.6: Kirchoff plot of $\Delta H_d$ dependence on temperature for EZnSH and ApoSS SOD1. (A) $\Delta H_d$ dependence on temperature of both EZnSH (black) and apoSS (grey) SOD1 are plotted. The $\Delta C_p$, determined from the slope of the data (solid line). Due to the tight dimer interface for apoSS SOD1, only a few mutants are observable at lower temperatures. (B) The $\Delta H_d$ dependence on temperature is assessed for all mutants studied here and coloured based on the location of the mutation. The nonmetal binding (yellow), dimer interface (red) and metal-binding (blue) mutants are shown relative to pWT (orange).

Since the dimer interface is tighter for the apoSS form than the EZnSH SOD1, we hypothesize the apoSS dimer dissociation involves more structural changes than the EZnSH form. The heat capacity of dimer dissociation, $\Delta C_{p,d}$ value represents the extent of conformation changes and/or changes in solvent binding upon dissociation.[312] Here, $\Delta C_{p,d}$ is calculated from the temperature dependence of $\Delta H_d$, also known as the Kirchoff plot. Due to the tight interface of the apoSS dimer, only some mutants (A4V, H43R, I113T, and V148G) had large enough heats of dissociation that were measurable at lower temperature ranges.[87] As hypothesized, the average apoSS SOD1 $\Delta C_{p,d}$ for all measurable mutant was 0.85±0.20 kcal (mol monomer)$^{-1}$, which was much higher than the average $\Delta C_{p,d}$ of 0.18±0.06 kcal (mol monomer)$^{-1}$ for the EZnSH SOD1 (Figure 7.6). Upon dimer dissociation, it appears the dissociation of the tight apoSS SOD1 interface results in significant structural rearrangement and/or change in solvent accessible surface area. However, the $\Delta H_d$ for the weaker EZnSH SOD1 interface has marginal temperature dependence, suggesting the monomer structure is minimally perturbed upon dissociation.

Interestingly, although the average EZnSH SOD1 $\Delta C_{p,d}$ value was ~4 fold smaller than the apoSS $\Delta C_{p,d}$, the value also had minimal differences among mutations unlike the significant variability for apoSS. The EZnSH $\Delta C_{p,d}$ values ranged from 0.10 to 0.35 kcal

(mol monomer)$^{-1}$, a difference of 0.25 kcal (mol monomer)$^{-1}$, while the apoSS values ranged from 0.70 to 1.14 kcal (mol monomer)$^{-1}$, a difference of 0.44 kcal (mol monomer)$^{-1}$ (Figure 7.6). Their respective range of values could be linked to the tightness of their interfaces. For a tight interface, such as for apoSS SOD1, its interface may be a significant structure stabilizer, and upon dissociation, there is significant structural rearrangement in response to the loss of structural stability.[53] On the other hand, a weak interface, like for EZnSH SOD1, may not provide significant structural stability, and in turn the monomers do not require substantial structural change upon dissociation. Overall, the EZnSH dimer dissociation was less endothermic and resulted in less structural changes than observed for the apoSS state.

### 7.4.5 DSC endotherm suggests zinc shuffling from the unfolded EZnSH SOD1 to form Zn,ZnSH SOD1

DSC was used to investigate the global unfolding of EZnSH SOD1 mutants. Previously, DSC endotherms were used to explore the global unfolding of apoSH, apoSS and CuZnSS SOD1 states, and each revealed a single, mostly narrow peak.[45,52,65,87] ApoSH and CuZnSS SOD1 states undergo 2-state unfolding, where the apoSH monomer and CuZnSS dimer unfold into an unfolded monomer. However, apoSS dimer unfolding better fit a 3-state unfolding model at lower concentrations, which involved the apoSS dimer dissociating into a monomer intermediate and then to an unfolded monomer. For the EZnSH SOD1, we hypothesized the unfolding would also be more 3-state for mutants that had pWT-like interfaces and 2-state for mutants that had perturbed interfaces

because they would exist as mostly monomers at the protein concentrations DSC experiments were performed.

Interestingly, neither 2-state or 3-state unfolding explained the EZnSH SOD1 endotherm. Rather than a narrow single peak, the control, EZnSH pWT, consisted of a broad peak with multiple transitions. The EZnSH pWT endotherms were collected at multiple concentrations, 0.68-2.3 mg $mL^{-1}$, which correspond to 34-53 % monomer based on the measured $K_d$ at 25 °C (see section 7.4.2). There appears to be two temperatures at which the maximum of the endotherm occurs, $T_{0.5}$, instead of one (Figure 7.7A). The first $T_{0.5}$ is at ~57 °C and does not change with protein concentration, suggestive of monomer unfolding; however, the second $T_{0.5}$ is at ~62 °C and shows some shift to higher temperatures with increasing protein concentration indicative of dimer-like behaviour. With many theories on this complex endotherm, including perhaps the EZnSH monomer unfolds earlier than the EZnSH dimer and/or there was different zinc bound states, more experiments were required to parse the thermal unfolding observed for the EZnSH pWT.

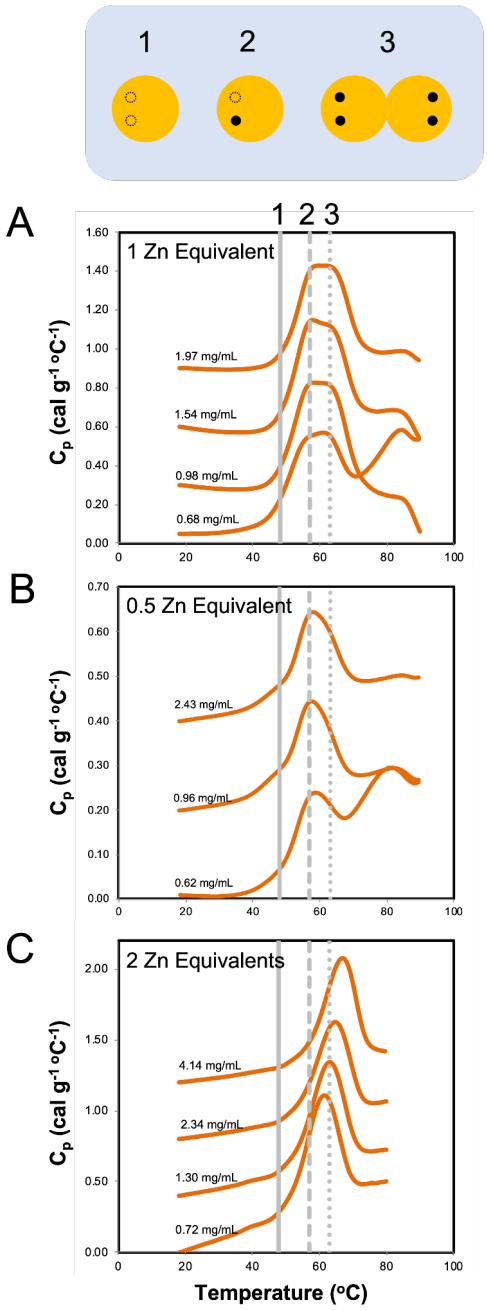



Figure 7.7: DSC of the ApoSH pWT as a function of varying equivalents of Zn. Each scan was offset for clarity, with the lowest concentration being the lower of the multiple scans and the highest concentration being at the top of each plot. The different dashed lines represent the $T_{0.5}$ of different forms, apoSH pWT monomer (grey solid line, obtained from Vassall *et al*., 2011), EZnSH pWT monomer (grey dashed line, maximum of the thermogram for 0.5 Zn equivalents), and ZnZnSH pWT dimer (grey dotted line, maximum of the thermogram for 2 Zn equivalents at 1.30 mg/mL), respectively. (A) Endotherm of 1 Zn equivalence at varying protein concentrations in 20 mM HEPES, 1 mM TCEP, pH 7.4. 1 Zn equivalence concentrations were 0.68 mg/mL, 0.98 mg/mL, 1.54 mg/mL and 2.30 mg/mL. (B) Thermal unfolding of disulfide reduced pWT with half a zinc equivalent at increasing protein concentrations. Half a Zn equivalents concentrations were 1.16 mg/mL, 1.85 mg/mL, and 2.96 mg/mL. (C) Consecutive thermograms of ZnZnSH-pWT are shown. 2 Zn equivalent concentrations included 0.72 mg/mL, 1.30 mg/mL, 2.34 mg/mL and 4.14 mg/mL.

Total protein unfolding is measured at different zinc equivalents relative to SOD1 protein concentration to explain the EZnSH pWT endotherm. Initially, only half an equivalence of zinc was added to protein samples that ranged in concentration from 1.16 to 2.96 mg $mL^{-1}$. In principle, multiple zinc-bound species can exist due to the following: affinity of different SOD1 species for zinc, protein concentration, and zinc concentration. Therefore, the protein sample can include monomers with 0-2 zinc bound, and dimers with 1-4 zinc bound. Due to the strong affinity of zinc for the SOD1 zinc-binding site, in the half zinc equivalent pWT sample, half of the population has one zinc bound (EZnSH) and the other half exists as apo (apoSH). As expected, the thermal unfolding of this mixture resulted in a complex endotherm (Figure 7.7B). The endotherm had a shoulder at ~48 °C that corresponds to the apoSH state since its expected $T_{0.5}$ was previously reported as 47.3 °C.[65] Also interestingly, there is a more prominent maximum to the half zinc equivalent endotherm at ~57 °C that was also observed in the 1 zinc equivalent endotherm above, and the $T_{0.5}$ does not shift with increased protein concentration. This $T_{0.5}$ is indicative of the EZnSH pWT monomer. This implies the dimer dissociation occurs at a lower temperature, and the larger heats of the initial endothermic transition of the EZnSH pWT is from monomer unfolding. There is also the presence of a shoulder at ~62 °C that appears to grow with increasing protein concentration. However, this more thermally stable species unfolding is also present in the 1 zinc equivalent experiments as a prominent peak compared to the shoulder observed in the half equivalent endotherm.

With the initial peak observed in the complex EZnSH pWT endotherm determined, the second, more thermally stable species remained unresolved. Since the initial peak represented the single zinc bound species, this second peak could represent a multi-zinc

bound state. Although the copper-binding site has affinity for zinc, it was clear from previous zinc titration experiments the zinc preferentially binds to all available zinc-binding sites prior to occupying the copper site.[56] With this, and the fact that metal shuffling is not observed in the CuZnSS SOD1 DSC endotherms, we did not consider zinc shuffling a likely outcome. However, upon unfolding pWT samples with 2 zinc equivalents, it is clear the more thermally stable species observed at ~62 °C is ZnZnSH pWT (Figure 7.7C). The single peak endotherm shifted to increasing temperature with increasing protein concentration, indicative of dimer unfolding. In summary, with various zinc equivalent experiments relative to SOD1 concentration, we hypothesize the complex EZnSH pWT endotherm consists of an initial unfolding of the EZnSH monomer and a subsequent zinc shuffling from unfolded EZnSH monomer to a folded EZnSH SOD1 to form a ZnZnSH dimer that unfolds in the latter half of the endotherm.

Interestingly, the fALS mutants showed even more complex behaviour, ranging from pWT-like endotherms to more individual peaks, that ultimately validated the proposed theory of zinc shuffling upon EZnSH monomer unfolding. Unsurprisingly EZnSH V148I endotherm was nearly identical to pWT, with the $T_{0.5}$ at ~57 °C and 62.9 °C for the EZnSH monomer and the ZnZnSH dimer, respectively (Figure 7.8 and Table 7.3). Although markedly more thermally destabilized than pWT and V148I as expected, EZnSH E100G endotherm consisted of a similar shape in terms of a broad endotherm with multiple transitions that were of comparable heights (Figure 7.8). The two transitions occurred at 51 °C and 57.6 °C (Table 7.3), which is the lowest $T_{0.5}$ across the mutant set and is consistent with E100G typically having one of the lowest $T_{0.5}$ in all other maturation

states. Only EZnSH forms of V148I and E100G showed pWT-like endotherms that consist of two transitions at comparable heights.

Table 7.2: Melting temperatures for the various states of SOD1.

| Mutant | ApoSH $T_{0.5}$ | EZnSH $T_{0.5}$ | ZnZnSH $T_{0.5}$ | ApoSS $T_{0.5}$ | HoloSS $T_{0.5}$ |
|---|---|---|---|---|---|
| **pWT** | 47.6 | 57 | 63 | 59.5 | 92.5 |
| **A4S** | 32.5 | 51.3 | 58.4 | 46.3 | 87.8 |
| **A4T** | 30.9 | 55 | 59.9 | 43.6 | 86.2 |
| **A4V** | 36.3 | 59.3 | 64.5 | 50.9 | 88.1 |
| **E100G** | 33.2 | 51 | 57.6 | 48 | 87.3 |
| **G93A** | 34.6 | 57.2 | 59.8 | 47.4 | 88 |
| **G93R** | 35.4 | 57 | 60.5 | 49.1 | 77.6 |
| **G93S** | 33.8 | 47.7 | 55 | 49.2 | 88.5 |
| **G85R** | 40.7 | 53.2 | - | - | - |
| **H43R** | 35.4 | 57.7 | 63 | 47.6 | - |
| **H46R** | 52.6 | 55.33 | - | 62.5 | - |
| **I113T** | 36 | 53 | 55.8 | 46.7 | 88.2 |
| **V148G** | 34.3 | 55.5 | 64.4 | 48.6 | - |
| **V148I** | 51 | 57 | 62.9 | 60.1 | 92.5 |

ApoSH, apoSS and holoSS SOD1 values are obtained from Vassal *et al*., 2011, Broom *et al*., 2013 and Stathopoulos *et al*., 2006. EZnSH values are based on visual observation of the maximum of the transition peak.
- represents the state is not observable or not measured.

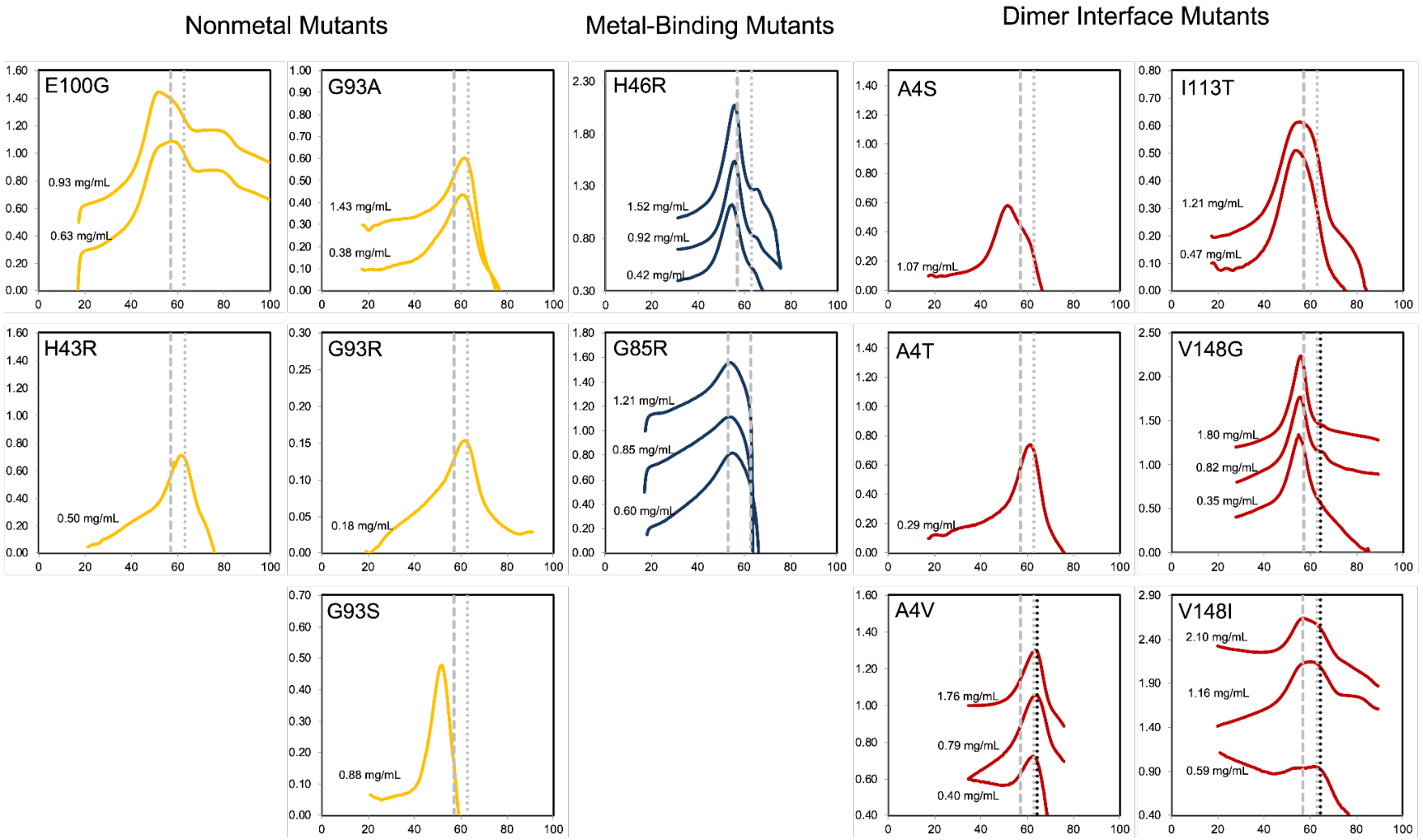



Figure 7.8: Thermogram for EZnSH SOD1 mutants at varying protein concentrations. Each scan is offset for clarity, with the lower scans representing the lower concentrations, to the highest scan belonging to the largest protein concentration. The grey dashed and dotted lines represent the $T_{0.5}$ of different forms of pWT SOD1, EZnSH pWT monomer, and ZnZnSH pWT dimer, respectively. For some mutants, the ZnZnSH SOD1 mutants are measured and represented as black dotted lines. The concentrations are labelled on the side of each plot.

Interestingly, mutants with known aberrant metal-binding further supported the theory on zinc shuffling. Although H43R is near metal-binding sites, it has no known impact on metal-binding, as observed in the zinc titration and dimer dissociation data as discussed above. However, EZnSH H43R endotherm did not look pWT-like, instead it consisted of an unusually broad endotherm, consistent with multiple transitions, with a single maximum (Figure 7.8). Additionally, the difference between $T_{0.5}$ of the two transitions (5.3 °C) is smaller compared to pWT (6 °C), and likely have different heat capacities of unfolding, which results in one broad endotherm with one maximum compared to the two shoulders observed in pWT. Moreover, due to the known metal-binding issues for G85R and H46R, the zinc shuffling may be perturbed. Both metal-binding mutants did not show zinc shuffling, rather only one transition was observed at 53.2 °C for G85R and 55.3 °C for H46R (Figure 7.7B and Table 7.3). Further equivalents of zinc were added to H46R, and still only one transition was observed at 55.3 °C (Table 7.3). The lack of zinc shuffling is expected as the zinc ion likely occupies the only non-perturbed metal-binding site, while the other is perturbed and therefore cannot allow for metal-binding. The metal-binding mutants confirm the second transition observed is from zinc shuffling.

Similar to their diverse dimer strength, A4 mutants also show significantly different thermal stability. Although the order of thermal stability goes from the highest to lowest for A4V, A4S and then A4T in other maturation states, here the thermal stabilities of EZnSH A4 mutants is inversely proportional to their dimer strength (Figure 7.8). A4V and A4T show endotherms that are comparable to H43R, which consists of a broad endotherm with a single maximum. Again, this likely correlates to the two transitions

having more overlap between them than observed for pWT. EZnSH and ZnZnSH states unfold at 55.0 °C and 59.9 °C for A4T, respectively, and 59.3 °C and 64.5 °C for A4V, respectively (Figure 7.8 and Table 7.3). However, the only A4 mutant that formed a dimer comparable to pWT was A4S, but $T_{0.5}$ was 51.3 °C and 58.4 °C for the EZnSH and ZnZnSH state, respectively, both much lower than pWT. However, the two states overlap less and do not show similar heights as observed for pWT, rather the ZnZnSH state had a lower height than the EZnSH transition. Moreover, the EZnSH A4S formed a stronger dimer than the other EZnSH A4 mutants but had the lowest thermal stability. Instead, EZnSH A4V, the weakest dimer among the A4 mutants, had the highest thermal stability, even more than pWT. Overall, the A4 mutants show an inverse relationship between EZnSH dimer strength and thermal stability.

In contrast to the A4 mutants, G93 mutants had comparable dimer strength to pWT, however their thermal stability was much lower. All three mutants showed a single maximum for their broad endotherm, suggesting the multiple transitions are overlapping. Both G93A and G93R had similar $T_{0.5}$ for both states, 57.2°C and 57°C for EZnSH state, respectively, and 59.8 °C and 60.5 °C for the ZnZnSH state, respectively (Figure 7.8 and Table 7.3). However, G93S $T_{0.5}$ was remarkably lower for both transitions, 47.7 °C and 55.0 °C. The trend of G93R having the highest thermal stability and G93S the lowest among the G93 mutants was also observed for the apoSH G93 thermal stabilities. G93 mutants offer insight on the impact of point mutations on monomer stability, considering all mutants had similar $K_d$.

Lastly, the other dimer-interface mutants, I113T and V148G, also showed an increase in thermal stability due to zinc-binding. I113T endotherms showed a shoulder

corresponding to the second transition at increased protein concentration, similar to pWT. The EZnSH and ZnZnSH $T_{0.5}$ was at 53.0 °C and 55.8 °C, which are lower than pWT (Figure 7.8 and Table 7.3). Unlike the zinc shuffling observed for I113T, V148G showed no zinc shuffling. A single transition was observed at 55.5 °C at multiple concentrations. The EZnSH V148G has a similar thermal stability to pWT, which is unlike V148G in other maturation states, where it typically has one of the lowest thermal stabilities. Interestingly, even though EZnSH V148G endotherm showed no zinc shuffling, V148G with multiple zinc equivalents showed a peak at 62.9 °C, likely corresponding to the ZnZnSH state (Table 7.3). Unlike the lack of zinc shuffling observed for metal-binding mutants due to a perturbed metal-binding site impacting shuffling, discussed above, V148G has the capability to bind both metals, but upon unfolding, the zinc does not occupy the second site. Here, we must consider all the factors that affect the observed endotherm.

The EZnSH endotherms vary to a great degree; some endotherms include two transitions that are of comparable heights resulting in two shoulders, others consist of two overlapping transitions with one maximum, and some mutants show only one transition with no observable zinc shuffling. If the transitions have similar $T_{0.5}$ and unfolding heats, the transitions will appear less distinct and result in a single broad peak, as observed for H43R, A4T, A4V, I113T, and G93 mutants. However, if the $T_{0.5}$ and unfolding heats are more distinct, it will result in more distinguishable transitions, as observed for pWT, V148I, E100G and A4S. There are multiple factors to consider that impact the observed endotherm, many surrounding the zinc shuffling. For the zinc to shuffle from the unfolded EZnSH monomer to a folded EZnSH monomer, the second metal-binding site cannot be perturbed as it often is in metal-binding mutants like G85R and H46R. Two other important

factors to consider that impact zinc shuffling is the affinity of the zinc at the copper site may vary based on the mutation, and also the dimer strength of the ZnZnSH state will impact the unfolding of the second transition. Lastly, the observed endotherm can also be impacted by the unfolding transition of each state. If either transition unfolds with 2-state unfolding or 3-state unfolding with a monomer intermediate, it will impact the endotherm and the equilibrium of EZnSH and ZnZnSH states observed. Thus, parsing this complex endotherm cannot be done without considering the multiple factors that impact the observed transitions.

### 7.4.6 The dimer interface and thermal stability of EZnSH SOD1 is functionally relevant

Upon addition of zinc, the monomer is significantly stabilized relative to apoSH SOD1[65], but the dimer interface is weaker than the apoSS SOD1.[53] Both differences are likely functionally important for EZnSH SOD1's cellular role. For instance, at cellular concentrations the protein likely exists predominantly as a monomer in this state which means it is accessible to bind to the CCS, but also leaves it susceptible to aggregation. In addition to the weakened dimer interface being important for binding the CCS, the ability to transiently dimerize can protect EZnSH SOD1 against protein aggregation. For the EZnSH SOD1 that exists as a monomer, its enhanced thermal stability relative to apoSH SOD1 favours the folded conformation under cellular conditions, further preventing the changes of aggregation. Therefore, both the enhanced thermal stability and the weakened dimer interface serve protective roles against aggregation.

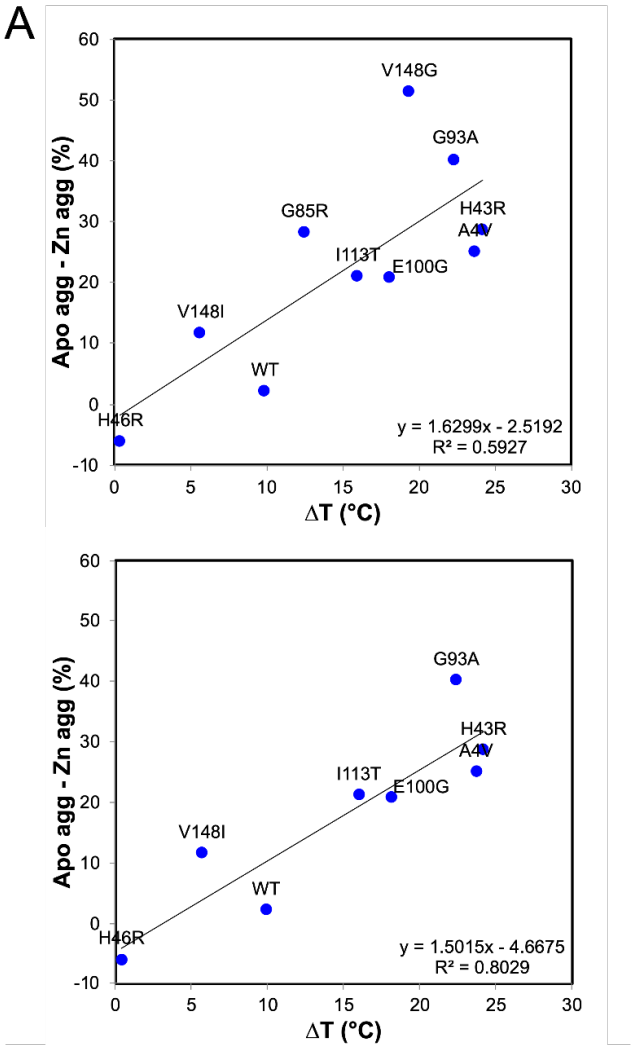
Although there is conflicting data on the CCS's role and which species the CCS binds to, our data provide critical insight on deciphering the conflicting data. Since the

CCS binds to the SOD1 monomer using the SOD1 dimer interface, a species with an ability to dimerize is critical. If apoSH SOD1 cannot readily dimerize, it is unlikely that the CCS can bind the SOD1. However, there is support for the idea that CCS binding to apoSH SOD1 and aiding in its ability to bind zinc.[43] Even though this seems not likely due to apoSH SOD1's inability to dimerize, there are certain situations where dimerization may be possible. For instance, if the cellular concentration of apoSH SOD1 is high, then the equilibrium will shift toward binding the CCS. Then with an apoSH-SOD1:CCS dimer, the zinc affinity likely increases due to increased structuring from dimer formation, as our data suggest (section 7.4.1). Therefore, the preferred species for CCS binding is the EZnSH SOD1 due to its ability to dimerize, under specific conditions, it is considerable the apoSH SOD1 can also dimerize with the CCS.

Moreover, the EZnSH SOD1 mutants show diverse thermal stability and dimer strength, which likely impairs their maturation and enhances protein association. A4V has a significantly weakened dimer interface, in addition to its perturbed native structure (see Chapter 3) and both factors likely explain its inability to bind to the CCS.[56] Similarly, any mutants that also show a weakened dimer interface should likely have perturbed affinity for the CCS, and therefore disrupting its maturation. Furthermore, *in vivo* SOD1 aggregation correlates with extent of protein unfolding, and aggregation rescue upon the addition of zinc correlates to thermal stabilization upon zinc binding.[25] For instance, the samples with more rescue often have a higher increase in thermal stability relative to other mutants. Overall, both the thermal stability and dimer interface strength of EZnSH SOD1 have consequences on maturation and protein aggregation.

### 7.4.7 Species without zinc may be a critical in SOD1 aggregation

Considering the zinc ion offers structural stability, the inability to bind zinc or the loss of zinc will understandably lead to a less structured SOD1 that is susceptible to aggregation. The zinc-deficient SOD1 mutants show decreased structure in both the metal-binding loop and the electrostatic loops.[215] This loss of structure leads to the edge-strands interacting to form crystals, specifically in H46R and H43R.[59,224,313] To further support that the zinc ions offer enough structural stability to shift the mutants away from aggregation, *in vivo* aggregation experiments are compared to the gain of thermal stability due to the zinc ion reported above.[25] In vivo aggregation experiments measured the extent of aggregation both without and with zinc presence. The rescue in aggregation upon zinc addition correlates with the gain in thermal stability from zinc addition to apoSH SOD1 (Figure 7.9A). The $R^2$ is 0.60 with all mutants considered, however removing G85R and V148G due to their perturbed thermograms increases the $R^2$ to 0.80 (Figure 7.9A). This suggests the monomer stability is critical in preventing the aggregation observed in this study.[25]


A
Apo agg - Zn agg (%)
ΔT (°C)
V148G
G93A
G85R
H43R
A4V
I113T
E100G
V148I
WT
H46R
y = 1.6299x - 2.5192
R² = 0.5927
y = 1.5015x - 4.6675
R² = 0.8029

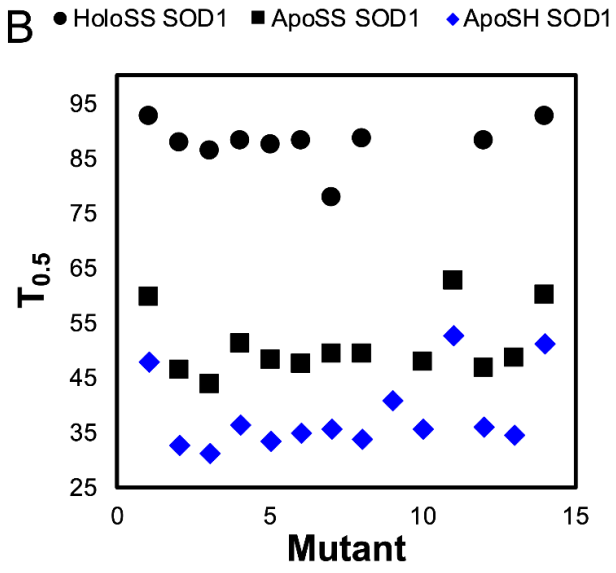

B
HoloSS SOD1
ApoSS SOD1
ApoSH SOD1
T0.5
Mutant

Figure 7.9: Aggregation rescue due to addition of zinc and the range of thermal stabilities for different SOD1 states. (A) The decrease in aggregation due to the addition of zinc is plotted against the gain in thermal stability between ApoSH and EZnSH SOD1. The top panel includes all mutants, and the bottom includes all mutants except for G85R and V148G due to their perturbed DSC endotherms, both are fit to a linear model (black solid line). (B) The $T_{0.5}$ are plotted for different SOD1 states to show the decrease in temperature range across mutants due to increased maturation.

While the EZnSH derives most of its stability from the stable monomer and not the weak dimer interface, which contrasts the apoSS SOD1 that gains more stability from its dimer[87], the rescue of aggregation due to zinc should correlate to monomer thermal stability and not dimer interface strength. Moreover, the decrease in thermal stability due to point mutations appears to lessen with the addition of maturation steps (Figure 7.9B). For instance, the apoSH mutants show the largest range in melting temperatures, which decreases for apoSS SOD1 and even more so for holoSS SOD1 (Figure 7.9B).[52,65,87] This suggests that the maturation steps offer enough stability to compensate for the destabilizing impact of any point mutation. Therefore, zinc binding is critical for increasing structural stability to help prevent aggregation.

## 7.5 Conclusion

Due to strong evidence suggesting the more immature forms of SOD1 are responsible for disease aggregation, understanding how mutations impair maturation is imperative. Here we analyze EZnSH SOD1 using a combination of isothermal titration calorimetry (ITC), to measure dimer dissociation as well as metal binding, with complementary differential scanning calorimetry (DSC), to measure global protein unfolding. Interestingly the mutations in EZnSH SOD1 had either a destabilizing effect or no effect on the dimer interface, which contrasts the exclusively destabilizing effect of all mutations on the apoSS SOD1. The mutations that show destabilized dimer interfaces

may have a direct consequence in further maturation of SOD1, as the zinc bound form is key in binding to the copper chaperone (CCS), which catalyzes formation of mature holo SOD1. Additionally, our data support that the zinc bound species is more thermal stable than the apoSH state, and this gain in stability is important in preventing SOD1 aggregation. These findings help rationalize maturation of SOD1 and potential roles for monomer forms of SOD1 in the neurotoxic aggregation linked to ALS.

# Chapter 8: Future Work

With that the data presented above, we have a well-rounded idea on both the subtle and major differences between apoSH SOD1 mutants as well as between solution conditions. With SOD1 showing significant aggregate structural polymorphisms, it is clear the mutants behave differently. Additionally, the data presented here, in addition to prior NMR data, show the apoSH SOD1 show local differences in dynamics, which may be responsible for differences in aggregation pathways or mechanisms that ultimately result in different aggregate characteristics. Now, the question remains, with these local regions of difference, how do they influence aggregation initiation? Are these regions responsible for initiating, templating, changing the folded-unfolded equilibrium, or impacting maturation?

To attempt to answer these questions, additional experiments are necessary. A vital experiment that may be important is to investigate how the local hotspots identified here impact our subunits. If HDX experiments are performed over changing conditions (e.g., protein concentration, salt concentration and/or with aggregates), the experiments may reveal transient interactions with other regions of the subunits. Additionally, using NMR experiments, such as TROSY experiments, which allow for us to explore self-association, can help decipher which segments of the protein are critical in initiating aggregation and which regions, if different, are vital in growing the aggregate. Although the 2D HSQC gave a wealth of information, as presented above, we can gain additional information with more complex NMR experiments.

Moreover, we also present the consequences a mutation can have on the dimer interface which in turn can complicate maturation. With these findings, it is now important

to decipher if all mutations with perturbed dimer interfaces also have perturbed CCS binding and subsequently maturation. With more experiments on the strength of CCS binding to ALS-associated mutant and the extent of maturation, we may be able to find potential reasons for the pools of immature forms of SOD1 found in disease-expressing cells.[309]

Lastly, the solution conditions impact the SOD1 aggregate structures observed. The need for more systematic assessment of these structures and probing for additional possible mechanisms is vital for us to understand what influences proteins down a certain pathway. Here, we have presented a look at SOD1 aggregates under different protein concentration, and different salts and salt concentrations across many mutants. To build on this framework, it would be interesting to now mimic intracellular crowding conditions, for instance introduce a popular crowding agent like Ficoll and polyethylene glycol (PEG), which when introduced, shows proteins behave comparable to those in cell lysate.[314] Collectively with this framework provided, it is essential to further probe SOD1 mutant dynamics, and their impact on maturation and protein aggregation, because these data may ultimately help elucidate the complexities driving protein aggregation.

# References:


(1) Dobson, C. M. Protein Folding and Misfolding. *Nature 2003 426:6968* **2003**, *426* (6968), 884–890. https://doi.org/10.1038/nature02261.

(2) Dobson, C.; Sali, A.; Karplus, M. Protein Folding: A Perspective from Theory and Experiment. *Angew Chem Int Ed Engl* **1998**. https://doi.org/10.1002/(SICI)1521-3773(19980420)37:7.

(3) Dinner, A. R.; Šalib, A.; Smith, L. J.; Dobson, C. M.; Karplus, M. Understanding Protein Folding via Free-Energy Surfaces from Theory and Experiment. *Trends Biochem Sci* **2000**, *25* (7), 331–339. https://doi.org/10.1016/S0968-0004(00)01610-8.

(4) Toal, S.; Schweitzer-Stenner, R. Local Order in the Unfolded State: Conformational Biases and Nearest Neighbor Interactions. *Biomolecules* **2014**, *4* (3), 725. https://doi.org/10.3390/BIOM4030725.

(5) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. Navigating the Folding Routes. *Science* **1995**, *267* (5204), 1619–1620. https://doi.org/10.1126/SCIENCE.7886447.

(6) Dill, K. A.; Chan, H. S. From Levinthal to Pathways to Funnels. *Nature Structural Biology 1997 4:1* **1997**, *4* (1), 10–19. https://doi.org/10.1038/nsb0197-10.

(7) Hartl, F. U.; Hayer-Hartl, M. Molecular Chaperones in the Cytosol: From Nascent Chain to Folded Protein. *Science* **2002**, *295* (5561), 1852–1858. https://doi.org/10.1126/SCIENCE.1068408.

(8) Kaufman, R. J.; Scheuner, D.; Schröder, M.; Shen, X.; Lee, K.; Liu, C. Y.; Arnold, S. M. The Unfolded Protein Response in Nutrient Sensing and Differentiation. *Nat Rev Mol Cell Biol* **2002**, *3* (6), 411–421. https://doi.org/10.1038/NRM829.

(9) True, H. L.; Lindquist, S. L. A Yeast Prion Provides a Mechanism for Genetic Variation and Phenotypic Diversity. *Nature* **2000**, *407* (6803), 477–483. https://doi.org/10.1038/35035005.

(10) Chapman, M. R.; Robinson, L. S.; Pinkner, J. S.; Roth, R.; Heuser, J.; Hammar, M.; Normark, S.; Hultgren, S. J. Role of Escherichia Coli Curli Operons in Directing Amyloid Fiber Formation. *Science* **2002**, *295* (5556), 851–855. https://doi.org/10.1126/SCIENCE.1067484.

(11) Kelly, J. W.; Balch, W. E. Amyloid as a Natural Product. *J Cell Biol* **2003**, *161* (3), 461. https://doi.org/10.1083/JCB.200304074.

(12) Deol, H. K.; Broom, H. R.; Siebeneichler, B.; Lee, B.; Leonenko, Z.; Meiering, E. M. Immature ALS-Associated Mutant Superoxide Dismutases Form Variable Aggregate Structures through Distinct Oligomerization Processes. *Biophys Chem* **2022**, *288*, 106844. https://doi.org/10.1016/J.BPC.2022.106844.

(13) Chiti, F.; Dobson, C. M. Amyloid Formation by Globular Proteins under Native Conditions. *Nat Chem Biol* **2009**, *5* (1), 15–22. https://doi.org/10.1038/NCHEMBIO.131.

(14) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem* **2017**, *86*, 27–68. https://doi.org/10.1146/ANNUREV-BIOCHEM-061516-045115.

(15) Sengupta, I.; Udgaonkar, J. B. Structural Mechanisms of Oligomer and Amyloid Fibril Formation by the Prion Protein. *Chemical Communications* **2018**, *54* (49), 6230–6242. https://doi.org/10.1039/C8CC03053G.

(16) Jahn, T. R.; Radford, S. E. Folding versus Aggregation: Polypeptide Conformations on Competing Pathways. *Arch Biochem Biophys* **2008**, *469* (1), 100–117. https://doi.org/10.1016/J.ABB.2007.05.015.

(17) Gray, F.; Chrétien, F.; Adle-Biassette, H.; Dorandeu, A.; Ereau, T.; Delisle, M. B.; Kopp, N.; Ironside, J. W.; Vital, C. Neuronal Apoptosis in Creutzfeldt-Jakob Disease. *J Neuropathol Exp Neurol* **1999**, *58* (4), 321–328. https://doi.org/10.1097/00005072-199904000-00002.

(18) Loo, D. T.; Copani, A.; Pike, C. J.; Whittemore, E. R.; Walencewicz, A. J.; Cotman, C. W. Apoptosis Is Induced by Beta-Amyloid in Cultured Central Nervous System Neurons. *Proc Natl Acad Sci U S A* **1993**, *90* (17), 7951–7955. https://doi.org/10.1073/PNAS.90.17.7951.

(19) Forloni, G.; Angeretti, N.; Chiesa, R.; Monzani, E.; Salmona, M.; Bugiani, O.; Tagliavini, F. Neurotoxicity of a Prion Protein Fragment. *Nature* **1993**, *362* (6420), 543–546. https://doi.org/10.1038/362543A0.

(20) El-Agnaf, O. M. A.; Jakes, R.; Curran, M. D.; Middleton, D.; Ingenito, R.; Bianchi, E.; Pessi, A.; Neill, D.; Wallace, A. Aggregates from Mutant and Wild-Type Alpha-Synuclein Proteins and NAC Peptide Induce Apoptotic Cell Death in Human Neuroblastoma Cells by Formation of Beta-Sheet and Amyloid-like Filaments. *FEBS Lett* **1998**, *440* (1–2), 71–75. https://doi.org/10.1016/S0014-5793(98)01418-5.

(21) Broom, H. R.; Rumfeldt, J. A. O.; Meiering, E. M. Many Roads Lead to Rome? Multiple Modes of Cu,Zn Superoxide Dismutase Destabilization, Misfolding and Aggregation in Amyotrophic Lateral Sclerosis. *Essays Biochem* **2014**, *56* (1), 149–165. https://doi.org/10.1042/BSE0560149.

(22) Wright, G. S. A.; Antonyuk, S. v.; Hasnain, S. S. The Biophysics of Superoxide Dismutase-1 and Amyotrophic Lateral Sclerosis. *Q Rev Biophys* **2019**, *52*, e12. https://doi.org/10.1017/S003358351900012X.

(23) Karamanos, T. K.; Jackson, M. P.; Calabrese, A. N.; Goodchild, S. C.; Cawood, E. E.; Thompson, G. S.; Kalverda, A. P.; Hewitt, E. W.; Radford, S. E. Structural Mapping of Oligomeric Intermediates in an Amyloid Assembly Pathway. *Elife* **2019**, *8*. https://doi.org/10.7554/ELIFE.46574.001.

(24) Lang, L.; Kurnik, M.; Danielsson, J.; Oliveberg, M. Fibrillation Precursor of Superoxide Dismutase 1 Revealed by Gradual Tuning of the Protein-Folding Equilibrium. *Proc Natl Acad Sci U S A* **2012**, *109* (44), 17868–17873. https://doi.org/10.1073/PNAS.1201795109/.

(25) Naser, D. Cellular Inclusion Bodies: Structure and Mechanisms. **2022**.

(26) Khurana, R.; Gillespie, J. R.; Talapatra, A.; Minert, L. J.; Ionescu-Zanetti, C.; Millett, I.; Fink, A. L. Partially Folded Intermediates as Critical Precursors of Light Chain Amyloid Fibrils and Amorphous Aggregates. *Biochemistry* **2001**, *40* (12), 3525–3535. https://doi.org/10.1021/BI001782B.

(27) Qin, Z.; Hu, D.; Zhu, M.; Fink, A. L. Structural Characterization of the Partially Folded Intermediates of an Immunoglobulin Light Chain Leading to Amyloid

Fibrillation and Amorphous Aggregation. *Biochemistry* **2007**, *46* (11), 3521–3531. https://doi.org/10.1021/BI061716V.

(28) Ayers, J. I.; Fromholt, S. E.; O'Neal, V. M.; Diamond, J. H.; Borchelt, D. R. Prion-like Propagation of Mutant SOD1 Misfolding and Motor Neuron Disease Spread along Neuroanatomical Pathways. *Acta Neuropathol* **2016**, *131* (1), 103–114. https://doi.org/10.1007/S00401-015-1514-0.

(29) Ayers, J. I.; Fromholt, S.; Koch, M.; DeBosier, A.; McMahon, B.; Xu, G.; Borchelt, D. R. Experimental Transmissibility of Mutant SOD1 Motor Neuron Disease. *Acta Neuropathol* **2014**, *128* (6), 791–803. https://doi.org/10.1007/S00401-014-1342-7.

(30) Breijyeh, Z.; Karaman, R. Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules* **2020**, *25* (24). https://doi.org/10.3390/MOLECULES25245789.

(31) Armstrong, M. J.; Okun, M. S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* **2020**, *323* (6), 548–560. https://doi.org/10.1001/JAMA.2019.22360.

(32) Hayashi, Y.; Homma, K.; Ichijo, H. SOD1 in Neurotoxicity and Its Controversial Roles in SOD1 Mutation-Negative ALS. *Advances in Biological Regulation*. 2016. https://doi.org/10.1016/j.jbior.2015.10.006.

(33) Joyce, P. I.; Mcgoldrick, P.; Saccon, R. A.; Weber, W.; Fratta, P.; West, S. J.; Zhu, N.; Carter, S.; Phatak, V.; Stewart, M.; Simon, M.; Kumar, S.; Heise, I.; Bros-Facer, V.; Dick, J.; Corrochano, S.; Stanford, M. J.; Luong, T. V.; Nolan, P. M.; Meyer, T.; Brandner, S.; Bennett, D. L. H.; Hande Ozdinler, P.; Greensmith, L.; Fisher, E. M. C.; Acevedo-Arozena, A. A Novel SOD1-ALS Mutation Separates Central and

Peripheral Effects of Mutant SOD1 Toxicity. *Hum Mol Genet* **2014**. https://doi.org/10.1093/hmg/ddu605.

(34) Karch, C. M.; Prudencio, M.; Winkler, D. D.; Hart, P. J.; Borchelt, D. R. Role of Mutant SOD1 Disulfide Oxidation and Aggregation in the Pathogenesis of Familial ALS. *Proc Natl Acad Sci U S A* **2009**, *106* (19), 7774–7779. https://doi.org/10.1073/PNAS.0902505106.

(35) Hardiman, O.; Al-Chalabi, A.; Chio, A.; Corr, E. M.; Logroscino, G.; Robberecht, W.; Shaw, P. J.; Simmons, Z.; van den Berg, L. H. Amyotrophic Lateral Sclerosis. *Nature Reviews Disease Primers 2017 3:1* **2017**, *3* (1), 1–19. https://doi.org/10.1038/nrdp.2017.71.

(36) Masrori, P.; van Damme, P. Amyotrophic Lateral Sclerosis: A Clinical Review. *Eur J Neurol* **2020**, *27* (10), 1918–1929. https://doi.org/10.1111/ENE.14393.

(37) Grad, L. I.; Rouleau, G. A.; Ravits, J.; Cashman, N. R. Clinical Spectrum of Amyotrophic Lateral Sclerosis (ALS). *Cold Spring Harb Perspect Med* **2017**, *7* (8). https://doi.org/10.1101/CSHPERSPECT.A024117.

(38) Robberecht, W.; Philips, T. The Changing Scene of Amyotrophic Lateral Sclerosis. *Nat Rev Neurosci* **2013**, *14* (4), 248–264. https://doi.org/10.1038/NRN3430.

(39) Zelko, I. N.; Mariani, T. J.; Folz, R. J. Superoxide Dismutase Multigene Family: A Comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) Gene Structures, Evolution, and Expression. *Free Radic Biol Med* **2002**, *33* (3), 337–349. https://doi.org/10.1016/S0891-5849(02)00905-X.

(40) Tainer, J. A.; Getzoff, E. D.; Richardson, J. S.; Richardson, D. C. Structure and Mechanism of Copper, Zinc Superoxide Dismutase. *Nature* **1983**, *306* (5940), 284–287. https://doi.org/10.1038/306284A0.

(41) Getzoff, E. D.; Tainer, J. A.; Stempien, M. M.; Bell, G. I.; Hallewell, R. A. Evolution of CuZn Superoxide Dismutase and the Greek Key Beta-Barrel Structural Motif. *Proteins* **1989**, *5* (4), 322–336. https://doi.org/10.1002/PROT.340050408.

(42) Goodsell, D. S.; Olson, A. J. Structural Symmetry and Protein Function. *Annu Rev Biophys Biomol Struct* **2000**, *29*, 105–153. https://doi.org/10.1146/ANNUREV.BIOPHYS.29.1.105.

(43) Luchinat, E.; Barbieri, L.; Banci, L. A Molecular Chaperone Activity of CCS Restores the Maturation of SOD1 FALS Mutants. *Scientific Reports 2017 7:1* **2017**, *7* (1), 1–8. https://doi.org/10.1038/s41598-017-17815-y.

(44) Potter, S. Z.; Zhu, H.; Shaw, B. F.; Rodriguez, J. A.; Doucette, P. A.; Sohn, S. H.; Durazo, A.; Faull, K. F.; Gralla, E. B.; Nersissian, A. M.; Valentine, J. S. Binding of a Single Zinc Ion to One Subunit of Copper−Zinc Superoxide Dismutase Apoprotein Substantially Influences the Structure and Stability of the Entire Homodimeric Protein. *J Am Chem Soc* **2007**, *129* (15), 4575–4583. https://doi.org/10.1021/ja066690+.

(45) Vassall, K. A.; Stathopulos, P. B.; Rumfeldt, J. A. O.; Lepock, J. R.; Meiering, E. M. Equilibrium Thermodynamic Analysis of Amyotrophic Lateral Sclerosis-Associated Mutant Apo Cu,Zn Superoxide Dismutases. *Biochemistry* **2006**, *45* (23), 7366–7379. https://doi.org/10.1021/BI0600953.

(46) Banci, L.; Bertini, I.; Cantini, F.; Kozyreva, T.; Massagni, C.; Palumaa, P.; Rubino, J. T.; Zovo, K. Human Superoxide Dismutase 1 (HSOD1) Maturation through Interaction with Human Copper Chaperone for SOD1 (HCCS). *Proc Natl Acad Sci U S A* **2012**, *109* (34), 13555–13560. https://doi.org/10.1073/PNAS.1207493109.

(47) Tsang, C. K. wan; Liu, Y.; Thomas, J.; Zhang, Y.; Zheng, X. F. S. Superoxide Dismutase 1 Acts as a Nuclear Transcription Factor to Regulate Oxidative Stress Resistance. *Nature Communications 2014 5:1* **2014**, *5* (1), 1–11. https://doi.org/10.1038/ncomms4446.

(48) da Ros, M.; Deol, H. K.; Savard, A.; Guo, H.; Meiering, E. M.; Gibbings, D. Wild-Type and Mutant SOD1 Localizes to RNA-Rich Structures in Cells and Mice but Does Not Bind RNA. *J Neurochem* **2021**, *156* (4), 524–538. https://doi.org/10.1111/JNC.15126.

(49) Cohen, N. R.; Zitzewitz, J. A.; Bilsel, O.; Robert Matthews, C. Nonnative Structure in a Peptide Model of the Unfolded State of Superoxide Dismutase 1 (SOD1): Implications for ALS-Linked Aggregation. *J Biol Chem* **2019**, *294* (37), 13708. https://doi.org/10.1074/JBC.RA119.008765.

(50) Arosio, P.; Knowles, T. P. J.; Linse, S. On the Lag Phase in Amyloid Fibril Formation. *Physical Chemistry Chemical Physics* **2015**, *17* (12), 7606–7618. https://doi.org/10.1039/C4CP05563B.

(51) Zhu, C.; Beck, M. v.; Griffith, J. D.; Deshmukh, M.; Dokholyan, N. v. Large SOD1 Aggregates, Unlike Trimeric SOD1, Do Not Impact Cell Viability in a Model of Amyotrophic Lateral Sclerosis. *Proc Natl Acad Sci U S A* **2018**, *115* (18), 4661–4665. https://doi.org/10.1073/PNAS.1800187115/.

(52) Stathopulos, P. B.; Rumfeldt, J. A. O.; Karbassi, F.; Siddall, C. A.; Lepock, J. R.; Meiering, E. M. Calorimetric Analysis of Thermodynamic Stability and Aggregation for Apo and Holo Amyotrophic Lateral Sclerosis-Associated Gly-93 Mutants of Superoxide Dismutase. *J Biol Chem* **2006**, *281* (10), 6184–6193. https://doi.org/10.1074/JBC.M509496200.

(53) Broom, H. R.; Rumfeldt, J. A. O.; Vassall, K. A.; Meiering, E. M. Destabilization of the Dimer Interface Is a Common Consequence of Diverse ALS-associated Mutations in Metal Free SOD1. *Protein Sci* **2015**, *24* (12), 2081. https://doi.org/10.1002/PRO.2803.

(54) Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Bouvignies, G.; Meiering, E. M.; Kay, L. E. Thermal Fluctuations of Immature SOD1 Lead to Separate Folding and Misfolding Pathways. *Elife* **2015**, *4*. https://doi.org/10.7554/ELIFE.07296.

(55) Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Sobering, R. E.; Meiering, E. M.; Kay, L. E. Probing the Free Energy Landscapes of ALS Disease Mutants of SOD1 by NMR Spectroscopy. *Proc Natl Acad Sci U S A* **2016**, *113* (45), E6939–E6945. https://doi.org/10.1073/PNAS.1611418113.

(56) Culik, R. M.; Sekhar, A.; Nagesh, J.; Deol, H.; Rumfeldt, J. A. O.; Meiering, E. M.; Kay, L. E. Effects of Maturation on the Conformational Free-Energy Landscape of SOD1. *Proc Natl Acad Sci U S A* **2018**, *115* (11), E2546–E2555. https://doi.org/10.1073/PNAS.1721022115/.

(57) Naser, D.; Tarasca, M. v.; Siebeneichler, B.; Schaefer, A.; Deol, H. K.; Soule, T. G. B.; Almey, J.; Kelso, S.; Mishra, G. G.; Simon, H.; Meiering, E. M. High-Resolution NMR H/D Exchange of Human Superoxide Dismutase Inclusion Bodies Reveals

Significant Native Features Despite Structural Heterogeneity. *Angewandte Chemie* **2022**, *134* (24), e202112645. https://doi.org/10.1002/ANGE.202112645.

(58) Chattopadhyay, M.; Durazo, A.; Se, H. S.; Strong, C. D.; Gralla, E. B.; Whitelegge, J. P.; Valentine, J. S. Initiation and Elongation in Fibrillation of ALS-Linked Superoxide Dismutase. *Proc Natl Acad Sci U S A* **2008**, *105* (48), 18663–18668. https://doi.org/10.1073/PNAS.0807058105.

(59) Elam, J. S.; Taylor, A. B.; Strange, R.; Antonyuk, S.; Doucette, P. A.; Rodriguez, J. A.; Hasnain, S. S.; Hayward, L. J.; Valentine, J. S.; Yeates, T. O.; Hart, P. J. Amyloid-like Filaments and Water-Filled Nanotubes Formed by SOD1 Mutant Proteins Linked to Familial ALS. *Nat Struct Biol* **2003**, *10* (6), 461–467. https://doi.org/10.1038/NSB935.

(60) Sangwan, S.; Sawaya, M. R.; Murray, K. A.; Hughes, M. P.; Eisenberg, D. S. Atomic Structures of Corkscrew-forming Segments of SOD1 Reveal Varied Oligomer Conformations. *Protein Sci* **2018**, *27* (7), 1231. https://doi.org/10.1002/PRO.3391.

(61) Uversky, V. N. Mysterious Oligomerization of the Amyloidogenic Proteins. *FEBS J* **2010**, *277* (14), 2940–2953. https://doi.org/10.1111/J.1742-4658.2010.07721.X.

(62) Kodali, R.; Wetzel, R. Polymorphism in the Intermediates and Products of Amyloid Assembly. *Curr Opin Struct Biol* **2007**, *17* (1), 48–57. https://doi.org/10.1016/J.SBI.2007.01.007.

(63) Rubin, J.; Khosravi, H.; Bruce, K. L.; Lydon, M. E.; Behrens, S. H.; Chernoff, Y. O.; Bommarius, A. S. Ion-Specific Effects on Prion Nucleation and Strain Formation. *J Biol Chem* **2013**, *288* (42), 30300. https://doi.org/10.1074/JBC.M113.467829.

(64) Sandelin, E.; Nordlund, A.; Andersen, P. M.; Marklund, S. S. L.; Oliveberg, M. Amyotrophic Lateral Sclerosis-Associated Copper/Zinc Superoxide Dismutase Mutations Preferentially Reduce the Repulsive Charge of the Proteins. *J Biol Chem* **2007**, *282* (29), 21230–21236. https://doi.org/10.1074/JBC.M700765200.

(65) Vassall, K. A.; Stubbs, H. R.; Primmer, H. A.; Tong, M. S.; Sullivan, S. M.; Sobering, R.; Srinivasan, S.; Briere, L. A. K.; Dunn, S. D.; Colón, W.; Meiering, E. M. Decreased Stability and Increased Formation of Soluble Aggregates by Immature Superoxide Dismutase Do Not Account for Disease Severity in ALS. *Proc Natl Acad Sci U S A* **2011**, *108* (6), 2210–2215. https://doi.org/10.1073/PNAS.0913021108/.

(66) Doyle, C. M.; Rumfeldt, J. A.; Broom, H. R.; Sekhar, A.; Kay, L. E.; Meiering, E. M. Concurrent Increases and Decreases in Local Stability and Conformational Heterogeneity in Cu, Zn Superoxide Dismutase Variants Revealed by Temperature-Dependence of Amide Chemical Shifts. *Biochemistry* **2016**, *55* (9), 1346–1361. https://doi.org/10.1021/ACS.BIOCHEM.5B01133.

(67) Hörnberg, A.; Logan, D. T.; Marklund, S. L.; Oliveberg, M. The Coupling between Disulphide Status, Metallation and Dimer Interface Strength in Cu/Zn Superoxide Dismutase. *J Mol Biol* **2007**, *365* (2), 333–342. https://doi.org/10.1016/J.JMB.2006.09.048.

(68) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys Rev* **2016**, *8* (4), 409–427. https://doi.org/10.1007/S12551-016-0218-6.

(69) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse

Field Gradient NMR Techniques. *Biochemistry* **1999**, *38* (50), 16424–16431. https://doi.org/10.1021/BI991765Q.

(70) Johnson, D.; Hilal, N.; Bowen, W. R. Basic Principles of Atomic Force Microscopy. *Atomic Force Microscopy in Process Engineering* **2009**, 1–30. https://doi.org/10.1016/B978-1-85617-517-3.00001-8.

(71) Gaczynska, M.; Karpowicz, P.; Stuart, C. E.; Norton, M. G.; Teckman, J. H.; Marszal, E.; Osmulski, P. A. AFM Imaging Reveals Topographic Diversity of Wild Type and Z Variant Polymers of Human A1-Proteinase Inhibitor. *PLoS One* **2016**, *11* (3), e0151902. https://doi.org/10.1371/JOURNAL.PONE.0151902.

(72) Barth, A. Infrared Spectroscopy of Proteins. *Biochim Biophys Acta* **2007**, *1767* (9), 1073–1101. https://doi.org/10.1016/J.BBABIO.2007.06.004.

(73) Stathopulos, P. B.; Scholz, G. A.; Hwang, Y.-M.; Rumfeldt, J. A. O.; Lepock, J. R.; Meiering, E. M. Sonication of Proteins Causes Formation of Aggregates That Resemble Amyloid. *Protein Sci* **2004**, *13* (11), 3017–3027. https://doi.org/10.1110/PS.04831804.

(74) Kleckner, I. R.; Foster, M. P. An Introduction to NMR-Based Approaches for Measuring Protein Dynamics. *Biochim Biophys Acta* **2011**, *1814* (8), 942. https://doi.org/10.1016/J.BBAPAP.2010.10.012.

(75) Gordon S. Rule; Keven T. Hitchens. *Fundamentals of Protein NMR Spectroscopy*; Springer-Verlag, 2006; Vol. 5. https://doi.org/10.1007/1-4020-3500-4.

(76) Tzeng, S. R.; Pai, M. T.; Kalodimos, C. G. NMR Studies of Large Protein Systems. *Methods Mol Biol* **2012**, *831*, 133. https://doi.org/10.1007/978-1-61779-480-3_8.

(77) Waudby, C. A.; Ramos, A.; Cabrita, L. D.; Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Scientific Reports 2016 6:1* **2016**, *6* (1), 1–8. https://doi.org/10.1038/srep24826.

(78) Kleckner, I. R.; Foster, M. P. An Introduction to NMR-Based Approaches for Measuring Protein Dynamics. *Biochim Biophys Acta* **2011**, *1814* (8), 942. https://doi.org/10.1016/J.BBAPAP.2010.10.012.

(79) Binsch, G.; Gutowsky, H. S.; McCall, D. W.; Slichter, C. P.; Chem, J.; Holm, C. H. A Unified Theory of Exchange Effects on Nuclear Magnetic Resonance Line Shapes. *J Am Chem Soc* **1969**, *91* (6), 1304–1309. https://doi.org/10.1021/JA01034A007.

(80) McConnell, H. M. Reaction Rates by Nuclear Magnetic Resonance. *J Chem Phys* **1958**, *28* (3), 430–431. https://doi.org/10.1063/1.1744152.

(81) Kern, D.; Kern, G.; Scherer, G.; Fischer, G.; Drakenberg, T. Kinetic Analysis of Cyclophilin-Catalyzed Prolyl Cis/Trans Isomerization by Dynamic NMR Spectroscopy. *Biochemistry* **1995**, *34* (41), 13594–13602. https://doi.org/10.1021/BI00041A039.

(82) Damry, A. M. *From Protein Sequence to Motion to Function: Towards the Rational Design of Functional Protein Dynamics*; 2019.

(83) Sugase, K.; Dyson, H. J.; Wright, P. E. Mechanism of Coupled Folding and Binding of an Intrinsically Disordered Protein. *Nature 2007 447:7147* **2007**, *447* (7147), 1021–1025. https://doi.org/10.1038/nature05858.

(84) Bain, A. D. Chemical Exchange in NMR. *Prog Nucl Magn Reson Spectrosc* **2003**, *43* (3–4), 63–103. https://doi.org/10.1016/J.PNMRS.2003.08.001.

(85) Beach, H.; Cole, R.; Gill, M. L.; Loria, J. P. Conservation of Mus-Ms Enzyme Motions in the Apo- and Substrate-Mimicked State. *J Am Chem Soc* **2005**, *127* (25), 9167–9176. https://doi.org/10.1021/JA0514949.

(86) Freyer, M. W.; Lewis, E. A. Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. *Methods Cell Biol* **2008**, *84*, 79–113. https://doi.org/10.1016/S0091-679X(07)84004-0.

(87) Broom, H. R.; Vassall, K. A.; Rumfeldt, J. A. O.; Doyle, C. M.; Tong, M. S.; Bonner, J. M.; Meiering, E. M. Combined Isothermal Titration and Differential Scanning Calorimetry Define Three-State Thermodynamics of FALS-Associated Mutant Apo SOD1 Dimers and an Increased Population of Folded Monomer. *Biochemistry* **2016**, *55* (3), 519–533. https://doi.org/10.1021/ACS.BIOCHEM.5B01187/.

(88) Rosen, D. R.; Siddique, T.; Patterson, D.; Figlewicz, D. A.; Sapp, P.; Hentati, A.; Donaldson, D.; Goto, J.; O'Regan, J. P.; Deng, H. X.; Rahmani, Z.; Krizus, A.; McKenna-Yasek, D.; Cayabyab, A.; Gaston, S. M.; Berger, R.; Tanzi, R. E.; Halperin, J. J.; Herzfeldt, B.; Den Bergh, R. Van; Hung, W. Y.; Bird, T.; Deng, G.; Mulder, D. W.; Smyth, C.; Laing, N. G.; Soriano, E.; Pericak-Vance, M. A.; Haines, J.; Rouleau, G. A.; Gusella, J. S.; Horvitz, H. R.; Brown, R. H. Mutations in Cu/Zn Superoxide Dismutase Gene Are Associated with Familial Amyotrophic Lateral Sclerosis. *Nature* **1993**, *362* (6415), 59–62. https://doi.org/10.1038/362059A0.

(89) Strong, M. J.; Kesavapany, S.; Pant, H. C. The Pathobiology of Amyotrophic Lateral Sclerosis: A Proteinopathy? *J Neuropathol Exp Neurol* **2005**, *64* (8), 649–664. https://doi.org/10.1097/01.JNEN.0000173889.71434.EA.

(90) Robberecht, W.; Philips, T. The Changing Scene of Amyotrophic Lateral Sclerosis. *Nat Rev Neurosci* **2013**, *14* (4), 248–264. https://doi.org/10.1038/NRN3430.

(91) Okamoto, Y.; Shirakashi, Y.; Ihara, M.; Urushitani, M.; Oono, M.; Kawamoto, Y.; Yamashita, H.; Shimohama, S.; Kato, S.; Hirano, A.; Tomimoto, H.; Ito, H.; Takahashi, R. Colocalization of 14-3-3 Proteins with SOD1 in Lewy Body-Like Hyaline Inclusions in Familial Amyotrophic Lateral Sclerosis Cases and the Animal Model. *PLoS One* **2011**, *6* (5), e20427. https://doi.org/10.1371/JOURNAL.PONE.0020427.

(92) Kato, S.; Takikawa, M.; Nakashima, K.; Hirano, A.; Cleveland, D. W.; Kusaka, H.; Shibata, N.; Kato, M.; Nakano, I.; Ohama, E. New Consensus Research on Neuropathological Aspects of Familial Amyotrophic Lateral Sclerosis with Superoxide Dismutase 1 (SOD1) Gene Mutations: Inclusions Containing SOD1 in Neurons and Astrocytes. *Amyotroph Lateral Scler Other Motor Neuron Disord* **2000**, *1* (3), 163–184. https://doi.org/10.1080/14660820050515160.

(93) Matsumoto, S.; Kusaka, H.; Ito, H.; Shibata, N.; Asayama, T.; Imai, T. Sporadic Amyotrophic Lateral Sclerosis with Dementia and Cu/Zn Superoxide Dismutase-Positive Lewy Body-like Inclusions. *Clin Neuropathol* **1996**, *15* (1), 41–46.

(94) Vassall, K. A.; Stubbs, H. R.; Primmer, H. A.; Tong, M. S.; Sullivan, S. M.; Sobering, R.; Srinivasan, S.; Briere, L. A. K.; Dunn, S. D.; Colón, W.; Meiering, E. M. Decreased Stability and Increased Formation of Soluble Aggregates by Immature Superoxide Dismutase Do Not Account for Disease Severity in ALS. *Proc Natl Acad Sci U S A* **2011**, *108* (6), 2210–2215. https://doi.org/10.1073/PNAS.0913021108/-/DCSUPPLEMENTAL.

(95) Wang, Q.; Johnson, J. L.; Agar, N. Y. R.; Agar, J. N. Protein Aggregation and Protein Instability Govern Familial Amyotrophic Lateral Sclerosis Patient Survival. *PLoS Biol* **2008**, *6* (7), e170. https://doi.org/10.1371/JOURNAL.PBIO.0060170.

(96) Roberts, C. J. Non-Native Protein Aggregation Kinetics. *Biotechnol Bioeng* **2007**, *98* (5), 927–938. https://doi.org/10.1002/BIT.21627.

(97) Coelho, T.; Merlini, G.; Bulawa, C. E.; Fleming, J. A.; Judge, D. P.; Kelly, J. W.; Maurer, M. S.; Planté-Bordeneuve, V.; Labaudinière, R.; Mundayat, R.; Riley, S.; Lombardo, I.; Huertas, P. Mechanism of Action and Clinical Application of Tafamidis in Hereditary Transthyretin Amyloidosis. *Neurol Ther* **2016**, *5* (1), 1–25. https://doi.org/10.1007/s40120-016-0040-x.

(98) Furukawa, Y.; O'Halloran, T. V. Amyotrophic Lateral Sclerosis Mutations Have the Greatest Destabilizing Effect on the Apo- and Reduced Form of SOD1, Leading to Unfolding and Oxidative Aggregation. *J Biol Chem* **2005**, *280* (17), 17266–17274. https://doi.org/10.1074/JBC.M500482200.

(99) Chattopadhyay, M.; Durazo, A.; Se, H. S.; Strong, C. D.; Gralla, E. B.; Whitelegge, J. P.; Valentine, J. S. Initiation and Elongation in Fibrillation of ALS-Linked Superoxide Dismutase. *Proc Natl Acad Sci U S A* **2008**, *105* (48), 18663–18668. https://doi.org/10.1073/PNAS.0807058105.

(100) Lindberg, M. J.; Tibell, L.; Oliveberg, M. Common Denominator of Cu/Zn Superoxide Dismutase Mutants Associated with Amyotrophic Lateral Sclerosis: Decreased Stability of the Apo State. *Proceedings of the National Academy of Sciences* **2002**, *99* (26), 16607–16612. https://doi.org/10.1073/PNAS.262527099.

(101) Oztug Durer, Z. A.; Cohlberg, J. A.; Dinh, P.; Padua, S.; Ehrenclou, K.; Downes, S.; Tan, J. K.; Nakano, Y.; Bowman, C. J.; Hoskins, J. L.; Kwon, C.; Mason, A. Z.; Rodriguez, J. A.; Doucette, P. A.; Shaw, B. F.; Valentine, J. S. Loss of Metal Ions, Disulfide Reduction and Mutations Related to Familial ALS Promote Formation of Amyloid-Like Aggregates from Superoxide Dismutase. *PLoS One* **2009**, *4* (3), e5004. https://doi.org/10.1371/JOURNAL.PONE.0005004.

(102) Broom, H. R.; Primmer, H. A.; Rumfeldt, J. A. O.; Stathopulos, P. B.; Vassall, K. A.; Hwang, Y.-M.; Meiering, E. M. Folding and Aggregation of Cu, Zn-Superoxide Dismutase. *Amyotrophic Lateral Sclerosis* **2012**. https://doi.org/10.5772/31629.

(103) Broom, H. R.; Rumfeldt, J. A. O.; Meiering, E. M. Many Roads Lead to Rome? Multiple Modes of Cu,Zn Superoxide Dismutase Destabilization, Misfolding and Aggregation in Amyotrophic Lateral Sclerosis. *Essays Biochem* **2014**, *56* (1), 149–165. https://doi.org/10.1042/BSE0560149.

(104) Lang, L.; Kurnik, M.; Danielsson, J.; Oliveberg, M. Fibrillation Precursor of Superoxide Dismutase 1 Revealed by Gradual Tuning of the Protein-Folding Equilibrium. *Proc Natl Acad Sci U S A* **2012**, *109* (44), 17868–17873. https://doi.org/10.1073/PNAS.1201795109/-/DCSUPPLEMENTAL.

(105) Wright, G. S. A.; Antonyuk, S. V.; Hasnain, S. S. The Biophysics of Superoxide Dismutase-1 and Amyotrophic Lateral Sclerosis. *Q Rev Biophys* **2019**, *52*, e12. https://doi.org/10.1017/S003358351900012X.

(106) Sandelin, E.; Nordlund, A.; Andersen, P. M.; Marklund, S. S. L.; Oliveberg, M. Amyotrophic Lateral Sclerosis-Associated Copper/Zinc Superoxide Dismutase Mutations Preferentially Reduce the Repulsive Charge of the Proteins. *Journal of*

*Biological Chemistry* **2007**, *282* (29), 21230–21236. https://doi.org/10.1074/jbc.M700765200.

(107) Ayers, J. I.; Fromholt, S. E.; O'Neal, V. M.; Diamond, J. H.; Borchelt, D. R. Prion-like Propagation of Mutant SOD1 Misfolding and Motor Neuron Disease Spread along Neuroanatomical Pathways. *Acta Neuropathol* **2016**, *131* (1), 103–114. https://doi.org/10.1007/S00401-015-1514-0.

(108) Ayers, J. I.; Fromholt, S.; Koch, M.; DeBosier, A.; McMahon, B.; Xu, G.; Borchelt, D. R. Experimental Transmissibility of Mutant SOD1 Motor Neuron Disease. *Acta Neuropathol* **2014**, *128* (6), 791–803. https://doi.org/10.1007/S00401-014-1342-7.

(109) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys Rev* **2016**, *8* (4), 409–427. https://doi.org/10.1007/S12551-016-0218-6.

(110) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **1999**, *38* (50), 16424–16431. https://doi.org/10.1021/BI991765Q.

(111) Deng, H. X.; Hentati, A.; Tainer, J. A.; Iqbal, Z.; Cayabyab, A.; Hung, W. Y.; Getzoff, E. D.; Hu, P.; Herzfeldt, B.; Roos, R. P.; Warner, C.; Deng, G.; Soriano, E.; Smyth, C.; Parge, H. E.; Ahmed, A.; Roses, A. D.; Hallewell, R. A.; Pericak-Vance, M. A.; Siddique, T. Amyotrophic Lateral Sclerosis and Structural Defects in Cu,Zn Superoxide Dismutase. *Science* **1993**, *261* (5124), 1047–1051. https://doi.org/10.1126/SCIENCE.8351519.

(112) Lindberg, M. J.; Normark, J.; Holmgren, A.; Oliveberg, M. Folding of Human Superoxide Dismutase: Disulfide Reduction Prevents Dimerization and Produces Marginally Stable Monomers. *Proceedings of the National Academy of Sciences* **2004**, *101* (45), 15893–15898. https://doi.org/10.1073/PNAS.0403979101.

(113) Majumdar, R.; Manikwar, P.; Hickey, J. M.; Samra, H. S.; Sathish, H. A.; Bishop, S. M.; Middaugh, C. R.; Volkin, D. B.; Weis, D. D. Effects of Salts from the Hofmeister Series on the Conformational Stability, Aggregation Propensity, and Local Flexibility of an IgG1 Monoclonal Antibody. *Biochemistry* **2013**, *52* (19), 3376–3389. https://doi.org/10.1021/BI400232P.

(114) Cacace, M. G.; Landau, E. M.; Ramsden, J. J. The Hofmeister Series: Salt and Solvent Effects on Interfacial Phenomena. *Q Rev Biophys* **1997**, *30* (3), 241–277. https://doi.org/10.1017/S0033583597003363.

(115) Baldwin, R. L. How Hofmeister Ion Interactions Affect Protein Stability. *Biophys J* **1996**, *71* (4), 2056. https://doi.org/10.1016/S0006-3495(96)79404-3.

(116) Zhang, Y.; Cremer, P. S. Interactions between Macromolecules and Ions: The Hofmeister Series. *Curr Opin Chem Biol* **2006**, *10* (6), 658–663. https://doi.org/10.1016/J.CBPA.2006.09.020.

(117) Mouat, M. F.; Manchester, K. L. The Intracellular Ionic Strength of Red Cells and the Influence of Complex Formation. *undefined* **2007**, *8* (1), 58–60. https://doi.org/10.1007/BF02628107.

(118) Mouat, M. F.; Manchester, K. L. The Intracellular Ionic Strength of Red Cells and the Influence of Complex Formation. *undefined* **2007**, *8* (1), 58–60. https://doi.org/10.1007/BF02628107.

(119) Gaczynska, M.; Karpowicz, P.; Stuart, C. E.; Norton, M. G.; Teckman, J. H.; Marszal, E.; Osmulski, P. A. AFM Imaging Reveals Topographic Diversity of Wild Type and Z Variant Polymers of Human A1-Proteinase Inhibitor. *PLoS One* **2016**, *11* (3), e0151902. https://doi.org/10.1371/JOURNAL.PONE.0151902.

(120) Weissgerber, T. L.; Milic, N. M.; Winham, S. J.; Garovic, V. D. Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PLoS Biol* **2015**, *13* (4), e1002128. https://doi.org/10.1371/JOURNAL.PBIO.1002128.

(121) Willard, L.; Ranjan, A.; Zhang, H.; Monzavi, H.; Boyko, R. F.; Sykes, B. D.; Wishart, D. S. VADAR: A Web Server for Quantitative Evaluation of Protein Structure Quality. *Nucleic Acids Res* **2003**, *31* (13), 3316. https://doi.org/10.1093/NAR/GKG565.

(122) Waner, M. J.; Gilchrist, M.; Schindler, M.; Dantus, M. Imaging the Molecular Dimensions and Oligomerization of Proteins at Liquid/Solid Interfaces. *Journal of Physical Chemistry B* **1998**, *102* (9), 1649–1657. https://doi.org/10.1021/JP9732219.

(123) Vosough, F.; Barth, A. Characterization of Homogeneous and Heterogeneous Amyloid-B42 Oligomer Preparations with Biochemical Methods and Infrared Spectroscopy Reveals a Correlation between Infrared Spectrum and Oligomer Size. *ACS Chem Neurosci* **2021**, *12* (3), 473–488. https://doi.org/10.1021/ACSCHEMNEURO.0C00642.

(124) Ami, D.; Sciandrone, B.; Mereghetti, P.; Falvo, J.; Catelani, T.; Visentin, C.; Tortora, P.; Ventura, S.; Natalello, A.; Regonesi, M. E. Pathological ATX3 Expression Induces Cell Perturbations in E. Coli as Revealed by Biochemical and Biophysical

Investigations. *Int J Mol Sci* **2021**, *22* (2), 1–21. https://doi.org/10.3390/IJMS22020943.

(125) Natalello, A.; Frana, A. M.; Relini, A.; Apicella, A.; Invernizzi, G.; Casari, C.; Gliozzi, A.; Doglia, S. M.; Tortora, P.; Regonesi, M. E. A Major Role for Side-Chain Polyglutamine Hydrogen Bonding in Irreversible Ataxin-3 Aggregation. *PLoS One* **2011**, *6* (4), e18789. https://doi.org/10.1371/JOURNAL.PONE.0018789.

(126) Klementieva, O.; Willén, K.; Martinsson, I.; Israelsson, B.; Engdahl, A.; Cladera, J.; Uvdal, P.; Gouras, G. K. Pre-Plaque Conformational Changes in Alzheimer's Disease-Linked Aβ and APP. *Nature Communications 2017 8:1* **2017**, *8* (1), 1–9. https://doi.org/10.1038/ncomms14726.

(127) Barth, A. Infrared Spectroscopy of Proteins. *Biochim Biophys Acta Bioenerg* **2007**, *1767* (9), 1073–1101. https://doi.org/10.1016/j.bbabio.2007.06.004.

(128) Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclayes, S.; Dufrênes, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. Antiparallel β-Sheet: A Signature Structure of the Oligomeric Amyloid β-Peptide. *Biochemical Journal* **2009**, *421* (3), 415–423. https://doi.org/10.1042/BJ20090379.

(129) Goormaghtigh, E.; Cabiaux, V.; Ruysschaert, J. M. Determination of Soluble and Membrane Protein Structure by Fourier Transform Infrared Spectroscopy. *Subcell Biochem* **1994**, *23*, 405–450. https://doi.org/10.1007/978-1-4615-1863-1_10.

(130) Biancalana, M.; Koide, S. Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochim Biophys Acta* **2010**, *1804* (7), 1405–1412. https://doi.org/10.1016/J.BBAPAP.2010.04.001.

(131) Groenning, M.; Olsen, L.; van de Weert, M.; Flink, J. M.; Frokjaer, S.; Jørgensen, F. S. Study on the Binding of Thioflavin T to Beta-Sheet-Rich and Non-Beta-Sheet Cavities. *J Struct Biol* **2007**, *158* (3), 358–369. https://doi.org/10.1016/J.JSB.2006.12.010.

(132) Groenning, M.; Norrman, M.; Flink, J. M.; van de Weert, M.; Bukrinsky, J. T.; Schluckebier, G.; Frokjaer, S. Binding Mode of Thioflavin T in Insulin Amyloid Fibrils. *J Struct Biol* **2007**, *159* (3), 483–497. https://doi.org/10.1016/J.JSB.2007.06.004.

(133) Zou, Y.; Li, Y.; Hao, W.; Hu, X.; Ma, G. Parallel β-Sheet Fibril and Antiparallel β-Sheet Oligomer: New Insights into Amyloid Formation of Hen Egg White Lysozyme under Heat and Acidic Condition from FTIR Spectroscopy. *Journal of Physical Chemistry B* **2013**, *117* (15), 4003–4013. https://doi.org/10.1021/JP4003559/SUPPL_FILE/JP4003559_SI_001.PDF.

(134) Waeytens, J.; Van Hemelryck, V.; Deniset-Besseau, A.; Ruysschaert, J. M.; Dazzi, A.; Raussens, V. Characterization by Nano-Infrared Spectroscopy of Individual Aggregated Species of Amyloid Proteins. *Molecules 2020, Vol. 25, Page 2899* **2020**, *25* (12), 2899. https://doi.org/10.3390/MOLECULES25122899.

(135) Sarroukh, R.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. ATR-FTIR: A “Rejuvenated” Tool to Investigate Amyloid Proteins. *Biochim Biophys Acta* **2013**, *1828* (10), 2328–2338. https://doi.org/10.1016/J.BBAMEM.2013.04.012.

(136) Zandomeneghi, G.; Krebs, M. R. H.; McCammon, M. G.; Fändrich, M. FTIR Reveals Structural Differences between Native β-Sheet Proteins and Amyloid Fibrils. *Protein Science* **2009**, *13* (12), 3314–3321. https://doi.org/10.1110/ps.041024904.

(137) Broom, H. R.; Rumfeldt, J. A. O.; Vassall, K. A.; Meiering, E. M. Destabilization of the Dimer Interface Is a Common Consequence of Diverse ALS- Associated Mutations in Metal Free SOD1. **2015**, *24*, 2081–2089. https://doi.org/10.1002/pro.2803.

(138) Chiti, F.; Dobson, C. M. Amyloid Formation by Globular Proteins under Native Conditions. *Nature Chemical Biology 2009 5:1* **2008**, *5* (1), 15–22. https://doi.org/10.1038/nchembio.131.

(139) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid and Human Disease. *Annu Rev Biochem* **2006**, *75*, 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901.

(140) Qin, Z.; Hu, D.; Zhu, M.; Fink, A. L. Structural Characterization of the Partially Folded Intermediates of an Immunoglobulin Light Chain Leading to Amyloid Fibrillation and Amorphous Aggregation. *Biochemistry* **2007**, *46* (11), 3521–3531. https://doi.org/10.1021/bi061716v.

(141) Khurana, R.; Gillespie, J. R.; Talapatra, A.; Minert, L. J.; Ionescu-Zanetti, C.; Millett, I.; Fink, A. L. Partially Folded Intermediates as Critical Precursors of Light Chain Amyloid Fibrils and Amorphous Aggregates. *Biochemistry* **2001**, *40* (12), 3525–3535. https://doi.org/10.1021/bi001782b.

(142) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem* **2017**, *86*, 27–68. https://doi.org/10.1146/ANNUREV-BIOCHEM-061516-045115.

(143) Boatz, J. C.; Whitley, M. J.; Li, M.; Gronenborn, A. M.; Van Der Wel, P. C. A. Cataract-Associated P23T ΓD-Crystallin Retains a Native-like Fold in Amorphous-

Looking Aggregates Formed at Physiological PH. *Nat Commun* **2017**, *8*. https://doi.org/10.1038/NCOMMS15137.

(144) Broom, A.; Jacobi, Z.; Trainor, K.; Meiering, E. M. Computational Tools Help Improve Protein Stability but with a Solubility Tradeoff. *Journal of Biological Chemistry* **2017**, *292* (35), 14349–14361. https://doi.org/10.1074/JBC.M117.784165/ATTACHMENT/C2CF7D02-2765-4E0C-9105-B3558B09B8DE/MMC1.ZIP.

(145) Trainor, K.; Gingras, Z.; Shillingford, C.; Malakian, H.; Gosselin, M.; Lipovšek, D.; Meiering, E. M. Ensemble Modeling and Intracellular Aggregation of an Engineered Immunoglobulin-Like Domain. *J Mol Biol* **2016**, *428* (6), 1365–1374. https://doi.org/10.1016/J.JMB.2016.02.016.

(146) Trainor, K.; Broom, A.; Meiering, E. M. Exploring the Relationships between Protein Sequence, Structure and Solubility. *Curr Opin Struct Biol* **2017**, *42*, 136–146. https://doi.org/10.1016/J.SBI.2017.01.004.

(147) Kerman, A.; Liu, H. N.; Croul, S.; Bilbao, J.; Rogaeva, E.; Zinman, L.; Robertson, J.; Chakrabartty, A. Amyotrophic Lateral Sclerosis Is a Non-Amyloid Disease in Which Extensive Misfolding of SOD1 Is Unique to the Familial Form. *Acta Neuropathol* **2010**, *119* (3), 335–344. https://doi.org/10.1007/S00401-010-0646-5.

(148) Housmans, J. A. J.; Wu, G.; Schymkowitz, J.; Rousseau, F. A Guide to Studying Protein Aggregation. *FEBS Journal* **2021**, 1–30. https://doi.org/10.1111/febs.16312.

(149) Sengupta, I.; Udgaonkar, J. B. Structural Mechanisms of Oligomer and Amyloid Fibril Formation by the Prion Protein. *Chemical Communications* **2018**, *54* (49), 6230–6242. https://doi.org/10.1039/C8CC03053G.

(150) Jahn, T. R.; Parker, M. J.; Homans, S. W.; Radford, S. E. Amyloid Formation under Physiological Conditions Proceeds via a Native-like Folding Intermediate. *Nat Struct Mol Biol* **2006**, *13* (3), 195–201. https://doi.org/10.1038/nsmb1058.

(151) Andjus, P.; Stamenković, S.; Dučić, T. Synchrotron Radiation-Based FTIR Spectro-Microscopy of the Brainstem of the HSOD1 G93A Rat Model of Amyotrophic Lateral Sclerosis. *Eur Biophys J* **2019**, *48* (5), 475–484. https://doi.org/10.1007/S00249-019-01380-5.

(152) Krishnan, S.; Chi, E. Y.; Webb, J. N.; Chang, B. S.; Shan, D.; Goldenberg, M.; Manning, M. C.; Randolph, T. W.; Carpenter, J. F. Aggregation of Granulocyte Colony Stimulating Factor under Physiological Conditions: Characterization and Thermodynamic Inhibition. *Biochemistry* **2002**, *41* (20), 6422–6431. https://doi.org/10.1021/BI012006M.

(153) Fields, G. B.; Alonso, D. O. V.; Stigter, D.; Dill, K. A. Theory for the Aggregation of Proteins and Copolymers. *Journal of Physical Chemistry* **2002**, *96* (10), 3974–3981. https://doi.org/10.1021/J100189A013.

(154) Treuheit, M. J.; Kosky, A. A.; Brems, D. N. Inverse Relationship of Protein Concentration and Aggregation. *Pharmaceutical Research 2002 19:4* **2002**, *19* (4), 511–516. https://doi.org/10.1023/A:1015108115452.

(155) Shire, S. J.; Shahrokh, Z.; Liu, J. Challenges in the Development of High Protein Concentration Formulations. *J Pharm Sci* **2004**, *93* (6), 1390–1402. https://doi.org/10.1002/JPS.20079.

(156) Alford, J. R.; Kwok, S. C.; Roberts, J. N.; Wuttke, D. S.; Kendrick, B. S.; Carpenter, J. F.; Randolph, T. W. High Concentration Formulations of Recombinant Human Interleukin-1 Receptor Antagonist: I. Physical Characterization. *J Pharm Sci* **2008**, *97* (8), 3035–3050. https://doi.org/10.1002/JPS.21199.

(157) Minton, A. P. Influence of Macromolecular Crowding upon the Stability and State of Association of Proteins: Predictions and Observations. *J Pharm Sci* **2005**, *94* (8), 1668–1675. https://doi.org/10.1002/JPS.20417.

(158) Roberts, C. J. Protein Aggregation and Its Impact on Product Quality. *Curr Opin Biotechnol* **2014**, *30*, 211–217. https://doi.org/10.1016/J.COPBIO.2014.08.001.

(159) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. Protein Crowding Tunes Protein Stability. *J Am Chem Soc* **2011**, *133* (18), 7116–7120. https://doi.org/10.1021/ja200067p.

(160) Liu, J.; Andya, J. D.; Shire, S. J. A Critical Review of Analytical Ultracentrifugation and Field Flow Fractionation Methods for Measuring Protein Aggregation. *AAPS J* **2006**, *8* (3), E580. https://doi.org/10.1208/AAPSJ080367.

(161) Amin, S.; Barnett, G. V.; Pathak, J. A.; Roberts, C. J.; Sarangapani, P. S. Protein Aggregation, Particle Formation, Characterization & Rheology. *Curr Opin Colloid Interface Sci* **2014**, *19* (5), 438–449. https://doi.org/10.1016/J.COCIS.2014.10.002.

(162) Klement, K.; Wieligmann, K.; Meinhardt, J.; Hortschansky, P.; Richter, W.; Fändrich, M. Effect of Different Salt Ions on the Propensity of Aggregation and on

the Structure of Alzheimer's Abeta(1-40) Amyloid Fibrils. *J Mol Biol* **2007**, *373* (5), 1321–1333. https://doi.org/10.1016/J.JMB.2007.08.068.

(163) Jain, S.; Udgaonkar, J. B. Salt-Induced Modulation of the Pathway of Amyloid Fibril Formation by the Mouse Prion Protein. *Biochemistry* **2010**, *49* (35), 7615–7624. https://doi.org/10.1021/BI100745J.

(164) Yeh, V.; Broering, J. M.; Romanyuk, A.; Chen, B.; Chernoff, Y. O.; Bommarius, A. S. The Hofmeister Effect on Amyloid Formation Using Yeast Prion Protein. *Protein Sci* **2010**, *19* (1), 47. https://doi.org/10.1002/PRO.281.

(165) Munishkina, L. A.; Henriques, J.; Uversky, V. N.; Fink, A. L. Role of Protein-Water Interactions and Electrostatics in α-Synuclein Fibril Formation. *Biochemistry* **2004**, *43* (11), 3289–3300. https://doi.org/10.1021/bi034938r.

(166) Jahn, T. R.; Radford, S. E. Folding versus Aggregation: Polypeptide Conformations on Competing Pathways. *Arch Biochem Biophys* **2008**, *469* (1), 100–117. https://doi.org/10.1016/J.ABB.2007.05.015.

(167) Zhu, C.; Beck, M. V.; Griffith, J. D.; Deshmukh, M.; Dokholyan, N. V. Large SOD1 Aggregates, Unlike Trimeric SOD1, Do Not Impact Cell Viability in a Model of Amyotrophic Lateral Sclerosis. *Proc Natl Acad Sci U S A* **2018**, *115* (18), 4661–4665. https://doi.org/10.1073/pnas.1800187115.

(168) Eisele, Y. S.; Monteiro, C.; Fearns, C.; Encalada, S. E.; Wiseman, R. L.; Powers, E. T.; Kelly, J. W. Targeting Protein Aggregation for the Treatment of Degenerative Diseases. *Nat Rev Drug Discov* **2015**, *14* (11), 759–780. https://doi.org/10.1038/nrd4593.

(169) Finsy, R. Particle Sizing by Quasi-Elastic Light Scattering. *Adv Colloid Interface Sci* **1994**, *52* (9), 79–143. https://doi.org/10.1016/0001-8686(94)80041-3.

(170) Furukawa, Y.; Kaneko, K.; Yamanaka, K.; Nukina, N. Mutation-Dependent Polymorphism of Cu,Zn-Superoxide Dismutase Aggregates in the Familial Form of Amyotrophic Lateral Sclerosis. *Journal of Biological Chemistry* **2010**, *285* (29), 22221–22231. https://doi.org/10.1074/jbc.M110.113597.

(171) Ivanova, M. I.; Sievers, S. A.; Guenther, E. L.; Johnson, L. M.; Winkler, D. D.; Galaleldeen, A.; Sawaya, M. R.; Hart, P. J.; Eisenberg, D. S. Aggregation-Triggering Segments of SOD1 Fibril Formation Support a Common Pathway for Familial and Sporadic ALS. *Proc Natl Acad Sci U S A* **2014**, *111* (1), 197–201. https://doi.org/10.1073/PNAS.1320786110/SUPPL_FILE/SAPP.PDF.

(172) Pickles, S.; Semmler, S.; Broom, H. R.; Destroismaisons, L.; Legroux, L.; Arbour, N.; Meiering, E.; Cashman, N. R.; Vande Velde, C. ALS-Linked Misfolded SOD1 Species Have Divergent Impacts on Mitochondria. *Acta Neuropathol Commun* **2016**, *4* (1), 43. https://doi.org/10.1186/s40478-016-0313-8.

(173) Chan, P. K.; Chattopadhyay, M.; Sharma, S.; Souda, P.; Gralla, E. B.; Borchelt, D. R.; Whitelegge, J. P.; Valentine, J. S. Structural Similarity of Wild-Type and ALS-Mutant Superoxide Dismutase-1 Fibrils Using Limited Proteolysis and Atomic Force Microscopy. *Proc Natl Acad Sci U S A* **2013**, *110* (27), 10934–10939. https://doi.org/10.1073/pnas.1309613110.

(174) Cohen, N. R.; Zitzewitz, J. A.; Bilsel, O.; Robert Matthews, C. Nonnative Structure in a Peptide Model of the Unfolded State of Superoxide Dismutase 1 (SOD1):

Implications for ALS-Linked Aggregation. *Journal of Biological Chemistry* **2019**, *294* (37), 13708–13717. https://doi.org/10.1074/jbc.RA119.008765.

(175) Arosio, P.; Knowles, T. P. J.; Linse, S. On the Lag Phase in Amyloid Fibril Formation. *Physical Chemistry Chemical Physics* **2015**, *17* (12), 7606–7618. https://doi.org/10.1039/c4cp05563b.

(176) Karamanos, T. K.; Jackson, M. P.; Calabrese, A. N.; Goodchild, S. C.; Cawood, E. E.; Thompson, G. S.; Kalverda, A. P.; Hewitt, E. W.; Radford, S. E. Structural Mapping of Oligomeric Intermediates in an Amyloid Assembly Pathway. *Elife* **2019**, *8*, 1–32. https://doi.org/10.7554/eLife.46574.001.

(177) Elam, J. S.; Taylor, A. B.; Strange, R.; Antonyuk, S.; Doucette, P. A.; Rodriguez, J. A.; Hasnain, S. S.; Hayward, L. J.; Valentine, J. S.; Yeates, T. O.; Hart, P. J. Amyloid-like Filaments and Water-Filled Nanotubes Formed by SOD1 Mutant Proteins Linked to Familial ALS. *Nat Struct Biol* **2003**, *10* (6), 461–467. https://doi.org/10.1038/nsb935.

(178) Strange, R. W.; Antonyuk, S.; Hough, M. A.; Doucette, P. A.; Rodriguez, J. A.; Hart, P. J.; Hayward, L. J.; Valentine, J. S.; Hasnain, S. S. The Structure of Holo and Metal-Deficient Wild-Type Human Cu, Zn Superoxide Dismutase and Its Relevance to Familial Amyotrophic Lateral Sclerosis. *J Mol Biol* **2003**, *328* (4), 877–891. https://doi.org/10.1016/S0022-2836(03)00355-3.

(179) Nordlund, A.; Leinartaite, L.; Saraboji, K.; Aisenbrey, C.; Gröbner, G.; Zetterström, P.; Danielsson, J.; Logan, D. T.; Oliveberg, M. Functional Features Cause Misfolding of the ALS-Provoking Enzyme SOD1. *Proc Natl Acad Sci U S A* **2009**, *106* (24), 9667–9672. https://doi.org/10.1073/PNAS.0812046106.

(180) Sangwan, S.; Zhao, A.; Adams, K. L.; Jayson, C. K.; Sawaya, M. R.; Guenther, E. L.; Pan, A. C.; Ngo, J.; Moore, D. M.; Soriaga, A. B.; Do, T. D.; Goldschmidt, L.; Nelson, R.; Bowers, M. T.; Koehler, C. M.; Shaw, D. E.; Novitch, B. G.; Eisenberg, D. S. Atomic Structure of a Toxic, Oligomeric Segment of SOD1 Linked to Amyotrophic Lateral Sclerosis (ALS). *Proc Natl Acad Sci U S A* **2017**, *114* (33), 8770–8775. https://doi.org/10.1073/PNAS.1705091114.

(181) Sangwan, S.; Sawaya, M. R.; Murray, K. A.; Hughes, M. P.; Eisenberg, D. S. Atomic Structures of Corkscrew-Forming Segments of SOD1 Reveal Varied Oligomer Conformations. *Protein Sci* **2018**, *27* (7), 1231–1242. https://doi.org/10.1002/PRO.3391.

(182) Uversky, V. N. Mysterious Oligomerization of the Amyloidogenic Proteins. *FEBS J* **2010**, *277* (14), 2940. https://doi.org/10.1111/J.1742-4658.2010.07721.X.

(183) Kodali, R.; Wetzel, R. Polymorphism in the Intermediates and Products of Amyloid Assembly. *Curr Opin Struct Biol* **2007**, *17* (1), 48–57. https://doi.org/10.1016/J.SBI.2007.01.007.

(184) Rubin, J.; Khosravi, H.; Bruce, K. L.; Lydon, M. E.; Behrens, S. H.; Chernoff, Y. O.; Bommarius, A. S. Ion-Specific Effects on Prion Nucleation and Strain Formation. *J Biol Chem* **2013**, *288* (42), 30300. https://doi.org/10.1074/JBC.M113.467829.

(185) Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Bouvignies, G.; Meiering, E. M.; Kay, L. E. Thermal Fluctuations of Immature SOD1 Lead to Separate Folding and Misfolding Pathways. *Elife* **2015**, *4* (JUNE). https://doi.org/10.7554/ELIFE.07296.

(186) Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Sobering, R. E.; Meiering, E. M.; Kay, L. E. Probing the Free Energy Landscapes of ALS Disease Mutants of SOD1 by NMR Spectroscopy. *Proc Natl Acad Sci U S A* **2016**, *113* (45), E6939–E6945. https://doi.org/10.1073/PNAS.1611418113.

(187) Lomakin, A.; Benedek, G. B.; Teplow, D. B. Monitoring Protein Assembly Using Quasielastic Light Scattering Spectroscopy. *Methods Enzymol* **1999**, *309* (99), 429–459. https://doi.org/10.1016/S0076-6879(99)09029-1.

(188) Nobbmann, U.; Connah, M.; Fish, B.; Varley, P.; Gee, C.; Mulot, S.; Chen, J.; Zhou, L.; Lu, Y.; Sheng, F.; Yi, J.; Harding, S. E. Dynamic Light Scattering as a Relative Tool for Assessing the Molecular Integrity and Stability of Monoclonal Antibodies. *Biotechnol Genet Eng Rev* **2007**, *24* (1), 117–128. https://doi.org/10.1080/02648725.2007.10648095.

(189) Inagaki, S.; Ghirlando, R.; Grisshammer, R. Biophysical Characterization of Membrane Proteins in Nanodiscs. *Methods* **2013**, *59* (3), 287–300. https://doi.org/10.1016/j.ymeth.2012.11.006.

(190) *Formulation stability evaluation using light scattering techniques | Malvern Panalytical*. https://www.malvernpanalytical.com/en/learn/knowledge-center/application-notes/AN110627FormulationStabilityEvaluation (accessed 2022-02-28).

(191) Koppel, D. E.; Carlson, C.; Smilowitz, H. Analysis of Heterogeneous Fluorescence Photobleaching by Video Kinetics Imaging: The Method of Cumulants. *J Microsc* **1989**, *155* (2), 199–206. https://doi.org/10.1111/j.1365-2818.1989.tb02882.x.

(192) Pecora, R. Dynamic Light Scattering Measurement of Nanometer Particles in Liquids. *Journal of Nanoparticle Research 2000 2:2* **2000**, *2* (2), 123–131. https://doi.org/10.1023/A:1010067107182.

(193) *Dynamic Light Scattering (DLS)* | *Common Terms Defined* | *Malvern Panalytical*. https://www.malvernpanalytical.com/en/learn/knowledge-center/Whitepapers/WP111214DLSTermsDefined.html (accessed 2022-02-28).

(194) *Size, charge, and molecular weight of ferritin* | *Malvern Panalytical*. https://www.malvernpanalytical.com/en/learn/knowledge-center/application-notes/AN101104CharacterizationFerritin.html (accessed 2022-02-28).

(195) Oberg, K. A.; Fink, A. L. A New Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy Method for the Study of Proteins in Solution. *Anal Biochem* **1998**, *256* (1), 92–106. https://doi.org/10.1006/abio.1997.2486.

(196) Yang, H.; Yang, S.; Kong, J.; Dong, A.; Yu, S. Obtaining Information about Protein Secondary Structures in Aqueous Solution Using Fourier Transform IR Spectroscopy. *Nat Protoc* **2015**, *10* (3), 382–396. https://doi.org/10.1038/nprot.2015.024.

(197) Stathopulos, P. B.; Rumfeldt, J. A. O.; Scholz, G. A.; Irani, R. A.; Frey, H. E.; Hallewell, R. A.; Lepock, J. R.; Meiering, E. M. Cu/Zn Superoxide Dismutase Mutants Associated with Amyotrophic Lateral Sclerosis Show Enhanced Formation of Aggregates in Vitro. *Proc Natl Acad Sci U S A* **2003**, *100* (12), 7021–7026. https://doi.org/10.1073/pnas.1237797100.

(198) Baxter, N. J.; Williamson, M. P. Temperature Dependence of 1H Chemical Shifts in Proteins. *Journal of Biomolecular NMR 1997 9:4* **1997**, *9* (4), 359–369. https://doi.org/10.1023/A:1018334207887.

(199) Cierpicki, T.; Otlewski, J. Amide Proton Temperature Coefficients as Hydrogen Bond Indicators in Proteins. *J Biomol NMR* **2001**, *21* (3), 249–261. https://doi.org/10.1023/A:1012911329730.

(200) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley, 1986.

(201) Heikkinen, S.; Toikka, M. M.; Karhunen, P. T.; Kilpeläinen, I. A. Quantitative 2D HSQC (Q-HSQC) via Suppression of J-Dependence of Polarization Transfer in NMR Spectroscopy: Application to Wood Lignin. *J Am Chem Soc* **2003**, *125* (14), 4362–4367. https://doi.org/10.1021/JA029035K.

(202) Tomlinson, J. H.; Williamson, M. P. Amide Temperature Coefficients in the Protein G B1 Domain. *J Biomol NMR* **2012**, *52* (1), 57–64. https://doi.org/10.1007/S10858-011-9583-4.

(203) Trainor, K.; Palumbo, J. A.; MacKenzie, D. W. S.; Meiering, E. M. Temperature Dependence of NMR Chemical Shifts: Tracking and Statistical Analysis. *Protein Sci* **2020**, *29* (1), 306–314. https://doi.org/10.1002/PRO.3785.

(204) Baxter, N. J.; Hosszu, L. L. P.; Waltho, J. P.; Williamson, M. P. Characterisation of Low Free-Energy Excited States of Folded Proteins. *J Mol Biol* **1998**, *284* (5), 1625–1639. https://doi.org/10.1006/JMBI.1998.2265.

(205) Chandra, K.; Sharma, Y.; Chary, K. V. R. Characterization of Low-Energy Excited States in the Native State Ensemble of Non-Myristoylated and Myristoylated

Neuronal Calcium Sensor-1. *Biochim Biophys Acta* **2011**, *1814* (2), 334–344. https://doi.org/10.1016/J.BBAPAP.2010.10.007.

(206) Srivastava, A. K.; Chary, K. V. R. Conformational Heterogeneity and Dynamics in a Bγ-Crystallin from Hahella Chejuensis. *Biophys Chem* **2011**, *157* (1–3), 7–15. https://doi.org/10.1016/J.BPC.2011.04.001.

(207) Williamson, M. P. Many Residues in Cytochrome c Populate Alternative States under Equilibrium Conditions. *Proteins* **2003**, *53* (3), 731–739. https://doi.org/10.1002/PROT.10464.

(208) McCord, J. M.; Fridovich, I. Superoxide Dismutase: AN ENZYMIC FUNCTION FOR ERYTHROCUPREIN (HEMOCUPREIN). *Journal of Biological Chemistry* **1969**, *244* (22), 6049–6055. https://doi.org/10.1016/S0021-9258(18)63504-5.

(209) Crow, J. P.; Sampson, J. B.; Zhuang, Y.; Thompson, J. A.; Beckman, J. S. Decreased Zinc Affinity of Amyotrophic Lateral Sclerosis-Associated Superoxide Dismutase Mutants Leads to Enhanced Catalysis of Tyrosine Nitration by Peroxynitrite. *J Neurochem* **1997**, *69* (5), 1936–1944. https://doi.org/10.1046/J.1471-4159.1997.69051936.X.

(210) Doyle, C. M.; Naser, D.; Bauman, H. A.; Rumfeldt, J. A. O.; Meiering, E. M. Spectrophotometric Method for Simultaneous Measurement of Zinc and Copper in Metalloproteins Using 4-(2-Pyridylazo)Resorcinol. *Anal Biochem* **2019**, *579*, 44–56. https://doi.org/10.1016/J.AB.2019.03.007.

(211) MacKenzie, D. W. S.; Schaefer, A.; Steckner, J.; Leo, C. A.; Naser, D.; Artikis, E.; Broom, A.; Ko, T.; Shah, P.; Ney, M. Q.; Tran, E.; Smith, M. T. J.; Fuglestad, B.; Wand, A. J.; Brooks, C. L.; Meiering, E. M. A Fine Balance of Hydrophobic-

Electrostatic Communication Pathways in a PH-Switching Protein. *Proc Natl Acad Sci U S A* **2022**, *119* (26), e2119686119. https://doi.org/10.1073/PNAS.2119686119/.

(212) Wang, Q.; Johnson, J. L.; Agar, N. Y. R.; Agar, J. N. Protein Aggregation and Protein Instability Govern Familial Amyotrophic Lateral Sclerosis Patient Survival. *PLoS Biol* **2008**, *6* (7), e170. https://doi.org/10.1371/JOURNAL.PBIO.0060170.

(213) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog Nucl Magn Reson Spectrosc* **2013**, *73*, 1–16. https://doi.org/10.1016/J.PNMRS.2013.02.001.

(214) Deol, H. K. Characterizing the Dimer Interface and Thermal Unfolding of Reduced, Zinc Bound Superoxide Dismutase-1. **2018**.

(215) Strange, R. W.; Antonyuk, S.; Hough, M. A.; Doucette, P. A.; Rodriguez, J. A.; Hart, P. J.; Hayward, L. J.; Valentine, J. S.; Hasnain, S. S. The Structure of Holo and Metal-Deficient Wild-Type Human Cu, Zn Superoxide Dismutase and Its Relevance to Familial Amyotrophic Lateral Sclerosis. *J Mol Biol* **2003**, *328* (4), 877–891. https://doi.org/10.1016/S0022-2836(03)00355-3.

(216) Szpryngiel, S.; Oliveberg, M.; Mäler, L. Diffuse Binding of Zn2+ to the Denatured Ensemble of Cu/Zn Superoxide Dismutase 1. *FEBS Open Bio* **2015**, *5*, 56. https://doi.org/10.1016/J.FOB.2014.12.003.

(217) Lam, S. L.; Hsu, V. L. NMR Identification of Left-Handed Polyproline Type II Helices. *Biopolymers* **2003**, *69* (2), 270–281. https://doi.org/10.1002/BIP.10354.

(218) Museth, A. K.; Brorsson, A. C.; Lundqvist, M.; Tibell, L. A. E.; Jonsson, B. H. The ALS-Associated Mutation G93A in Human Copper-Zinc Superoxide Dismutase

Selectively Destabilizes the Remote Metal Binding Region. *Biochemistry* **2009**, *48* (37), 8817–8829. https://doi.org/10.1021/BI900703V.

(219) Cierpicki, T.; Zhukov, I.; Byrd, R. A.; Otlewski, J. Hydrogen Bonds in Human Ubiquitin Reflected in Temperature Coefficients of Amide Protons. *Journal of Magnetic Resonance* **2002**, *157* (2), 178–180. https://doi.org/10.1006/JMRE.2002.2597.

(220) Winkler, D. D.; Schuermann, J. P.; Cao, X.; Holloway, S. P.; Borchelt, D. R.; Carroll, M. C.; Proescher, J. B.; Culotta, V. C.; Hart, P. J. Structural and Biophysical Properties of the Pathogenic SOD1 Variant H46R/H48Q. *Biochemistry* **2009**, *48* (15), 3436–3447. https://doi.org/10.1021/BI8021735/.

(221) Zetterström, P.; Stewart, H. G.; Bergemalm, D.; Jonsson, P. A.; Graffmo, K. S.; Andersen, P. M.; Brännström, T.; Oliveberg, M.; Marklund, S. L. Soluble Misfolded Subfractions of Mutant Superoxide Dismutase-1s Are Enriched in Spinal Cords throughout Life in Murine ALS Models. *Proc Natl Acad Sci U S A* **2007**, *104* (35), 14157–14162. https://doi.org/10.1073/PNAS.0700477104/.

(222) Healy, E. F.; Cervantes, L. An in Silico Study of the Effect of SOD1 Electrostatic Loop Dynamics on Amyloid-like Filament Formation. *Eur Biophys J* **2016**, *45* (8), 853. https://doi.org/10.1007/S00249-016-1163-9.

(223) Elam, J. S.; Malek, K.; Rodriguez, J. A.; Doucette, P. A.; Taylor, A. B.; Hayward, L. J.; Cabelli, D. E.; Valentine, J. S.; Hart, P. J. An Alternative Mechanism of Bicarbonate-Mediated Peroxidation by Copper-Zinc Superoxide Dismutase. *Journal of Biological Chemistry* **2003**, *278* (23), 21032–21039. https://doi.org/10.1074/jbc.m300484200.

(224) Antonyuk, S.; Elam, J. S.; Hough, M. A.; Strange, R. W.; Doucette, P. A.; Rodriguez, J. A.; Hayward, L. J.; Valentine, J. S.; Hart, P. J.; Hasnain, S. S. Structural Consequences of the Familial Amyotrophic Lateral Sclerosis SOD1 Mutant His46Arg. *Protein Sci* **2005**, *14* (5), 1201. https://doi.org/10.1110/PS.041256705.

(225) Cao, X.; Antonyuk, S. v.; Seetharaman, S. v.; Whitson, L. J.; Taylor, A. B.; Holloway, S. P.; Strange, R. W.; Doucette, P. A.; Valentine, J. S.; Tiwari, A.; Hayward, L. J.; Padua, S.; Cohlberg, J. A.; Hasnain, S. S.; Hart, P. J. Structures of the G85R Variant of SOD1 in Familial Amyotrophic Lateral Sclerosis. *J Biol Chem* **2008**, *283* (23), 16169–16177. https://doi.org/10.1074/JBC.M801522200.

(226) Hayward, L. J.; Rodriguez, J. A.; Kim, J. W.; Tiwari, A.; Goto, J. J.; Cabelli, D. E.; Valentine, J. S.; Brown, R. H. Decreased Metallation and Activity in Subsets of Mutant Superoxide Dismutases Associated with Familial Amyotrophic Lateral Sclerosis. *J Biol Chem* **2002**, *277* (18), 15923–15931. https://doi.org/10.1074/JBC.M112087200.

(227) Jonsson, P. A.; Graffmo, K. S.; Andersen, P. M.; Brännström, T.; Lindberg, M.; Oliveberg, M.; Marklund, S. L. Disulphide-Reduced Superoxide Dismutase-1 in CNS of Transgenic Amyotrophic Lateral Sclerosis Models. *Brain* **2006**, *129* (Pt 2), 451–464. https://doi.org/10.1093/BRAIN/AWH704.

(228) Dellarole, M.; Caro, J. A.; Roche, J.; Fossat, M.; Barthe, P.; García-Moreno E., B.; Royer, C. A.; Roumestand, C. Evolutionarily Conserved Pattern of Interactions in a Protein Revealed by Local Thermal Expansion Properties. *J Am Chem Soc* **2015**, *137* (29), 9354–9362. https://doi.org/10.1021/JACS.5B04320.

(229) Mera-Adasme, R.; Suomivuori, C. M.; Fierro, A.; Pesonen, J.; Sundholm, D. The Role of Solvent Exclusion in the Interaction between D124 and the Metal Site in SOD1: Implications for ALS. *Journal of Biological Inorganic Chemistry* **2013**, *18* (8), 931–938. https://doi.org/10.1007/S00775-013-1039-8.

(230) Sengupta, I.; Udgaonkar, J. Monitoring Site-Specific Conformational Changes in Real-Time Reveals a Misfolding Mechanism of the Prion Protein. *Elife* **2019**, *8*. https://doi.org/10.7554/ELIFE.44698.

(231) Otting, G. Experimental NMR Techniques for Studies of Protein-Ligand Interactions. *Curr Opin Struct Biol* **1993**, *3* (5), 760–768. https://doi.org/10.1016/0959-440X(93)90061-O.

(232) Pellecchia, M.; Montgomery, D. L.; Stevens, S. Y.; vander Kooi, C. W.; Feng, H. P.; Gierasch, L. M.; Zuiderweg, E. R. P. Structural Insights into Substrate Binding by the Molecular Chaperone DnaK. *Nat Struct Biol* **2000**, *7* (4), 298–303. https://doi.org/10.1038/74062.

(233) Zuiderweg, E. R. P. Mapping Protein-Protein Interactions in Solution by NMR Spectroscopy. *Biochemistry* **2002**, *41* (1), 1–7. https://doi.org/10.1021/BI011870B.

(234) Arai, M.; Ferreon, J. C.; Wright, P. E. Quantitative Analysis of Multisite Protein-Ligand Interactions by NMR: Binding of Intrinsically Disordered P53 Transactivation Subdomains with the TAZ2 Domain of CBP. *J Am Chem Soc* **2012**, *134* (8), 3792–3803. https://doi.org/10.1021/JA209936U/.

(235) Kato, H.; van Ingen, H.; Zhou, B. R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y. Architecture of the High Mobility Group Nucleosomal Protein 2-Nucleosome

Complex as Revealed by Methyl-Based NMR. *Proc Natl Acad Sci U S A* **2011**, *108* (30), 12283–12288. https://doi.org/10.1073/PNAS.1105848108/.

(236) Wang, M.; Tang, Y.; Sato, S.; Vugmeyster, L.; McKnight, C. J.; Raleigh, D. P. Dynamic NMR Line-Shape Analysis Demonstrates That the Villin Headpiece Subdomain Folds on the Microsecond Time Scale. *J Am Chem Soc* **2003**, *125* (20), 6032–6033. https://doi.org/10.1021/JA028752B/.

(237) O'Connor, C.; Kovrigin, E. L. Characterization of the Second Ion-Binding Site in the G Domain of H-Ras. *Biochemistry* **2012**, *51* (48), 9638–9646. https://doi.org/10.1021/BI301304G.

(238) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The Ensemble Nature of Allostery. *Nature* **2014**, *508* (7496), 331–339. https://doi.org/10.1038/NATURE13001.

(239) Andersen, N. H.; Chen, C.; Marschner, T. M.; Krystek, S. R.; Bassolino, D. A. Conformational Isomerism of Endothelin in Acidic Aqueous Media: A Quantitative NOESY Analysis. *Biochemistry* **1992**, *31* (5), 1280–1295. https://doi.org/10.1021/BI00120A003.

(240) Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z.; Tong, H. Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts. 1. The Importance of Conformational Averaging. *Biochem. Biophys. Res. Commun* **1997**, *20* (1), 11.

(241) Bouvignies, G.; Vallurupalli, P.; Cordes, M. H. J.; Hansen, D. F.; Kay, L. E. Measuring 1HN Temperature Coefficients in Invisible Protein States by Relaxation Dispersion NMR Spectroscopy. *J Biomol NMR* **2011**, *50* (1), 13. https://doi.org/10.1007/S10858-011-9498-0.

(242) Bah, A.; Vernon, R. M.; Siddiqui, Z.; Krzeminski, M.; Muhandiram, R.; Zhao, C.; Sonenberg, N.; Kay, L. E.; Forman-Kay, J. D. Folding of an Intrinsically Disordered Protein by Phosphorylation as a Regulatory Switch. *Nature* **2015**, *519* (7541), 106–109. https://doi.org/10.1038/NATURE13999.

(243) Okazaki, H.; Matsuo, N.; Tenno, T.; Goda, N.; Shigemitsu, Y.; Ota, M.; Hiroaki, H. Using 1HN Amide Temperature Coefficients to Define Intrinsically Disordered Regions: An Alternative NMR Method. *Protein Sci* **2018**, *27* (10), 1821. https://doi.org/10.1002/PRO.3485.

(244) Suárez, I. P.; Gauto, D. F.; Hails, G.; Mascali, F. C.; Crespo, R.; Zhao, L.; Wang, J.; Rasia, R. M. Conformational Sampling of the Intrinsically Disordered DsRBD-1 Domain from Arabidopsis Thaliana DCL1. *Physical Chemistry Chemical Physics* **2018**, *20* (16), 11237–11246. https://doi.org/10.1039/C7CP07908G.

(245) de Dios, A. C.; Pearson, J. G.; Oldfield, E. Secondary and Tertiary Structural Effects on Protein NMR Chemical Shifts: An Ab Initio Approach. *Science (1979)* **1993**, *260* (5113), 1491–1496. https://doi.org/10.1126/SCIENCE.8502992.

(246) Wang, Y.; Jardetzky, O. Investigation of the Neighboring Residue Effects on Protein Chemical Shifts. *J Am Chem Soc* **2002**, *124* (47), 14075–14084. https://doi.org/10.1021/JA026811F/.

(247) Wang, Y.; Jardetzky, O. Predicting 15N Chemical Shifts in Proteins Using the Preceding Residue-Specific Individual Shielding Surfaces from φ, Ψi−1, and X1torsion Angles. *Journal of Biomolecular NMR 2004 28:4* **2004**, *28* (4), 327–340. https://doi.org/10.1023/B:JNMR.0000015397.82032.2A.

(248) Wang, L.; Markley, J. L. Empirical Correlation Between Protein Backbone 15N And 13C Secondary Chemical Shifts and Its Application to Nitrogen Chemical Shift Re-Referencing. *J Biomol NMR* **2009**, *44* (2), 95. https://doi.org/10.1007/S10858-009-9324-0.

(249) Miletti, T.; di Trani, J.; Levros, L. C.; Mittermaier, A. Conformational Plasticity Surrounding the Active Site of NADH Oxidase from Thermus Thermophilus. *Protein Sci* **2015**, *24* (7), 1114–1128. https://doi.org/10.1002/PRO.2693.

(250) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. 1H, 13C and 15N Random Coil NMR Chemical Shifts of the Common Amino Acids. I. Investigations of Nearest-Neighbor Effects. *J Biomol NMR* **1995**, *5* (1), 67–81. https://doi.org/10.1007/BF00227471.

(251) Glushka, J.; Lee, M.; Coffin, S.; Cowburn, D. 15N Chemical Shifts of Backbone Amides in Bovine Pancreatic Trypsin Inhibitor and Apamin. *J Am Chem Soc* **1989**, *111* (20), 7716–7722. https://doi.org/10.1021/JA00202A009/.

(252) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. Sequence-Dependent Correction of Random Coil NMR Chemical Shifts. *J Am Chem Soc* **2001**, *123* (13), 2970–2978. https://doi.org/10.1021/JA003760I.

(253) Kjaergaard, M.; Brander, S.; Poulsen, F. M. Random Coil Chemical Shift for Intrinsically Disordered Proteins: Effects of Temperature and PH. *J Biomol NMR* **2011**, *49* (2), 139–149. https://doi.org/10.1007/S10858-011-9472-X.

(254) Tamiola, K.; Acar, B.; Mulder, F. A. A. Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. *J Am Chem Soc* **2010**, *132* (51), 18000–18003. https://doi.org/10.1021/JA105656T.

(255) Nielsen, J. T.; Mulder, F. A. A. POTENCI: Prediction of Temperature, Neighbor and PH-Corrected Chemical Shifts for Intrinsically Disordered Proteins. *J Biomol NMR* **2018**, *70* (3), 141–165. https://doi.org/10.1007/S10858-018-0166-5.

(256) Braun, D.; Wider, G.; Wüthrich, K. Sequence-Corrected 15N "Random Coil" Chemical Shifts. *J Am Chem Soc* **1994**, *116* (19), 8466–8469. https://doi.org/10.1021/JA00098A005.

(257) Obara, M.; Kang, M. S.; Yamada, K. M. Site-Directed Mutagenesis of the Cell-Binding Domain of Human Fibronectin: Separable, Synergistic Sites Mediate Adhesive Function. *Cell* **1988**, *53* (4), 649–657. https://doi.org/10.1016/0092-8674(88)90580-6.

(258) Chou, P. Y.; Fasman, G. D. Prediction of Protein Conformation. *Biochemistry* **1974**, *13* (2), 222–245. https://doi.org/10.1021/BI00699A002.

(259) Trainor, K.; Doyle, C. M.; Metcalfe-Roach, A.; Steckner, J.; Lipovšek, D.; Malakian, H.; Langley, D.; Krystek, S. R.; Meiering, E. M. Design for Solubility May Reveal Induction of Amide Hydrogen/Deuterium Exchange by Protein Self-Association. *J Mol Biol* **2022**, *434* (2). https://doi.org/10.1016/j.jmb.2021.167398.

(260) Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M. Sensitivity Improvement in Proton-Detected Two-Dimensional Heteronuclear Correlation NMR Spectroscopy. *Journal of Magnetic Resonance (1969)* **1991**, *93* (1), 151–170. https://doi.org/10.1016/0022-2364(91)90036-S.

(261) Kay, L. E.; Keifer, P.; Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity.

*J Am Chem Soc* **1992**, *114* (26), 10663–10665. https://doi.org/10.1021/JA00052A088.

(262) Grzesiek, S.; Bax, A. The Importance of Not Saturating H2O in Protein NMR. Application to Sensitivity Enhancement and NOE Measurements. *J Am Chem Soc* **1993**, *115* (26), 12593–12594. https://doi.org/10.1021/JA00079A052.

(263) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S. J.; Sørensen, O. W.; Griesinger, C. A General Enhancement Scheme in Heteronuclear Multidimensional NMR Employing Pulsed Field Gradients. *J Biomol NMR* **1994**, *4* (2), 301–306. https://doi.org/10.1007/BF00175254.

(264) Wishart, D. S.; Sykes, B. D. The 13C Chemical-Shift Index: A Simple Method for the Identification of Protein Secondary Structure Using 13C Chemical-Shift Data. *J Biomol NMR* **1994**, *4* (2), 171–180. https://doi.org/10.1007/BF00175245.

(265) Suarez, I. P.; Burdisso, P.; Benoit, M. P. M. H.; Boisbouvier, J.; Rasia, R. M. Induced Folding in RNA Recognition by Arabidopsis Thaliana DCL1. *Nucleic Acids Res* **2015**, *43* (13), 6607. https://doi.org/10.1093/NAR/GKV627.

(266) Baldwin, R. L. How Hofmeister Ion Interactions Affect Protein Stability. *Biophys J* **1996**, *71* (4), 2056. https://doi.org/10.1016/S0006-3495(96)79404-3.

(267) Zhang, Y.; Cremer, P. S. Interactions between Macromolecules and Ions: The Hofmeister Series. *Curr Opin Chem Biol* **2006**, *10* (6), 658–663. https://doi.org/10.1016/J.CBPA.2006.09.020.

(268) Yeh, V.; Broering, J. M.; Romanyuk, A.; Chen, B.; Chernoff, Y. O.; Bommarius, A. S. The Hofmeister Effect on Amyloid Formation Using Yeast Prion Protein. *Protein Sci* **2010**, *19* (1), 47. https://doi.org/10.1002/PRO.281.

(269) Cacace, M. G.; Landau, E. M.; Ramsden, J. J. The Hofmeister Series: Salt and Solvent Effects on Interfacial Phenomena. *Q Rev Biophys* **1997**, *30* (3), 241–277. https://doi.org/10.1017/S0033583597003363.

(270) Majumdar, R.; Manikwar, P.; Hickey, J. M.; Samra, H. S.; Sathish, H. A.; Bishop, S. M.; Middaugh, C. R.; Volkin, D. B.; Weis, D. D. Effects of Salts from the Hofmeister Series on the Conformational Stability, Aggregation Propensity, and Local Flexibility of an IgG1 Monoclonal Antibody. *Biochemistry* **2013**, *52* (19), 3376–3389. https://doi.org/10.1021/BI400232P.

(271) Rakhit, R.; Chakrabartty, A. Structure, Folding, and Misfolding of Cu,Zn Superoxide Dismutase in Amyotrophic Lateral Sclerosis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2006**, *1762* (11–12), 1025–1037. https://doi.org/10.1016/J.BBADIS.2006.05.004.

(272) Cierpicki, T.; Otlewski, J. Amide Proton Temperature Coefficients as Hydrogen Bond Indicators in Proteins. *J Biomol NMR* **2001**, *21*, 249–261.

(273) Lee, S.; Choi, M. C.; al Adem, K.; Lukman, S.; Kim, T. Y. Aggregation and Cellular Toxicity of Pathogenic or Non-Pathogenic Proteins. *Scientific Reports 2020 10:1* **2020**, *10* (1), 1–14. https://doi.org/10.1038/s41598-020-62062-3.

(274) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid and Human Disease. *Annu Rev Biochem* **2006**, *75*, 333–366. https://doi.org/10.1146/annurev.biochem.75.101304.123901.

(275) Wang, W.; Nema, S.; Teagarden, D. Protein Aggregation--Pathways and Influencing Factors. *Int J Pharm* **2010**, *390* (2), 89–99. https://doi.org/10.1016/J.IJPHARM.2010.02.025.

(276) Hellstrand, E.; Boland, B.; Walsh, D. M.; Linse, S. Amyloid β-Protein Aggregation Produces Highly Kinetic Data and Occurs by a Two-Phase Process. *ACS Chem Neurosci* **2010**, *1* (1), 13. https://doi.org/10.1021/CN900015V.

(277) Treuheit, M. J.; Kosky, A. A.; Brems, D. N. Inverse Relationship of Protein Concentration and Aggregation. *Pharm Res* **2002**, *19* (4), 511–516. https://doi.org/10.1023/A:1015108115452.

(278) Laber, J. R.; Dear, B. J.; Martins, M. L.; Jackson, D. E.; Divenere, A.; Gollihar, J. D.; Ellington, A. D.; Truskett, T. M.; Johnston, K. P.; Maynard, J. A. Charge Shielding Prevents Aggregation of Supercharged GFP Variants at High Protein Concentration. *Mol Pharm* **2017**, *14* (10), 3269–3280. https://doi.org/10.1021/ACS.MOLPHARMACEUT.7B00322.

(279) Gupta, P.; Hall, C. K.; Voegler, A. C. Effect of Denaturant and Protein Concentrations upon Protein Refolding and Aggregation: A Simple Lattice Model. *Protein Sci* **1998**, *7* (12), 2642–2652. https://doi.org/10.1002/PRO.5560071218.

(280) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. A Robust and Cost-Effective Method for the Production of Val, Leu, Ile (Delta 1) Methyl-Protonated 15N-, 13C-, 2H-Labeled Proteins. *J Biomol NMR* **1999**, *13* (4), 369–374. https://doi.org/10.1023/A:1008393201236.

(281) Wang, W.; Roberts, C. J. Protein Aggregation – Mechanisms, Detection, and Control. *Int J Pharm* **2018**, *550* (1–2), 251–268. https://doi.org/10.1016/J.IJPHARM.2018.08.043.

(282) Klement, K.; Wieligmann, K.; Meinhardt, J.; Hortschansky, P.; Richter, W.; Fändrich, M. Effect of Different Salt Ions on the Propensity of Aggregation and on

the Structure of Alzheimer's Abeta(1-40) Amyloid Fibrils. *J Mol Biol* **2007**, *373* (5), 1321–1333. https://doi.org/10.1016/J.JMB.2007.08.068.

(283) Jain, S.; Udgaonkar, J. B. Salt-Induced Modulation of the Pathway of Amyloid Fibril Formation by the Mouse Prion Protein. *Biochemistry* **2010**, *49* (35), 7615–7624. https://doi.org/10.1021/BI100745J.

(284) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. Protein Crowding Tunes Protein Stability. *J Am Chem Soc* **2011**, *133* (18), 7116–7120. https://doi.org/10.1021/JA200067P.

(285) Minton, A. P. Influence of Macromolecular Crowding upon the Stability and State of Association of Proteins: Predictions and Observations. *J Pharm Sci* **2005**, *94* (8), 1668–1675. https://doi.org/10.1002/JPS.20417.

(286) Ściskalska, M.; Ołdakowska, M.; Marek, G.; Milnerowicz, H. Changes in the Activity and Concentration of Superoxide Dismutase Isoenzymes (Cu/Zn SOD, MnSOD) in the Blood of Healthy Subjects and Patients with Acute Pancreatitis. *Antioxidants* **2020**, *9* (10), 1–16. https://doi.org/10.3390/ANTIOX9100948.

(287) Gertsman, I.; Wuu, J.; McAlonis-Downes, M.; Ghassemian, M.; Ling, K.; Rigo, F.; Bennett, F.; Benatar, M.; Miller, T. M.; da Cruz, S. An Endogenous Peptide Marker Differentiates SOD1 Stability and Facilitates Pharmacodynamic Monitoring in SOD1 Amyotrophic Lateral Sclerosis. *JCI Insight* **2019**, *4* (10). https://doi.org/10.1172/JCI.INSIGHT.122768.

(288) Wang, Y.; Branicky, R.; Noë, A.; Hekimi, S. Superoxide Dismutases: Dual Roles in Controlling ROS Damage and Regulating ROS Signaling. *J Cell Biol* **2018**, *217* (6), 1915. https://doi.org/10.1083/JCB.201708007.

(289) Andjus, P.; Stamenković, S.; Dučić, T. Synchrotron Radiation-Based FTIR Spectro-Microscopy of the Brainstem of the HSOD1 G93A Rat Model of Amyotrophic Lateral Sclerosis. *Eur Biophys J* **2019**, *48* (5), 475–484. https://doi.org/10.1007/S00249-019-01380-5.

(290) Sugaya, K.; Nakano, I. Prognostic Role of "Prion-like Propagation" in SOD1-Linked Familial ALS: An Alternative View. *Front Cell Neurosci* **2014**, *8* (October), 1–10. https://doi.org/10.3389/FNCEL.2014.00359.

(291) Pokrishevsky, E.; Hong, R. H.; Mackenzie, I. R.; Cashman, N. R. Spinal Cord Homogenates from SOD1 Familial Amyotrophic Lateral Sclerosis Induce SOD1 Aggregation in Living Cells. *PLoS One* **2017**, *12* (9), e0184384. https://doi.org/10.1371/JOURNAL.PONE.0184384.

(292) Polymenidou, M.; Cleveland, D. W. The Seeds of Neurodegeneration: Prion-like Spreading in ALS. *Cell* **2011**, *147* (3), 498–508. https://doi.org/10.1016/J.CELL.2011.10.011.

(293) Wright, G. S. A.; Antonyuk, S. v.; Hasnain, S. S. A Faulty Interaction between SOD1 and HCCS in Neurodegenerative Disease. *Sci Rep* **2016**, *6*. https://doi.org/10.1038/SREP27691.

(294) Karch, C. M.; Borchelt, D. R. A Limited Role for Disulfide Cross-Linking in the Aggregation of Mutant SOD1 Linked to Familial Amyotrophic Lateral Sclerosis. *J Biol Chem* **2008**, *283* (20), 13528–13537. https://doi.org/10.1074/JBC.M800564200.

(295) Leal, S. S.; Cristóvão, J. S.; Biesemeier, A.; Cardoso, I.; Gomes, C. M. Aberrant Zinc Binding to Immature Conformers of Metal-Free Copper-Zinc Superoxide

Dismutase Triggers Amorphous Aggregation. *Metallomics* **2015**, *7* (2), 333–346. https://doi.org/10.1039/C4MT00278D.

(296) Lepock, J. R.; Frey, H. E.; Hallewell, R. A. Contribution of Conformational Stability and Reversibility of Unfolding to the Increased Thermostability of Human and Bovine Superoxide Dismutase Mutated at Free Cysteines. *Journal of Biological Chemistry* **1990**, *265* (35), 21612–21618. https://doi.org/10.1016/s0021-9258(18)45784-5.

(297) Rumfeldt, J. A. O.; Stathopulos, P. B.; Chakrabarrty, A.; Lepock, J. R.; Meiering, E. M. Mechanism and Thermodynamics of Guanidinium Chloride-Induced Denaturation of ALS-Associated Mutant Cu,Zn Superoxide Dismutases. *J Mol Biol* **2006**, *355* (1), 106–123. https://doi.org/10.1016/J.JMB.2005.10.042.

(298) Lyons, T. J.; Liu, H.; Goto, J. J.; Nersissian, A.; Roe, J. A.; Graden, J. A.; Café, C.; Ellerby, L. M.; Bredesen, D. E.; Gralla, E. B.; Valentine, J. S. Mutations in Copper-Zinc Superoxide Dismutase That Cause Amyotrophic Lateral Sclerosis Alter the Zinc Binding Site and the Redox Behavior of the Protein. *Proc Natl Acad Sci U S A* **1996**, *93* (22), 12240. https://doi.org/10.1073/PNAS.93.22.12240.

(299) Furukawa, Y.; O'Halloran, T. v. Amyotrophic Lateral Sclerosis Mutations Have the Greatest Destabilizing Effect on the Apo- and Reduced Form of SOD1, Leading to Unfolding and Oxidative Aggregation. *J Biol Chem* **2005**, *280* (17), 17266–17274. https://doi.org/10.1074/JBC.M500482200.

(300) Mulligan, V. K.; Kerman, A.; Ho, S.; Chakrabartty, A. Denaturational Stress Induces Formation of Zinc-Deficient Monomers of Cu,Zn Superoxide Dismutase:

Implications for Pathogenesis in Amyotrophic Lateral Sclerosis. *J Mol Biol* **2008**, *383* (2), 424–436. https://doi.org/10.1016/J.JMB.2008.08.024.

(301) Velázquez-Campoy, A.; Ohtaka, H.; Nezami, A.; Muzammil, S.; Freire, E. Isothermal Titration Calorimetry. *Curr Protoc Cell Biol* **2004**, *23* (1), 17.8.1-17.8.24. https://doi.org/10.1002/0471143030.CB1708S23.

(302) Burrows, S. D.; Doyle, M. L.; Murphy, K. P.; Franklin, S. G.; White, J. R.; Brooks, I.; McNulty, D. E.; Scott, M. O.; Knutson, J. R.; Porter, D.; Young, P. R.; Hensley, P. Determination of the Monomer-Dimer Equilibrium of Interleukin-8 Reveals It Is a Monomer at Physiological Concentrations. *Biochemistry* **1994**, *33* (43), 12741–12745. https://doi.org/10.1021/BI00209A002.

(303) Marky, L. A.; Breslauer, K. J. Calculating Thermodynamic Data for Transitions of Any Molecularity from Equilibrium Melting Curves. *Biopolymers* **1987**, *26* (9), 1601–1620. https://doi.org/10.1002/BIP.360260911.

(304) Kayatekin, C.; Zitzewitz, J. A.; Matthews, C. R. Zinc Binding Modulates the Entire Folding Free Energy Surface of Human Cu,Zn Superoxide Dismutase. *J Mol Biol* **2008**, *384* (2), 540–555. https://doi.org/10.1016/J.JMB.2008.09.045.

(305) Hough, M. A.; Grossmann, J. G.; Antonyuk, S. v.; Strange, R. W.; Doucette, P. A.; Rodriguez, J. A.; Whitson, L. J.; Hart, P. J.; Hayward, L. J.; Valentine, J. S.; Hasnain, S. S. Dimer Destabilization in Superoxide Dismutase May Result in Disease-Causing Properties: Structures of Motor Neuron Disease Mutants. *Proc Natl Acad Sci U S A* **2004**, *101* (16), 5976–5981. https://doi.org/10.1073/PNAS.0305143101.

(306) Wang, J.; Caruano-Yzermans, A.; Rodriguez, A.; Scheurmann, J. P.; Slunt, H. H.; Cao, X.; Gitlin, J.; Hart, P. J.; Borchelt, D. R. Disease-Associated Mutations at Copper Ligand Histidine Residues of Superoxide Dismutase 1 Diminish the Binding of Copper and Compromise Dimer Stability. *J Biol Chem* **2007**, *282* (1), 345–352. https://doi.org/10.1074/JBC.M604503200.

(307) Brown, N. M.; Torres, A. S.; Doan, P. E.; O'Halloran, T. v. Oxygen and the Copper Chaperone CCS Regulate Posttranslational Activation of Cu,Zn Superoxide Dismutase. *Proceedings of the National Academy of Sciences* **2004**, *101* (15), 5518–5523. https://doi.org/10.1073/PNAS.0401175101.

(308) Furukawa, Y.; Torres, A. S.; O'Halloran, T. v. Oxygen-Induced Maturation of SOD1: A Key Role for Disulfide Formation by the Copper Chaperone CCS. *EMBO J* **2004**, *23* (14), 2872–2881. https://doi.org/10.1038/SJ.EMBOJ.7600276.

(309) Seetharaman, S. v.; Winkler, D. D.; Taylor, A. B.; Cao, X.; Whitson, L. J.; Doucette, P. A.; Valentine, J. S.; Schirf, V.; Demeler, B.; Carroll, M. C.; Culotta, V. C.; Hart, P. J. Disrupted Zinc-Binding Sites in Structures of Pathogenic SOD1 Variants D124V and H80R. *Biochemistry* **2010**, *49* (27), 5714. https://doi.org/10.1021/BI100314N.

(310) Paramita, V. D.; Panyoyai, N.; Kasapis, S. Molecular Functionality of Plant Proteins from Low- to High-Solid Systems with Ligand and Co-Solute. *Int J Mol Sci* **2020**, *21* (7). https://doi.org/10.3390/IJMS21072550.

(311) Prabhu, N. v.; Sharp, K. A. Heat Capacity in Proteins. *Annu Rev Phys Chem* **2005**, *56*, 521–548. https://doi.org/10.1146/ANNUREV.PHYSCHEM.56.092503.141202.

(312) DiDonato, M.; Craig, L.; Huff, M. E.; Thayer, M. M.; Cardoso, R. M. F.; Kassmann, C. J.; Lo, T. P.; Bruns, C. K.; Powers, E. T.; Kelly, J. W.; Getzoff, E. D.; Tainer, J. A. ALS Mutants of Human Superoxide Dismutase Form Fibrous Aggregates via Framework Destabilization. *J Mol Biol* **2003**, *332* (3), 601–615. https://doi.org/10.1016/S0022-2836(03)00889-1.

(313) Davis, C. M.; Gruebele, M. Non-Steric Interactions Predict the Trend and Steric Interactions the Offset of Protein Stability in Cells. *Chemphyschem* **2018**, *19* (18), 2290–2294. https://doi.org/10.1002/CPHC.201800534.

# APPENDIX

APPENDIX A: Temperature coefficients for apoSH pWT and mutants for the folded temperature range of 9–18 °C.

| Residue | pWT | A4V | E100G | G85R | V148G | V148I |
|---|---|---|---|---|---|---|
| **1** | - | - | - | - | - | - |
| **2** | -0.93 | - | - | -1.50 | - | -1.48 |
| **3** | -7.12 | -7.59 | -8.33 | -8.05 | -9.14 | -8.71 |
| **4** | -0.74 | -6.20 | -0.23 | -0.90 | -0.44 | -0.38 |
| **5** | -3.66 | -5.53 | -3.48 | -4.30 | -3.82 | -3.53 |
| **6** | -3.27 | - | -2.81 | -3.40 | -3.28 | -3.09 |
| **7** | -6.62 | -6.95 | -6.22 | -6.70 | -9.01 | -6.90 |
| **8** | -4.18 | -4.06 | -4.35 | -4.14 | -5.43 | -3.86 |
| **9** | -5.15 | -6.38 | - | -6.76 | - | - |
| **10** | -9.53 | -10.02 | -9.48 | -8.49 | -10.01 | -9.14 |
| **11** | -7.31 | -1.19 | -6.68 | -6.74 | -6.57 | -6.60 |
| **12** | 0.91 | 0.76 | 1.34 | -0.21 | 1.06 | 1.04 |
| **13** | - | - | - | - | - | - |
| **14** | -5.48 | -5.11 | -4.87 | -5.29 | -5.05 | -5.12 |
| **15** | 0.52 | 0.25 | 0.35 | 0.33 | 0.97 | 0.46 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **16** | 1.07 | 2.47 | 1.31 | -0.59 | 1.81 | 1.34 |
| **17** | -2.88 | -3.48 | -2.18 | -2.38 | -2.86 | -2.88 |
| **18** | -2.77 | -2.75 | -3.04 | -5.58 | -3.81 | -3.13 |
| **19** | -1.91 | -1.93 | -1.31 | -2.69 | -2.50 | -2.44 |
| **20** | -1.21 | -1.45 | -1.22 | -1.39 | -1.87 | -1.42 |
| **21** | -2.53 | -3.12 | -3.15 | -3.33 | -2.73 | -2.76 |
| **22** | -4.96 | -6.02 | -4.35 | -4.46 | -5.11 | -4.59 |
| **23** | - | -3.52 | - | - | - | - |
| **24** | -7.39 | -6.60 | -6.82 | -7.32 | - | -7.18 |
| **25** | - | - | - | - | - | - |
| **26** | -4.69 | -5.47 | -5.38 | -7.68 | -5.38 | -4.81 |
| **27** | -3.81 | -4.40 | -4.05 | -3.96 | -4.03 | -3.84 |
| **28** | - | - | - | - | - | - |
| **29** | -3.99 | -5.81 | -4.50 | -4.38 | -4.40 | -4.42 |
| **30** | -3.04 | -3.19 | -3.22 | -3.75 | -3.01 | -2.70 |
| **31** | -1.16 | -2.05 | - | -3.51 | -1.17 | -1.45 |
| **32** | -2.23 | -1.73 | -2.28 | -2.91 | -2.48 | -2.11 |
| **33** | -3.74 | -5.66 | -5.58 | -5.66 | -5.80 | -2.98 |
| **34** | -0.25 | -0.71 | 1.05 | -0.56 | 0.08 | -0.22 |
| **35** | -1.62 | -1.00 | -0.83 | -0.49 | -1.14 | -1.35 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **36** | -2.54 | -3.32 | -2.52 | -2.64 | -3.06 | -2.82 |
| **37** | -3.26 | -3.58 | -2.09 | -4.04 | -1.92 | -2.78 |
| **38** | -2.92 | - | -5.98 | - | -8.85 | -6.07 |
| **39** | -16.94 | -4.85 | - | - | - | - |
| **40** | - | - | - | -6.02 | -9.96 | -9.88 |
| **41** | -6.91 | -6.90 | -6.81 | -7.34 | -7.16 | -6.86 |
| **42** | - | -7.81 | - | -7.80 | -8.91 | - |
| **43** | - | -2.96 | - | - | 0.86 | - |
| **44** | -2.82 | -2.43 | -2.93 | -8.04 | -2.04 | -1.50 |
| **45** | -11.69 | -7.74 | -13.13 | -8.55 | -15.82 | -21.90 |
| **46** | -3.61 | - | -1.73 | - | -2.99 | -3.29 |
| **47** | -3.89 | -5.46 | -4.35 | - | -4.39 | -4.05 |
| **48** | -4.22 | -4.88 | -3.43 | -4.25 | -4.99 | -4.47 |
| **49** | - | - | - | - | - | - |
| **50** | - | - | - | - | - | - |
| **51** | -7.22 | -6.86 | -5.73 | -6.07 | -5.87 | -5.89 |
| **52** | -6.13 | -5.97 | -5.19 | -5.19 | -6.27 | -6.14 |
| **53** | -6.00 | -7.31 | -5.62 | -7.47 | -5.62 | -6.03 |
| **54** | -4.09 | -3.99 | -3.60 | -4.44 | -3.86 | -4.10 |
| **55** | -6.89 | -7.62 | -6.63 | -7.11 | -6.85 | -6.93 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **56** | -6.20 | -6.09 | -5.84 | -6.44 | -6.38 | -6.02 |
| **57** | -4.02 | -3.94 | -4.06 | -3.46 | -4.23 | -3.97 |
| **58** | -6.32 | -6.26 | -5.33 | -3.45 | -6.17 | -6.25 |
| **59** | - | -8.89 | -8.25 | -6.18 | - | -4.53 |
| **60** | -5.78 | -5.30 | -7.38 | -6.74 | -7.70 | -5.64 |
| **61** | -6.03 | -5.94 | -5.59 | -5.64 | -6.14 | -5.82 |
| **62** | - | - | - | - | - | - |
| **63** | - | - | - | - | - | - |
| **64** | - | - | - | - | - | - |
| **65** | - | - | - | - | - | - |
| **66** | - | - | - | - | - | - |
| **67** | -4.07 | -4.56 | -3.76 | -3.83 | -5.49 | -3.89 |
| **68** | -3.09 | -3.38 | -2.84 | -2.60 | -3.50 | -2.96 |
| **69** | -3.74 | -5.89 | - | -8.01 | -5.66 | - |
| **70** | -8.51 | -7.60 | - | -7.66 | -8.08 | -8.23 |
| **71** | - | - | - | - | - | - |
| **72** | - | - | - | - | - | - |
| **73** | -5.66 | -5.43 | -5.54 | -5.72 | -5.87 | -5.74 |
| **74** | - | - | - | - | - | - |
| **75** | -7.42 | -7.75 | -7.30 | -5.95 | -7.54 | -6.64 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **76** | -5.21 | -6.11 | - | -8.53 | - | -5.68 |
| **77** | -6.22 | -6.82 | -5.26 | -8.03 | -6.52 | -5.90 |
| **78** | -6.76 | -5.61 | -7.51 | -7.83 | -8.12 | -6.77 |
| **79** | - | - | - | - | - | - |
| **80** | - | - | - | - | - | - |
| **81** | - | - | - | - | - | - |
| **82** | - | - | - | - | - | - |
| **83** | - | - | - | - | - | - |
| **84** | - | - | - | - | - | - |
| **85** | - | -6.04 | - | - | -5.73 | - |
| **86** | -5.79 | -6.28 | -5.73 | - | -6.12 | -5.92 |
| **87** | -1.73 | -2.22 | -2.77 | -4.59 | -1.63 | -1.79 |
| **88** | -5.77 | -5.55 | -5.32 | -5.37 | -5.77 | -5.40 |
| **89** | -5.02 | -5.07 | -5.53 | -4.40 | -5.12 | -5.08 |
| **90** | -4.44 | -4.52 | -3.77 | -5.12 | -4.98 | -4.93 |
| **91** | -6.47 | -6.66 | -6.16 | -6.83 | -6.43 | -6.65 |
| **92** | -2.40 | -1.28 | -1.08 | -2.69 | -1.66 | -1.39 |
| **93** | -1.89 | -2.01 | -1.01 | -1.39 | -0.76 | -1.37 |
| **94** | -1.15 | -0.82 | -0.91 | -0.92 | -0.82 | -0.95 |
| **95** | -6.91 | -7.89 | -7.09 | -6.82 | -8.08 | -7.61 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **96** | -4.03 | -4.15 | -4.41 | -4.77 | -4.51 | -4.17 |
| **97** | -3.63 | -3.88 | -2.18 | -3.75 | -3.35 | -3.85 |
| **98** | -8.86 | -8.45 | -7.80 | -7.14 | -8.52 | -8.39 |
| **99** | -1.90 | -3.14 | -2.83 | -2.72 | -2.05 | -1.89 |
| **100** | -7.44 | -8.21 | - | -9.48 | -7.32 | -7.89 |
| **101** | -5.25 | -5.84 | -5.92 | -7.35 | -5.28 | -5.42 |
| **102** | -7.89 | -7.56 | - | -7.13 | -7.72 | -7.53 |
| **103** | -7.06 | -8.06 | - | - | -7.12 | -7.01 |
| **104** | -5.50 | -4.36 | - | - | -5.28 | -5.99 |
| **105** | -1.06 | -4.28 | -1.00 | -8.85 | -0.37 | -0.62 |
| **106** | -4.01 | -7.43 | - | -6.67 | -5.46 | -3.11 |
| **107** | -1.49 | -4.15 | -1.32 | -1.59 | -1.21 | -1.40 |
| **108** | -6.13 | -6.06 | -5.51 | -5.55 | -6.10 | -5.86 |
| **109** | -8.12 | -7.79 | -7.93 | -8.39 | -8.18 | -7.82 |
| **110** | -6.44 | -7.46 | -6.62 | -7.09 | -6.08 | -6.63 |
| **111** | -4.41 | -0.81 | - | - | - | - |
| **112** | -4.64 | -5.35 | -4.17 | -4.33 | -4.49 | -4.29 |
| **113** | -3.09 | -3.54 | -2.84 | -3.08 | -3.00 | -2.88 |
| **114** | -2.59 | -2.11 | -2.09 | -2.66 | -6.51 | -2.90 |
| **115** | -3.14 | -4.98 | -9.04 | -4.69 | -3.53 | -5.84 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **116** | -2.32 | -0.63 | - | -6.28 | -7.10 | -5.21 |
| **117** | -3.15 | -2.45 | -2.65 | -3.84 | -3.47 | -3.30 |
| **118** | -1.31 | -1.46 | -0.46 | - | -0.46 | -1.37 |
| **119** | -6.48 | -6.50 | - | - | - | - |
| **120** | - | -4.65 | - | - | -7.86 | - |
| **121** | -8.82 | -9.72 | -3.52 | -7.06 | -8.89 | -7.59 |
| **122** | -7.39 | - | - | - | - | -6.62 |
| **123** | -4.44 | -9.31 | -7.57 | -6.82 | -10.09 | -2.99 |
| **124** | - | - | - | - | - | - |
| **125** | -8.42 | -7.99 | -7.53 | - | - | -10.85 |
| **126** | -12.47 | -11.66 | -11.54 | - | -7.95 | -11.59 |
| **127** | -4.07 | -4.17 | -3.37 | -5.41 | -4.36 | -3.85 |
| **128** | -4.83 | -7.51 | -7.91 | -8.81 | -6.94 | -5.22 |
| **129** | -6.71 | -6.45 | -6.40 | -6.99 | -6.64 | -6.46 |
| **130** | -6.93 | -6.00 | -6.22 | -6.51 | -6.09 | -6.66 |
| **131** | -7.03 | -6.63 | -6.00 | -5.78 | -8.34 | -6.09 |
| **132** | -6.25 | -6.58 | -6.02 | -6.67 | -6.08 | -4.80 |
| **133** | -4.81 | -4.70 | -4.66 | -5.96 | -4.40 | -5.13 |
| **134** | -3.20 | -3.97 | -2.63 | -4.40 | -4.02 | -3.25 |
| **135** | -3.05 | -6.01 | -2.68 | -5.67 | -3.15 | -2.88 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **136** | -2.87 | -8.82 | -4.47 | -8.68 | -9.44 | -2.26 |
| **137** | -5.93 | -6.76 | -6.93 | -7.77 | -6.66 | -6.53 |
| **138** | -3.56 | -4.79 | -3.97 | -2.51 | -4.73 | -4.24 |
| **139** | -6.57 | -7.86 | -6.71 | -7.61 | -8.54 | -6.93 |
| **140** | -1.84 | -7.17 | - | - | -3.50 | - |
| **141** | -4.61 | -3.89 | -3.62 | -4.46 | -3.96 | -3.95 |
| **142** | -9.99 | -8.56 | -10.27 | -11.84 | -10.37 | -8.99 |
| **143** | -5.53 | -6.07 | - | - | - | -5.97 |
| **144** | -4.98 | -8.22 | - | -3.19 | - | -5.59 |
| **145** | -7.01 | -5.80 | -5.13 | -5.71 | -7.64 | -6.58 |
| **146** | -1.70 | -1.50 | -1.10 | -1.04 | -1.29 | -1.56 |
| **147** | -4.06 | -3.35 | -2.73 | -2.46 | -2.06 | -5.88 |
| **148** | -7.30 | -7.00 | -7.02 | -7.62 | - | -7.33 |
| **149** | -3.87 | -2.65 | -3.46 | -3.94 | - | -2.57 |
| **150** | -1.25 | -1.38 | -0.99 | -1.78 | -5.13 | -0.99 |
| **151** | -5.47 | -8.38 | -7.43 | -5.92 | -5.52 | -7.24 |
| **152** | -4.35 | -5.79 | -4.00 | -4.47 | - | -4.03 |
| **153** | -9.25 | -9.69 | -8.89 | -9.26 | -9.39 | -9.05 |

- no assignment available

APPENDIX B: Folded and unfolded temperature coefficients for apoSH mutants.

| Residue | A4V F | A4V U | G85R F | G85R U | E100G F | E100G U | V148G F | V148G U |
|---|---|---|---|---|---|---|---|---|
| **1** | - | - | - | - | - | - | - | - |
| **2** | - | - | -1.57 | - | - | - | - | - |
| **3** | -7.54 | -8.08 | -8.49 | -9.22 | -7.79 | -9.29 | -9.09 | -8.54 |
| **4** | -6.76 | -2.63 | -1.17 | -0.69 | -0.18 | -1.39 | -0.37 | -0.53 |
| **5** | -5.41 | -5.22 | -4.12 | -3.86 | -3.36 | -4.98 | -3.74 | -4.26 |
| **6** | - | - | -3.60 | -2.54 | -2.72 | -3.85 | -3.04 | -3.18 |
| **7** | -6.90 | -8.05 | -7.01 | -8.20 | -6.28 | -8.45 | -8.90 | -10.14 |
| **8** | -4.26 | -3.85 | -4.25 | -2.77 | -4.11 | -4.23 | -5.54 | -4.07 |
| **9** | -6.48 | -5.83 | -6.58 | - | - | - | - | - |
| **10** | -10.04 | -10.21 | -8.86 | -9.41 | -9.33 | -10.72 | -9.79 | -9.73 |
| **11** | -1.73 | -3.16 | -6.52 | -6.06 | -6.52 | -6.90 | -6.60 | -6.53 |
| **12** | 0.96 | 0.65 | -0.25 | 0.38 | 1.28 | -0.27 | 1.22 | 0.04 |
| **13** | - | - | - | - | - | - | - | - |
| **14** | -5.36 | -4.59 | -4.95 | -4.35 | -4.70 | -5.41 | -5.17 | -4.49 |
| **15** | 0.25 | 1.47 | 0.45 | 0.88 | 0.68 | 2.87 | 0.64 | 0.58 |
| **16** | 2.20 | 1.96 | -0.50 | -0.52 | 1.97 | 3.62 | 1.53 | 1.15 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **17** | -4.86 | -1.98 | -2.19 | -1.60 | -1.71 | -1.92 | -3.03 | -2.19 |
| **18** | -2.69 | -3.54 | -4.64 | -4.25 | -3.24 | -4.88 | -3.60 | -3.99 |
| **19** | -2.11 | -1.53 | -2.49 | -3.93 | -1.40 | -2.08 | -2.72 | -2.01 |
| **20** | -1.37 | -0.98 | -1.47 | -1.75 | -1.20 | -1.74 | -1.87 | -1.12 |
| **21** | -3.06 | -3.05 | -3.00 | -2.57 | -3.12 | -4.09 | -2.67 | -2.99 |
| **22** | -6.18 | -6.18 | -4.39 | -4.12 | -4.29 | -5.67 | -5.00 | -5.14 |
| **23** | -3.59 | -5.74 | - | - | - | - | -1.87 | - |
| **24** | -6.82 | -6.22 | -7.09 | -5.49 | -6.48 | -6.69 | - | - |
| **25** | - | - | - | - | - | - | 3.98 | -0.88 |
| **26** | -5.50 | -5.20 | -7.72 | -7.08 | -5.33 | -6.07 | -5.88 | -5.33 |
| **27** | -4.43 | -4.37 | -4.21 | -3.62 | -3.74 | -4.88 | -3.98 | -4.12 |
| **28** | - | - | - | - | - | - | - | - |
| **29** | -5.50 | -5.80 | -4.44 | -5.04 | -4.56 | -5.61 | -3.99 | -5.18 |
| **30** | -3.12 | -3.27 | -3.35 | -2.96 | -2.65 | -3.67 | -2.85 | -2.77 |
| **31** | -1.93 | -2.26 | -3.42 | -2.38 | - | - | -0.63 | -1.68 |
| **32** | -1.77 | -2.44 | -2.91 | -3.20 | -2.38 | -3.98 | -2.35 | -2.88 |
| **33** | -5.79 | -5.47 | -5.66 | -4.70 | -5.18 | -5.92 | -4.55 | -5.35 |
| **34** | -0.69 | -0.62 | -0.52 | 0.26 | 0.96 | 0.65 | -0.14 | 0.00 |
| **35** | -0.78 | -0.35 | -0.32 | -0.39 | -0.73 | -0.50 | -1.04 | -0.47 |
| **36** | -3.11 | -3.41 | -2.78 | -2.14 | -2.55 | -3.98 | -2.83 | -3.25 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **37** | -3.24 | -3.07 | -3.39 | -2.00 | -2.13 | -4.02 | -2.42 | -3.44 |
| **38** | - | - | - | - | -5.63 | -7.03 | -9.86 | -7.63 |
| **39** | -4.85 | -17.46 | - | - | -19.59 | -17.29 | - | -17.61 |
| **40** | - | - | -7.10 | -9.72 | - | - | -8.24 | -9.72 |
| **41** | -6.90 | -7.72 | -7.52 | -8.49 | -6.67 | -7.66 | -7.02 | -7.66 |
| **42** | -8.06 | -7.23 | -6.44 | - | - | - | -8.75 | -7.02 |
| **43** | -3.19 | -2.94 | - | - | - | - | 1.14 | 0.54 |
| **44** | -2.88 | -1.22 | -9.69 | -9.48 | -2.93 | -3.01 | -1.99 | -0.94 |
| **45** | -7.74 | -7.76 | -8.42 | -4.07 | -11.76 | -13.79 | -15.70 | -12.97 |
| **46** | - | - | - | - | -1.34 | -1.14 | -3.79 | -1.35 |
| **47** | -5.59 | -6.27 | - | - | -4.80 | -6.14 | -4.14 | -5.25 |
| **48** | -4.43 | -5.51 | -4.40 | -4.21 | -3.44 | -5.40 | -5.26 | -6.20 |
| **49** | - | - | - | - | - | - | - | - |
| **50** | - | - | - | - | - | - | - | - |
| **51** | -6.70 | -8.49 | -6.13 | -7.16 | -6.00 | -5.34 | -5.72 | -6.64 |
| **52** | -6.14 | -5.28 | -5.25 | -4.81 | -5.13 | -5.08 | -5.57 | -8.70 |
| **53** | -7.13 | -7.52 | -7.52 | -6.19 | -5.59 | -6.82 | -5.65 | -7.55 |
| **54** | -4.04 | -3.98 | -4.34 | -3.76 | -3.45 | -4.21 | -3.97 | -3.93 |
| **55** | -8.40 | -5.53 | -6.95 | -5.47 | -6.30 | -6.25 | -6.65 | -5.62 |
| **56** | -6.20 | -5.05 | -6.06 | -5.86 | -5.72 | -5.23 | -6.29 | -5.18 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **57** | -4.10 | -4.85 | -3.61 | - | -3.55 | -3.11 | -4.09 | -4.64 |
| **58** | -6.40 | -6.26 | -3.71 | -4.38 | -5.54 | -6.14 | -6.37 | -6.32 |
| **59** | -8.50 | -7.78 | -6.24 | -6.92 | -7.91 | -8.25 | - | - |
| **60** | -5.47 | -7.67 | -6.91 | -6.29 | -7.31 | -8.75 | -7.73 | -6.23 |
| **61** | -6.15 | -4.50 | -5.74 | -6.73 | -5.28 | -5.47 | -6.16 | -5.14 |
| **62** | - | - | - | - | - | - | - | - |
| **63** | - | - | - | - | - | - | - | - |
| **64** | - | - | - | - | - | - | - | - |
| **65** | - | - | - | - | - | - | - | - |
| **66** | - | - | - | - | - | - | - | - |
| **67** | -4.56 | -4.17 | -3.72 | -4.70 | -3.64 | -4.05 | -5.90 | -4.08 |
| **68** | -3.27 | -2.60 | -2.60 | -3.57 | -2.60 | -3.52 | -3.22 | -4.98 |
| **69** | -5.85 | -5.44 | -8.07 | -6.88 | -7.57 | -8.17 | -5.83 | -5.62 |
| **70** | -7.72 | -6.93 | -7.60 | -6.35 | - | - | -8.33 | -6.64 |
| **71** | - | - | - | - | - | - | - | - |
| **72** | - | - | - | - | - | - | - | - |
| **73** | -5.34 | -4.85 | -5.72 | -4.01 | -5.65 | -5.98 | -5.75 | -4.84 |
| **74** | - | - | - | - | - | - | - | - |
| **75** | -7.61 | -6.23 | -5.55 | -5.26 | -7.13 | -6.48 | -7.27 | -6.53 |
| **76** | -6.07 | -4.52 | -7.67 | -7.64 | - | - | - | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **77** | -7.12 | -6.33 | -7.99 | -6.85 | -5.00 | -7.92 | -6.78 | -7.11 |
| **78** | -5.66 | -6.12 | -7.08 | - | -7.54 | -8.04 | -7.62 | -7.59 |
| **79** | - | - | - | - | - | - | - | - |
| **80** | - | - | - | - | - | - | - | - |
| **81** | - | - | - | - | - | - | - | - |
| **82** | - | - | - | - | - | - | - | - |
| **83** | - | - | - | - | - | - | - | - |
| **84** | - | - | - | - | - | - | - | - |
| **85** | -5.98 | -5.36 | - | - | - | - | -5.42 | -5.41 |
| **86** | -5.76 | -5.48 | - | - | -5.31 | -5.74 | -5.89 | -5.58 |
| **87** | -2.26 | -3.01 | -5.17 | -5.54 | -2.69 | -4.56 | -1.87 | -3.04 |
| **88** | -5.94 | -5.74 | -5.24 | -5.37 | -5.08 | -6.24 | -5.93 | -5.77 |
| **89** | -4.89 | -5.99 | -4.68 | -3.85 | -5.12 | -6.64 | -4.73 | -6.01 |
| **90** | -4.56 | -4.63 | -4.95 | -4.20 | -3.76 | -5.07 | -5.21 | -4.52 |
| **91** | -6.68 | -5.86 | -6.37 | -5.48 | -5.70 | -5.74 | -6.39 | -5.68 |
| **92** | -1.88 | -1.15 | -2.57 | -1.86 | -1.17 | -0.98 | -1.61 | -0.84 |
| **93** | -1.82 | -1.14 | -1.34 | -0.75 | -0.58 | -1.60 | -0.99 | -0.83 |
| **94** | -1.21 | -4.45 | -0.93 | -1.21 | -0.80 | -0.97 | -0.96 | -4.58 |
| **95** | -7.91 | -7.98 | -7.06 | -7.09 | -6.97 | -8.84 | -7.78 | -8.24 |
| **96** | -4.18 | -5.61 | -4.87 | -4.98 | -4.33 | -5.76 | -4.14 | -5.56 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **97** | -3.96 | -4.02 | -4.51 | -5.56 | -2.20 | -2.98 | -3.28 | -5.10 |
| **98** | -8.58 | -8.00 | -7.63 | -7.52 | -7.61 | -8.84 | -8.68 | -8.12 |
| **99** | -3.22 | -3.00 | -2.90 | -2.80 | -2.63 | -4.68 | -1.89 | -2.39 |
| **100** | -8.13 | -6.32 | -9.42 | -8.25 | - | - | -7.49 | -6.45 |
| **101** | -4.28 | -7.35 | -7.12 | -8.19 | -6.21 | -8.67 | -4.71 | -5.20 |
| **102** | -7.35 | -7.26 | -7.09 | -6.04 | - | - | -7.55 | -7.07 |
| **103** | -7.69 | -6.80 | - | - | - | - | -6.79 | -6.86 |
| **104** | -4.66 | -3.39 | -6.22 | -5.62 | - | - | -4.75 | -4.41 |
| **105** | -4.60 | -1.82 | -8.11 | - | -1.72 | -1.98 | -0.14 | -0.54 |
| **106** | -7.83 | -7.78 | -5.88 | -3.23 | - | - | -5.62 | -5.90 |
| **107** | -4.08 | -4.21 | -1.55 | -1.95 | -1.20 | -2.89 | -1.08 | -1.70 |
| **108** | -6.01 | -7.01 | -5.98 | - | -5.80 | - | -6.15 | - |
| **109** | -8.12 | -7.10 | -8.69 | -9.52 | -7.61 | -9.66 | -8.09 | -5.85 |
| **110** | -7.20 | -2.93 | -7.83 | -9.97 | -6.60 | -7.39 | -6.53 | -8.17 |
| **111** | -0.80 | -0.89 | -1.09 | -0.77 | - | - | - | -0.72 |
| **112** | -5.57 | -4.30 | -4.25 | -3.71 | -4.32 | -5.62 | -4.22 | -3.83 |
| **113** | -3.43 | -3.31 | -3.87 | -2.72 | -2.70 | -4.05 | -2.89 | -3.22 |
| **114** | -2.31 | -1.70 | -2.29 | -2.43 | -2.07 | -3.21 | -6.53 | -5.72 |
| **115** | -4.83 | -4.11 | -3.78 | -1.91 | -7.95 | -7.90 | -3.59 | -3.12 |
| **116** | -0.63 | -0.34 | -7.01 | -6.23 | - | - | -6.89 | -6.89 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **117** | -2.95 | -2.73 | -3.60 | -2.88 | -2.45 | -3.38 | -3.24 | -2.75 |
| **118** | -1.45 | -1.51 | - | - | -0.44 | -2.00 | -0.87 | -1.26 |
| **119** | -6.63 | -5.93 | - | - | - | - | - | - |
| **120** | -4.65 | -7.80 | - | - | - | - | -7.38 | -7.66 |
| **121** | -9.91 | -8.01 | -7.23 | -7.89 | -3.48 | -7.67 | -8.67 | -7.71 |
| **122** | - | - | - | - | - | - | - | - |
| **123** | -9.38 | -8.32 | -6.20 | -6.57 | -5.77 | -4.12 | -9.51 | -8.63 |
| **124** | - | - | - | - | - | - | - | - |
| **125** | -8.03 | -7.83 | - | - | -7.28 | -8.22 | - | - |
| **126** | -12.08 | -11.80 | - | - | -11.55 | -12.55 | -7.84 | -7.68 |
| **127** | -3.72 | -5.75 | -5.05 | -4.71 | -3.60 | -5.95 | -3.85 | -6.64 |
| **128** | -7.78 | -4.77 | -8.27 | -4.00 | -6.70 | -8.04 | -7.23 | -4.31 |
| **129** | -6.59 | -6.63 | -6.77 | -6.76 | -6.08 | -7.87 | -6.58 | -6.50 |
| **130** | -6.10 | -5.51 | -6.44 | -4.38 | -5.80 | -6.07 | -6.05 | -5.60 |
| **131** | -6.53 | -6.00 | -5.94 | -4.19 | -6.04 | -6.01 | -7.01 | -4.18 |
| **132** | -6.69 | -5.11 | -6.37 | -5.06 | -5.62 | -5.66 | -6.25 | -6.94 |
| **133** | -4.75 | -3.22 | -5.92 | -6.77 | -4.25 | -5.54 | -4.49 | -5.24 |
| **134** | -4.23 | -4.95 | -4.05 | -3.35 | -3.18 | -4.92 | -3.59 | -5.17 |
| **135** | -6.30 | -5.24 | -5.41 | -4.86 | -2.73 | -4.43 | -2.75 | -4.82 |
| **136** | -8.86 | -7.37 | -8.63 | -7.71 | -4.95 | -6.90 | -9.06 | -7.14 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **137** | -7.00 | -6.14 | -8.40 | -8.39 | -6.47 | -6.55 | -6.92 | -6.43 |
| **138** | -5.01 | -5.18 | -3.06 | -3.87 | -4.23 | -8.72 | -5.21 | -5.42 |
| **139** | -8.37 | -7.16 | -7.60 | -8.88 | -6.93 | -6.79 | -7.99 | -4.89 |
| **140** | -7.17 | -3.57 | -1.96 | -2.10 | -1.91 | -3.93 | -3.01 | -2.85 |
| **141** | -4.35 | -3.17 | -4.53 | -4.01 | -3.11 | -3.80 | -4.20 | -3.29 |
| **142** | -8.57 | -9.53 | -11.87 | -13.37 | -9.85 | -10.04 | -9.72 | -9.24 |
| **143** | -5.72 | -8.02 | -8.75 | -8.51 | -5.87 | -7.74 | -4.04 | -6.33 |
| **144** | -8.07 | -8.58 | -2.63 | - | -5.04 | -6.95 | - | - |
| **145** | -5.57 | -3.70 | -6.19 | -5.65 | -5.45 | -7.59 | -5.55 | -6.61 |
| **146** | -1.44 | -1.33 | -1.46 | -0.27 | -1.16 | -2.14 | -1.29 | -1.39 |
| **147** | -3.30 | -1.70 | -2.13 | -2.03 | -2.32 | -2.71 | -1.72 | -2.21 |
| **148** | -7.37 | -7.50 | -7.86 | -7.52 | -6.89 | -8.40 | - | - |
| **149** | -2.46 | -2.51 | -4.31 | -3.91 | -3.52 | -4.88 | - | - |
| **150** | -2.88 | - | -1.54 | -0.96 | -0.88 | -1.62 | -5.47 | - |
| **151** | -8.33 | -7.25 | -6.51 | -6.94 | -6.46 | -8.30 | -5.51 | -5.18 |
| **152** | -5.90 | -5.88 | -4.81 | -4.24 | -3.87 | -4.92 | - | - |
| **153** | -9.88 | -9.15 | -9.29 | -8.44 | -8.57 | -8.94 | -9.44 | -8.46 |

- no assignment available

F and U denote folded and unfolded temperature ranges for the proteins based on Table 3.6

APPENDIX C: Urea denatured SOD1 and calculated random coil temperature coefficients.

| Residue | Amino Acid | $^{1}H$ | $^{15}N$ | RC TC |
|---|---|---|---|---|
| **1** | | | | |
| **2** | | | | -8.10 |
| **3** | LYS | 8.57 | 124.35 | -7.40 |
| **4** | ALA | 8.50 | 126.19 | -8.00 |
| **5** | VAL | 8.22 | 119.77 | -8.50 |
| **6** | ALA | 8.43 | 127.87 | -8.80 |
| **7** | VAL | 8.23 | 120.42 | -8.50 |
| **8** | LEU | 8.49 | 126.88 | -8.70 |
| **9** | LYS | 8.59 | 123.07 | -8.20 |
| **10** | GLY | 8.50 | 110.32 | -7.00 |
| **11** | ASP | 8.38 | 120.31 | -7.00 |
| **12** | GLY | 8.29 | 108.63 | -6.70 |
| **13** | | 0.00 | 0.00 | 0.00 |
| **14** | VAL | 8.40 | 120.76 | -9.40 |
| **15** | GLN | 8.59 | 124.34 | -8.90 |

| | | | | |
|---|---|---|---|---|
| **16** | GLY | 8.48 | 110.31 | -7.30 |
| **17** | ILE | 8.12 | 120.38 | -7.60 |
| **18** | ILE | 8.39 | 125.47 | -9.10 |
| **19** | ASN | 8.55 | 123.41 | -8.30 |
| **20** | PHE | 8.41 | 121.86 | -7.70 |
| **21** | GLU | 8.50 | 122.26 | -6.90 |
| **22** | GLN | 8.45 | 122.10 | -7.60 |
| **23** | LYS | 8.53 | 123.39 | -8.00 |
| **24** | GLU | 8.68 | 122.15 | -7.00 |
| **25** | SER | 8.51 | 116.88 | -7.80 |
| **26** | ASN | 8.61 | 120.98 | -7.30 |
| **27** | GLY | 8.27 | 108.76 | -6.60 |
| **28** | | | | |
| **29** | VAL | 8.34 | 120.86 | -9.40 |
| **30** | LYS | 8.50 | 126.19 | -8.50 |
| **31** | VAL | 8.25 | 122.12 | -8.30 |
| **32** | TRP | 8.51 | 125.83 | -9.40 |
| **33** | GLY | 8.42 | 110.59 | -6.60 |
| **34** | SER | 8.29 | 115.57 | -7.30 |
| **35** | ILE | 8.38 | 122.89 | -7.80 |

| | | | | |
|---|---|---|---|---|
| **36** | LYS | 8.53 | 125.89 | -8.80 |
| **37** | GLY | 8.50 | 110.34 | -7.00 |
| **38** | LEU | 8.23 | 121.69 | -7.20 |
| **39** | THR | 8.35 | 114.37 | -8.10 |
| **40** | GLU | 8.51 | 123.37 | -7.00 |
| **41** | GLY | 8.47 | 109.76 | -6.80 |
| **42** | LEU | 8.16 | 121.39 | -7.20 |
| **43** | HIS | 8.59 | 119.74 | -8.20 |
| **44** | GLY | 8.42 | 109.69 | -6.70 |
| **45** | PHE | 8.29 | 120.18 | -7.40 |
| **46** | HIS | 8.56 | 121.80 | -7.60 |
| **47** | VAL | 8.28 | 122.32 | -8.00 |
| **48** | HIS | 8.70 | 123.25 | -8.50 |
| **49** | GLU | 8.67 | 123.63 | -6.70 |
| **50** | GLU | 8.68 | 121.93 | -8.20 |
| **51** | GLU | 8.53 | 121.78 | -6.90 |
| **52** | ASP | 8.50 | 121.45 | -6.70 |
| **53** | ASN | 8.62 | 120.68 | -7.20 |
| **54** | THR | 8.38 | 113.60 | -7.50 |
| **55** | ALA | 8.27 | 125.75 | -8.00 |

| | | | | |
|---|---|---|---|---|
| **56** | GLY | 8.34 | 107.31 | -7.20 |
| **57** | CYS | 8.44 | 118.51 | -6.60 |
| **58** | THR | 8.49 | 115.57 | -8.20 |
| **59** | SER | 8.49 | 118.15 | -7.70 |
| **60** | ALA | 8.41 | 125.81 | -7.80 |
| **61** | GLY | 8.22 | 107.75 | -7.20 |
| **62** | | | | 0.00 |
| **63** | HIS | | | -8.80 |
| **64** | PHE | | | -7.80 |
| **65** | ASN | | | -7.40 |
| **66** | | | | -0.30 |
| **67** | LEU | 8.28 | 120.85 | -9.00 |
| **68** | SER | 8.21 | 116.28 | -8.20 |
| **69** | ARG | 8.47 | 123.34 | -3.90 |
| **70** | LYS | | | -7.70 |
| **71** | HIS | | | -8.00 |
| **72** | GLY | | | -7.00 |
| **73** | GLY | 8.40 | 108.95 | -6.30 |
| **74** | | | | 0.00 |
| **75** | LYS | 8.60 | 121.19 | -8.80 |

| | | | | |
|---|---|---|---|---|
| **76** | ASP | 8.45 | 121.51 | -7.40 |
| **77** | GLU | 8.47 | 120.82 | -6.70 |
| **78** | GLU | 8.44 | 121.99 | -6.80 |
| **79** | ARG | | | -4.20 |
| **80** | HIS | | | -7.70 |
| **81** | VAL | | | -8.30 |
| **82** | GLY | 8.53 | 112.17 | -7.80 |
| **83** | ASP | 8.31 | 120.56 | -6.70 |
| **84** | LEU | 8.43 | 122.69 | -7.90 |
| **85** | GLY | 8.52 | 108.65 | -7.20 |
| **86** | ASN | 8.39 | 118.88 | -6.80 |
| **87** | VAL | 8.25 | 120.29 | -7.90 |
| **88** | THR | 8.32 | 117.50 | -8.70 |
| **89** | ALA | 8.36 | 126.20 | -7.70 |
| **90** | ASP | 8.42 | 119.96 | -7.60 |
| **91** | LYS | 8.38 | 121.42 | -7.40 |
| **92** | ASP | 8.45 | 120.76 | -7.70 |
| **93** | GLY | 8.33 | 108.88 | -6.70 |
| **94** | VAL | 8.09 | 119.17 | -7.60 |
| **95** | ALA | 8.48 | 127.62 | -8.80 |

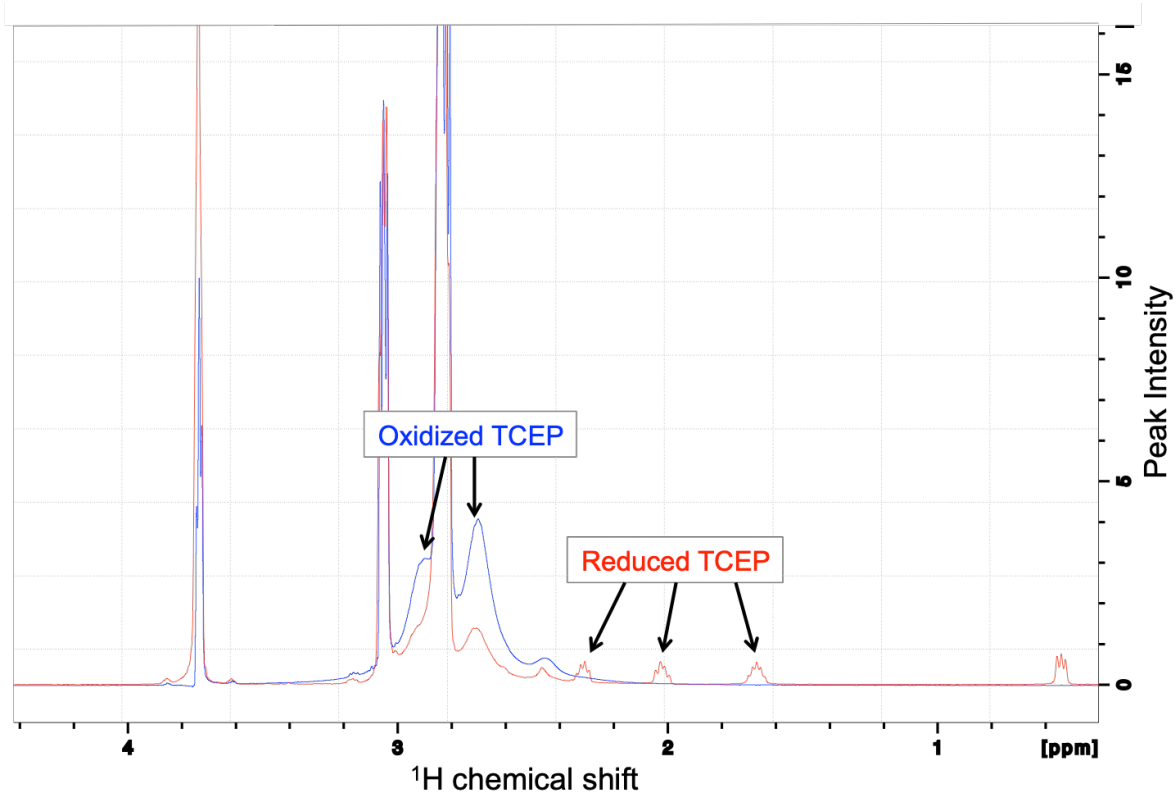
| | | | | |
|---|---|---|---|---|
| **96** | ASP | 8.41 | 120.45 | -7.60 |
| **97** | VAL | 8.14 | 119.46 | -8.00 |
| **98** | SER | 8.58 | 119.61 | -8.80 |
| **99** | ILE | 8.31 | 123.04 | -7.80 |
| **100** | GLU | 8.55 | 124.15 | -7.80 |
| **101** | ASP | 8.37 | 121.46 | -7.20 |
| **102** | SER | 8.30 | 115.82 | -7.70 |
| **103** | VAL | 8.21 | 121.75 | -7.80 |
| **104** | ILE | 8.25 | 123.97 | -9.10 |
| **105** | SER | 8.46 | 119.79 | -8.80 |
| **106** | LEU | 8.51 | 125.17 | -7.40 |
| **107** | SER | 8.40 | 115.77 | -8.50 |
| **108** | GLY | 8.44 | 110.58 | -6.50 |
| **109** | ASP | 8.32 | 120.31 | -6.70 |
| **110** | HIS | 8.42 | 119.14 | -7.40 |
| **111** | ALA | 8.32 | 124.43 | -7.70 |
| **112** | ILE | 8.28 | 120.61 | -7.80 |
| **113** | ILE | 8.37 | 125.57 | -9.40 |
| **114** | GLY | 8.54 | 113.14 | -7.80 |
| **115** | ARG | 8.31 | 120.66 | -3.70 |

| | | | | |
|---|---|---|---|---|
| **116** | THR | 8.43 | 116.47 | -7.90 |
| **117** | LEU | 8.47 | 125.72 | -7.60 |
| **118** | VAL | 8.37 | 123.07 | -8.50 |
| **119** | VAL | 8.37 | 125.11 | -9.10 |
| **120** | HIS | 8.65 | 124.01 | -8.50 |
| **121** | GLU | 8.67 | 123.32 | -6.70 |
| **122** | LYS | 8.60 | 122.88 | -7.50 |
| **123** | ALA | 8.52 | 125.87 | -8.00 |
| **124** | ASP | 8.43 | 119.81 | -7.60 |
| **125** | ASP | | | -7.10 |
| **126** | LEU | 8.32 | 122.38 | -7.90 |
| **127** | GLY | 8.54 | 109.00 | -7.20 |
| **128** | LYS | 8.22 | 120.52 | -7.30 |
| **129** | GLY | 8.56 | 109.72 | -7.30 |
| **130** | GLY | 8.41 | 108.67 | -6.30 |
| **131** | ASN | 8.52 | 118.64 | -6.80 |
| **132** | GLU | 8.71 | 121.43 | -6.60 |
| **133** | GLU | 8.57 | 121.67 | -6.80 |
| **134** | SER | 8.44 | 116.56 | -7.80 |
| **135** | THR | 8.27 | 115.54 | -7.40 |

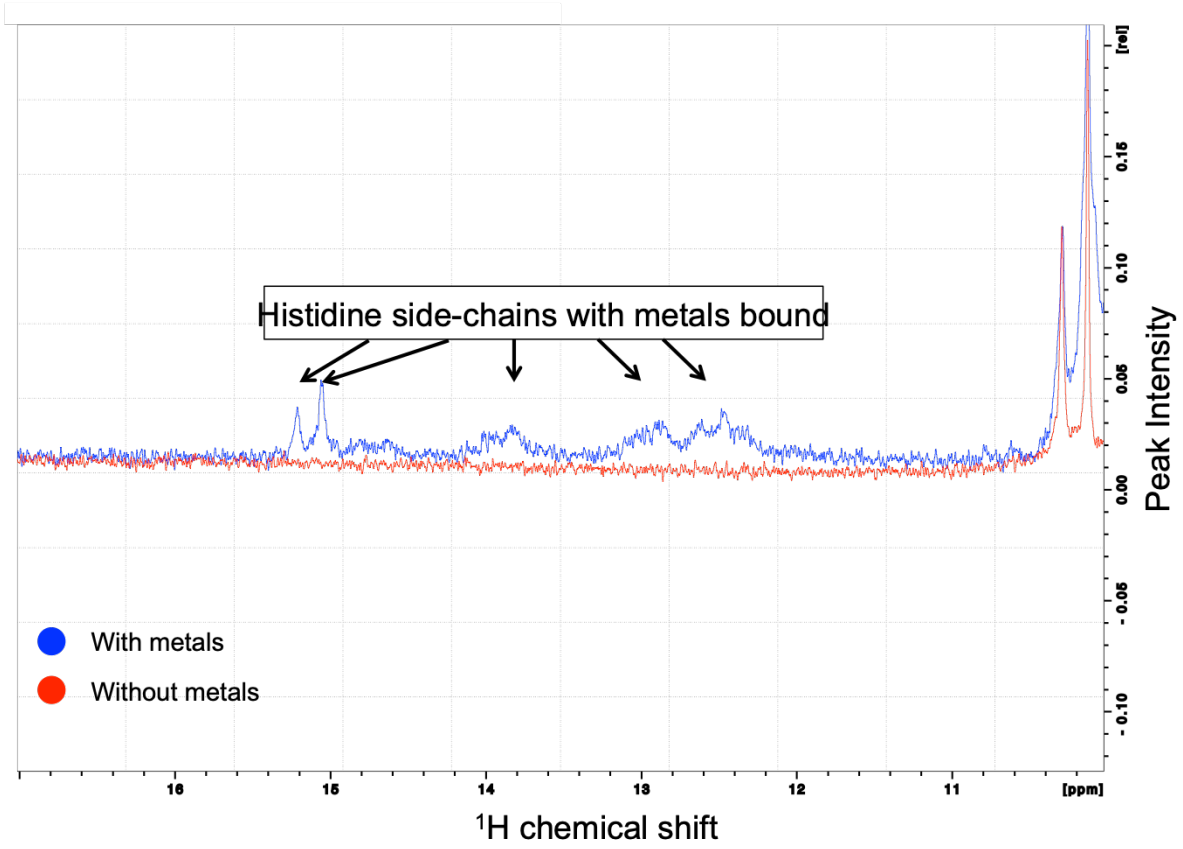
| | | | | |
|---|---|---|---|---|
| **136** | LYS | 8.40 | 123.52 | -7.40 |
| **137** | THR | 8.25 | 114.34 | -8.20 |
| **138** | GLY | 8.46 | 110.63 | -6.70 |
| **139** | ASN | 8.47 | 118.88 | -6.80 |
| **140** | ALA | 8.51 | 124.87 | -7.90 |
| **141** | GLY | 8.42 | 107.65 | -7.20 |
| **142** | SER | 8.27 | 115.57 | -7.30 |
| **143** | ARG | 8.53 | 123.22 | -3.90 |
| **144** | LEU | 8.41 | 123.66 | -7.90 |
| **145** | ALA | 8.49 | 125.23 | -8.20 |
| **146** | CYS | 8.49 | 118.10 | -7.80 |
| **147** | GLY | 8.62 | 110.85 | -7.30 |
| **148** | VAL | 8.10 | 119.43 | -7.60 |
| **149** | ILE | 8.39 | 125.07 | -9.40 |
| **150** | GLY | 8.52 | 113.11 | -7.80 |
| **151** | ILE | 8.08 | 119.94 | -7.60 |
| **152** | ALA | 8.47 | 128.75 | -8.80 |
| **153** | GLN | 8.09 | 125.44 | |

# APPENDIX D: 1 mM oxidized and reduced TCEP in 20 mM HEPES at pH 7.4 $^{1}$H spectra.

Reduced and oxidized peaks used to evaluate extent of reduction potential throughout the experiments.



APPENDIX E: $^{1}H$ spectra of imidazole peak intensity with and without metals.



# APPENDIX F: Pulse Program: zgesgp

```
;zgesgp
;avance-version (12/01/11)
;1D sequence
;water suppression using excitation sculpting with gradients
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;   Series A 112 275-279 (1995)
;
;$CLASS=HighRes
;$DIM=1D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


prosol relations=<triple>


#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"d12=20u"


"TAU=de+p1*2/3.1416+50u"


"acqt0=0"
baseopt_echo


1 ze
2 30m
  d12 pl1:f1 BLKGRAD
  d1
  p1 ph1

  50u UNBLKGRAD
  p16:gp1
```

```
  d16 pl0:f1
  (p12:sp1 ph2:r):f1
  4u
  d12 pl1:f1

  p2 ph3

  4u
  p16:gp1
  d16
  TAU
  p16:gp2
  d16 pl0:f1
  (p12:sp1 ph4:r):f1
  4u
  d12 pl1:f1

  p2 ph5

  4u
  p16:gp2
  d16

  go=2 ph31
  30m mc #0 to 2 F0(zd)
  4u BLKGRAD
exit


ph1=0
ph2=0 1
ph3=2 3
ph4=0 0 1 1
ph5=2 2 3 3
ph31=0 2 2 0


;pl0 : 0W
;pl1 : f1 channel - power level for pulse (default)
;sp1 : f1 channel - shaped pulse 180 degree
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p12: f1 channel - 180 degree shaped pulse (Squa100.1000)   [2 msec]
;p16: homospoil/gradient pulse
```

```
;d1 : relaxation delay; 1-5 * T1
;d12: delay for power switching                  [20 usec]
;d16: delay for homospoil/gradient recovery
;ns: 8 * n, total number of scans: NS * TD0
;ds: 4


;use gradient ratio:    gp 1 : gp 2
;                       31 :   11

;for z-only gradients:
;gpz1: 31%
;gpz2: 11%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100



;$Id: zgesgp,v 1.9 2012/01/31 17:49:32 ber Exp $
```

# APPENDIX G: Pulse program: hsqcfpf3gpphwg

```
;hsqcfpf3gpphwg
;avance-version (12/01/11)
;HSQC
;2D H-1/X correlation via double inept transfer
;phase sensitive
;with decoupling during acquisition
;using f3 - channel
;using flip-back pulse
;water suppression using watergate sequence
;similar to fhsqc
;(use parameterset HSQCFPF3GPPHWG)
;
;G. Bodenhausen & D.J. Ruben, Chem. Phys. Lett. 69, 185 (1980)
;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2, 661 - 666 (1992)
;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson.,
;   Series A 102, 241 -245 (1993)
;S. Mori, C. Abeygunawardana, M. O'Neil-Johnson & P.C.M. van Zijl,
;   J. Magn. Reson. B 108, 94-98 (1995)
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


prosol relations=<triple>


#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>


"p2=p1*2"
"p22=p21*2"
"d11=30m"
"d12=20u"
"d26=1s/(cnst4*4)"
```

```
"d0=3u"

"in0=inf1/2"


"DELTA1=d26-p16-d16"
"DELTA2=d26-p16-d16-p11-12u"

#   ifdef LABEL_CN
"DELTA=d0*2+larger(p2,p8)"
#   else
"DELTA=d0*2+p2"
#   endif /*LABEL_CN*/


1 ze
  d11 pl16:f3
2 d1 do:f3
3 d12 pl1:f1 pl3:f3
  50u UNBLKGRAD

  (p1 ph1)
  p16:gp3
  d16
  DELTA1
  (center (p2 ph2) (p22 ph6):f3 )
  DELTA1
  p16:gp3
  d16
  (p1 ph2)

  4u pl0:f1
  (p11:sp1 ph8:r):f1
  4u
  p16:gp4
  d16 pl1:f1

  (p21 ph3):f3
  d0

#   ifdef LABEL_CN
  (center (p2 ph5) (p8:sp13 ph1):f2 )
#   else
  (p2 ph5)
```

```
#   endif /*LABEL_CN*/

  d0
  (p22 ph4):f3
  DELTA
  (p21 ph4):f3

  4u
  p16:gp4
  d16 pl0:f1
  (p11:sp1 ph7:r):f1
  4u
  4u pl1:f1

  (p1 ph1)
  4u
  p16:gp5
  d16
  DELTA2 pl0:f1
  (p11:sp1 ph7:r):f1
  4u
  4u pl1:f1
  (center (p2 ph1) (p22 ph1):f3 )
  4u pl0:f1
  (p11:sp1 ph7:r):f1
  4u
  p16:gp5
  d16
  DELTA2 pl16:f3
  4u BLKGRAD

  go=2 ph31 cpd3:f3
  d1 do:f3 mc #0 to 2 F1PH(calph(ph3, +90) & calph(ph6, +90), caldel(d0, +in0))
exit


ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph7=2
ph8=0
```

```
ph31=0 2 0 2 2 0 2 0


;pl0 : 0W
;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl16: f3 channel - power level for CPD/BB decoupling
;sp1: f1 channel - shaped pulse  90 degree
;sp13: f2 channel - shaped pulse 180 degree (adiabatic)
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p8 : f2 channel - 180 degree shaped pulse for inversion (adiabatic)
;p11: f1 channel -  90 degree shaped pulse
;p16: homospoil/gradient pulse
;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;d0 : incremented delay (2D)                    [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                       [30 msec]
;d12: delay for power switching                 [20 usec]
;d16: delay for homospoil/gradient recovery
;d26 : 1/(4J)YH
;cnst4: = J(YH)
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 4 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence


;use gradient ratio:    gp 3 : gp 4 : gp 5
;                       50 :   80 :   30

;for z-only gradients:
;gpz3: 50%
;gpz4: 80%
;gpz5: 30%

;use gradient files:
;gpnam3: SMSQ10.100
```

```
;gpnam4: SMSQ10.100
;gpnam5: SMSQ10.100


                                ;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labeled samples start experiment with
;               option -DLABEL_CN (eda: ZGOPTNS)
                                ;preprocessor-flags-end



;$Id: hsqcfpf3gpphwg,v 1.12 2012/01/31 17:49:26 ber Exp $
```

# APPENDIX H: Pulse program: p3919gp

```
;p3919gp
;avance-version (12/01/11)
;1D sequence
;water suppression using 3-9-19 pulse sequence with gradients
;
;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2, 661 - 666 (1992)
;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson.,
;   Series A 102, 241 -245 (1993)
;
;$CLASS=HighRes
;$DIM=1D
;$TYPE=
;$SUBTYPE=
;$COMMENT=



#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>



"TAU=p1*2/3.1416-(p0-p27)*0.231-de+46u"
```

```
"acqt0=0"
baseopt_echo


1 ze
2 30m pl1:f1
  d1
  p1 ph1
  50u UNBLKGRAD
  p16:gp1
  d16 pl18:f1
  p27*0.231 ph3
  d19*2
  p27*0.692 ph3
  d19*2
  p27*1.462 ph3
  d19*2
  p27*1.462 ph4
  d19*2
  p27*0.692 ph4
  d19*2
```

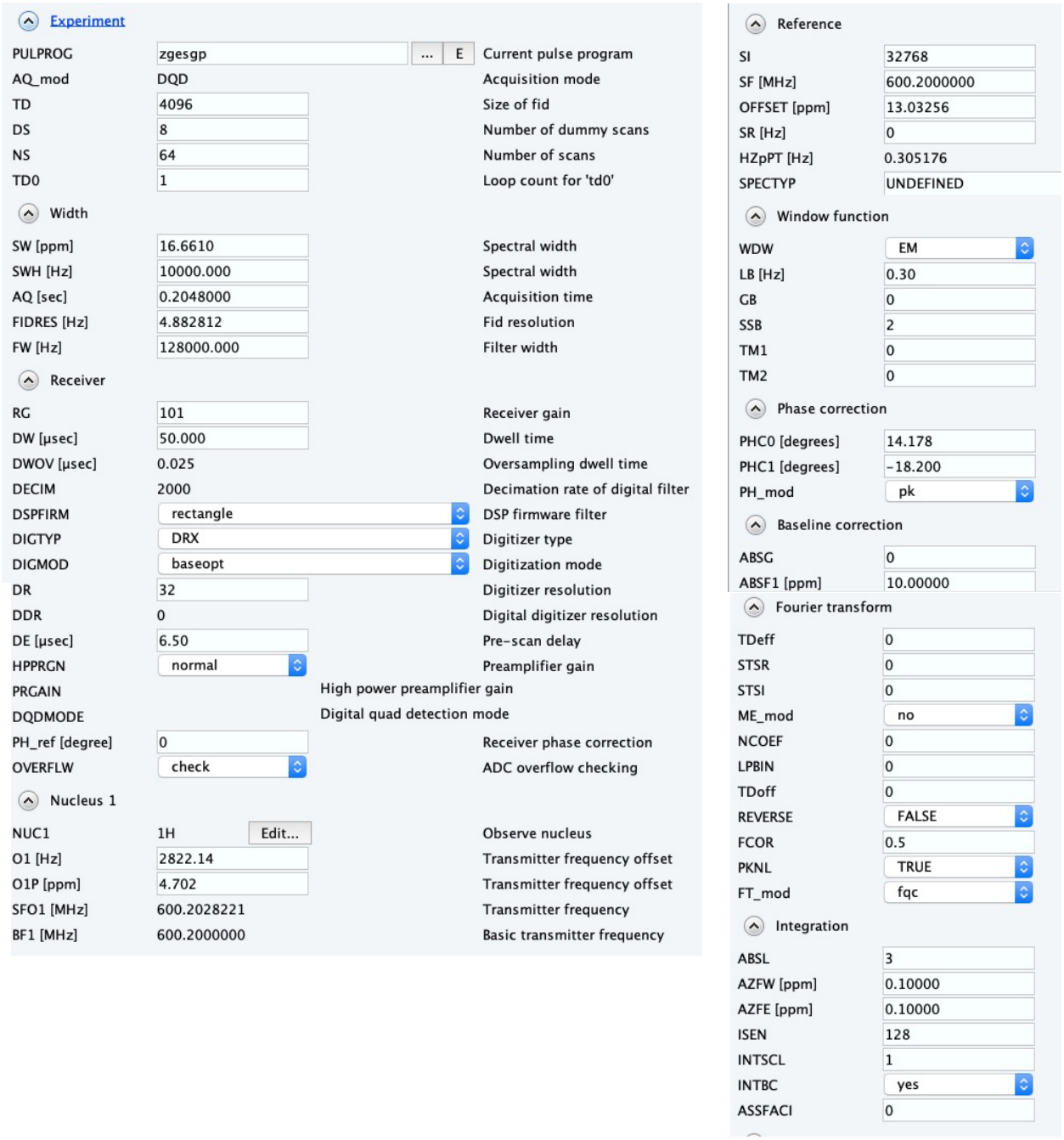
```
  p0*0.231 ph4
  TAU
  p16:gp1
  d16
  4u BLKGRAD
  go=2 ph31
  30m pl1:f1 mc #0 to 2 F0(zd)
exit


ph1=0 2
ph3=0 0 1 1 2 2 3 3
ph4=2 2 3 3 0 0 1 1
ph31=0 2 2 0


;pl1 : f1 channel - power level for pulse (default)
;pl18: f1 channel - power level for 3-9-19-pulse (watergate)
;p0 : f1 channel -  90 degree pulse at pl18
;                   use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p16: homospoil/gradient pulse
;p27: f1 channel -  90 degree pulse at pl18
```

```
;d1 : relaxation delay; 1-5 * T1
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;ns: 8 * n, total number of scans: NS * TD0
;ds: 4



;use gradient ratio:   gp 1
;                      20

;for z-only gradients:
;gpz1: 20%

;use gradient files:
;gpnam1: SMSQ10.100



;$Id: p3919gp,v 1.14 2012/01/31 17:49:28 ber Exp $
```

# APPENDIX I: Acquisition and processing parameters for 1D spectra.

ACQUISITION

PROCESSING



Experiment

| | | | |
|---|---|---|---|
| PULPROG | zgesgp | ... E | Current pulse program |
| AQ_mod | DQD | | Acquisition mode |
| TD | 4096 | | Size of fid |
| DS | 8 | | Number of dummy scans |
| NS | 64 | | Number of scans |
| TD0 | 1 | | Loop count for 'td0' |

Width

| | | |
|---|---|---|
| SW [ppm] | 16.6610 | Spectral width |
| SWH [Hz] | 10000.000 | Spectral width |
| AQ [sec] | 0.2048000 | Acquisition time |
| FIDRES [Hz] | 4.882812 | Fid resolution |
| FW [Hz] | 128000.000 | Filter width |

Receiver

| | | |
|---|---|---|
| RG | 101 | Receiver gain |
| DW [µsec] | 50.000 | Dwell time |
| DWOV [µsec] | 0.025 | Oversampling dwell time |
| DECIM | 2000 | Decimation rate of digital filter |
| DSPFIRM | rectangle | DSP firmware filter |
| DIGTYP | DRX | Digitizer type |
| DIGMOD | baseopt | Digitization mode |
| DR | 32 | Digitizer resolution |
| DDR | 0 | Digital digitizer resolution |
| DE [µsec] | 6.50 | Pre-scan delay |
| HPPRGN | normal | Preamplifier gain |
| PRGAIN | | High power preamplifier gain |
| DQDMODE | | Digital quad detection mode |
| PH_ref [degree] | 0 | Receiver phase correction |
| OVERFLW | check | ADC overflow checking |

Nucleus 1

| | | |
|---|---|---|
| NUC1 | 1H Edit... | Observe nucleus |
| O1 [Hz] | 2822.14 | Transmitter frequency offset |
| O1P [ppm] | 4.702 | Transmitter frequency offset |
| SFO1 [MHz] | 600.2028221 | Transmitter frequency |
| BF1 [MHz] | 600.2000000 | Basic transmitter frequency |

Reference

| | |
|---|---|
| SI | 32768 |
| SF [MHz] | 600.2000000 |
| OFFSET [ppm] | 13.03256 |
| SR [Hz] | 0 |
| HZpPT [Hz] | 0.305176 |
| SPECTYP | UNDEFINED |

Window function

| | |
|---|---|
| WDW | EM |
| LB [Hz] | 0.30 |
| GB | 0 |
| SSB | 2 |
| TM1 | 0 |
| TM2 | 0 |

Phase correction

| | |
|---|---|
| PHC0 [degrees] | 14.178 |
| PHC1 [degrees] | -18.200 |
| PH_mod | pk |

Baseline correction

| | |
|---|---|
| ABSG | 0 |
| ABSF1 [ppm] | 10.00000 |

Fourier transform

| | |
|---|---|
| TDeff | 0 |
| STSR | 0 |
| STSI | 0 |
| ME_mod | no |
| NCOEF | 0 |
| LPBIN | 0 |
| TDoff | 0 |
| REVERSE | FALSE |
| FCOR | 0.5 |
| PKNL | TRUE |
| FT_mod | fqc |

Integration

| | |
|---|---|
| ABSL | 3 |
| AZFW [ppm] | 0.10000 |
| AZFE [ppm] | 0.10000 |
| ISEN | 128 |
| INTSCL | 1 |
| INTBC | yes |
| ASSFACI | 0 |

# APPENDIX J: Acquisition and processing parameters for 2D $^{1}$H-$^{15}$N HSQC.

ACQUISITION

| | F2 | F1 | Frequency axis |
|---|---|---|---|
| **Experiment** | | | |
| PULPROG | hsqcfpf3gpphwg ... E | | Current pulse program |
| AQ_mod | DQD | | Acquisition mode |
| FnTYPE | traditional(planes) | | nD acquisition mode for 3D etc. |
| FnMODE | | States-TPPI | Acquisition mode for 2D, 3D etc. |
| TD | 4096 | 128 | Size of fid |
| DS | 16 | | Number of dummy scans |
| NS | 8 | | Number of scans |
| TD0 | 1 | | Loop count for 'td0' |
| TDav | 0 | | Average loop counter for nD experiments |
| **Width** | | | |
| SW [ppm] | 24.5015 | 35.0001 | Spectral width |
| SWH [Hz] | 14705.883 | 2128.870 | Spectral width |
| IN_F [µsec] | | 469.7328 | Increment for delay |
| AQ [sec] | 0.1392640 | 0.0300629 | Acquisition time |
| FIDRES [Hz] | 7.180607 | 33.263592 | Fid resolution |
| FW [Hz] | 128000.000 | | Filter width |
| **Receiver** | | | |
| RG | 101 | | Receiver gain |
| DW [µsec] | 34.000 | | Dwell time |
| DWOV [µsec] | 0.025 | | Oversampling dwell time |
| DECIM | 1360 | | Decimation rate of digital filter |
| DSPFIRM | sharp(standard) | | DSP firmware filter |
| DIGTYP | DRX | | Digitizer type |
| DIGMOD | digital | | Digitization mode |
| DR | 32 | | Digitizer resolution |
| DDR | 0 | | Digital digitizer resolution |
| DE [µsec] | 6.50 | | Pre-scan delay |
| NBL | 1 | | Number of blocks (of acquisition memory) |
| HPPRGN | normal | | Preamplifier gain |
| PRGAIN | | | High power preamplifier gain |
| DQDMODE | | | Digital quad detection mode |
| PH_ref [degree] | 0 | | Receiver phase correction |
| OVERFLW | check | | ADC overflow checking |
| **Nucleus 1** | | | |
| NUC1 | 1H Edit... | 15N | Observe nucleus |
| O1 [Hz] | 2822.14 | 6994.04 | Transmitter frequency offset |
| O1P [ppm] | 4.702 | 115.000 | Transmitter frequency offset |
| **Nucleus 2** | | | |
| NUC2 | 13C Edit... | | 2nd nucleus |
| O2 [Hz] | 15242.96 | | Frequency offset of 2nd nucleus |
| O2P [ppm] | 101.000 | | Frequency offset of 2nd nucleus |
| SFO2 [MHz] | 150.9356530 | | Frequency of 2nd nucleus |
| BF2 [MHz] | 150.9204100 | | Basic frequency of 2nd nucleus |
| **Nucleus 3** | | | |
| NUC3 | 15N Edit... | | 3rd nucleus |
| O3 [Hz] | 6994.04 | | Frequency offset of 3rd nucleus |
| O3P [ppm] | 115.000 | | Frequency offset of 3rd nucleus |
| SFO3 [MHz] | 60.8247320 | | Frequency of 3rd nucleus |
| BF3 [MHz] | 60.8177380 | | Basic frequency of 3rd nucleus |

| | F2 | F1 | Frequency axis |
|---|---|---|---|
| **Reference** | | | |
| SI | 2048 | 256 | Size of real spectrum |
| SF [MHz] | 600.2000000 | 60.8177380 | Spectrometer frequency |
| OFFSET [ppm] | 16.95282 | 132.50191 | Low field limit of spectrum |
| SR [Hz] | 0 | 0 | Spectrum reference frequency |
| HZpPT [Hz] | 7.180607 | 8.315815 | Spectral resolution |
| SPECTYP | UNDEFINED | | Type of spectrum e.g. COSY, HMQC, ... |
| **Window function** | | | |
| WDW | QSINE | QSINE | Window functions for trf, xfb,... |
| LB [Hz] | 0.30 | 0.30 | Line broadening for em |
| GB | 0 | 0.1 | Gaussian max. position for gm, 0<GB< |
| SSB | 2 | 2 | Sine bell shift SSB (0,1,2,..) |
| TM1 | 0 | 0.1 | Left limit for tm 0<TM1<1 |
| TM2 | 0 | 0.9 | Right limit for tm 0<TM2<1 |
| **Phase correction** | | | |
| PHC0 [degrees] | -86.400 | 0 | 0th order correction for pk |
| PHC1 [degrees] | 0 | 0 | 1st order correction for pk |
| PH_mod | pk | pk | Phasing modes for trf, xfb, ... |
| **Fourier transform** | | | |
| TDeff | 2048 | 0 | Number of fid data points used by ft |
| STSR | 0 | 0 | First output point of strip transform |
| STSI | 0 | 0 | Total number of output points of strip transform |
| ME_mod | no | no | Linear prediction for ft, xfb, ... |
| NCOEF | 0 | 12 | Number of LP coefficients |
| LPBIN | 0 | 256 | Number of output points for LP |
| TDoff | 0 | 0 | Number of back-predicted points |
| REVERSE | FALSE | FALSE | Reverse spectrum during transform |
| FCOR | 0.5 | 0.5 | Weighting factor for first fid point |
| PKNL | TRUE | | Group delay compensation |
| FT_mod | fqc | no | Fourier transform mode for trf, xtrf* |
| MC2 | | States-TPPI | Acquisition mode (FnMODE) for 2D, 3D, etc. |

APPENDIX K: Temperature coefficients for apoSH and holoSS SOD1 when only considering identically assigned residues across each state

| | ApoSH | | | HoloSS | | |
|---|---|---|---|---|---|---|
| | **rapWT** | **raE100G** | **raV148I** | **hpWT** | **hE100G** | **hV148I** |
| **average** | -4.64 (106) | -4.52 (106) | -4.69 (106) | -2.93 (75) | -3.09 (75) | -2.81 (75) |
| | | | | | | |
| **N** | - | - | - | - | - | - |
| **B1** | -4.26 (6) | -4.26 (6) | -4.41 (6) | -4.54 (3) | -4.54 (3) | -4.48 (3) |
| **loop I** | -5.31 (3) | -5.31 (3) | -4.90 (3) | - | - | - |
| **B2** | -2.24 (9) | -2.24 (9) | -2.28 (9) | -3.25 (5) | -3.41 (5) | -3.00 (5) |
| **loop II** | -5.30 (3) | -5.30 (3) | -5.28 (3) | -5.12 (1) | -5.32 (1) | -4.93 (1) |
| **B3** | -2.49 (7) | -2.49 (7) | -2.37 (7) | -2.54 (5) | -2.91 (5) | -2.33 (5) |
| **loop III** | -3.09 (2) | -3.09 (2) | -4.42 (2) | -4.02 (2) | -4.22 (2) | -3.98 (2) |
| **B4** | -5.52 (6) | -5.52 (6) | -7.01 (6) | -2.72 (6) | -2.86 (6) | -2.57 (6) |
| **loop IV** | -5.74 (16) | -5.74 (16) | -5.54 (16) | -3.13 (16) | -3.27 (16) | -3.00 (16) |
| **B5** | -4.58 (4) | -4.58 (4) | -4.55 (4) | -3.40 (6) | -3.54 (6) | -3.29 (6) |
| **loop V** | -3.80 (4) | -3.80 (4) | -3.59 (4) | -1.85 (2) | -1.88 (2) | -1.84 (2) |
| **B6** | -4.53 (7) | -4.53 (7) | -4.61 (7) | -3.86 (4) | -3.96 (4) | -3.80 (4) |
| **loop VI** | -4.19 (8) | -4.19 (8) | -4.05 (8) | -2.61 (7) | -2.86 (7) | -2.61 (7) |
| **B7** | -2.53 (3) | -2.53 (3) | -3.50 (3) | -1.86 (4) | -2.02 (4) | -1.81 (4) |
| **loop VII** | -6.03 (19) | -6.03 (19) | -5.80 (19) | -2.21 (9) | -2.30 (9) | -2.14 (9) |
| **B8** | -4.38 (7) | -4.38 (7) | -4.59 (7) | -2.98 (4) | -3.14 (4) | -2.62 (4) |
| **C** | -6.80 (2) | -6.80 (2) | -6.54 (2) | -0.70 (1) | -0.92 (1) | -0.38 (1) |

APPENDIX L: Temperature coefficients for apoSH and holoSS SOD1 when only considering identically assigned residues between both states

| | ApoSH | | | HoloSS | | |
|---|---|---|---|---|---|---|
| | **rapWT** | **raE100G** | **raV148I** | **hpWT** | **hE100G** | **hV148I** |
| **average** | -4.52 (51) | -4.46 (51) | -4.72 (51) | -3.11 (51) | -3.24 (51) | -2.97 (51) |
| | | | | | | |
| **N** | - | - | - | - | - | - |
| **B1** | -4.99 (3) | -4.99 (3) | -5.37 (3) | -4.54 (3) | -4.54 (3) | -4.48 (3) |
| **loop I** | - | - | - | - | - | - |
| **B2** | -3.41 (5) | -3.41 (5) | -3.35 (5) | -3.25 (5) | -3.41 (5) | -3.00 (5) |
| **loop II** | -3.81 (1) | -3.81 (1) | -3.84 (1) | -5.12 (1) | -5.32 (1) | -4.93 (1) |
| **B3** | -2.40 (4) | -2.40 (4) | -2.24 (4) | -2.67 (4) | -2.75 (4) | -2.48 (4) |
| **loop III** | -3.09 (2) | -3.09 (2) | -4.42 (2) | -4.02 (2) | -4.22 (2) | -3.98 (2) |
| **B4** | -6.33 (4) | -6.33 (4) | -8.58 (4) | -2.68 (4) | -2.77 (4) | -2.52 (4) |
| **loop IV** | -5.57 (7) | -5.57 (7) | -5.32 (7) | -2.89 (7) | -3.04 (7) | -2.65 (7) |
| **B5** | -4.58 (4) | -4.58 (4) | -4.55 (4) | -4.81 (4) | -4.95 (4) | -4.69 (4) |
| **loop V** | -2.15 (2) | -2.15 (2) | -1.38 (2) | -1.85 (2) | -1.88 (2) | -1.84 (2) |
| **B6** | -3.40 (4) | -3.40 (4) | -3.57 (4) | -3.86 (4) | -3.96 (4) | -3.80 (4) |
| **loop VI** | -3.44 (3) | -3.44 (3) | -3.36 (3) | -4.03 (3) | -4.33 (3) | -4.00 (3) |
| **B7** | -3.15 (2) | -3.15 (2) | -4.57 (2) | -2.24 (2) | -2.42 (2) | -2.44 (2) |
| **loop VII** | -7.53 (6) | -7.53 (6) | -7.67 (6) | -1.51 (6) | -1.62 (6) | -1.40 (6) |
| **B8** | -4.04 (3) | -4.04 (3) | -3.38 (3) | -3.05 (3) | -3.23 (3) | -2.65 (3) |
| **C** | -4.35 (1) | -4.35 (1) | -4.03 (1) | -0.70 (1) | -0.92 (1) | -0.38 (1) |

# APPENDIX M: Representative spectrum of apoSH pWT at 25 °C