UNIVERSITY OF CALIFORNIA, IRVINE

# Characterization of the Dynamics and Thermostability of Y-family translesion DNA polymerase Dbh

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Molecular Biology and Biochemistry

by

Sean Leo Moro

Dissertation Committee:

Dr. Melanie Cocco (committee chair)

Dr. Shiou-Chuan (Sheryl) Tsai

Dr. Ray Luo

2015

# TABLE OF CONTENTS

# LIST OF FIGURES

# ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Melanie Cocco, for all of her guidance, patience, and passion for science, and for her continual support. She has taught me the value of perseverance when experiments do not proceed as planned.

I would to thank my committee members, Professor Ray Luo and Professor Sheryl Tsai, for their kindness and support. I thank Professor Luo especially for the use of his computational resources, without which we could not have obtained the molecular dynamics simulation data for Dbh.

I thank Wesley Botello-Smith of the Luo laboratory for his teaching and guidance on how to perform MD simulations, which expedited the work significantly

I thank all the members of the Cocco laboratory with whom I have worked for their friendship and kindness over these years. Without good company, this would have certainly been a less enjoyable experience.

Portions of chapter 2 were originally published in Biomolecular NMR assignments, Volume 9, Issue 2, pages 441-445, with Melanie J. Cocco. The final publication is available at link.springer.com, doi 10.1007/s12104-015-9626-y. The copyright information for Biomolecular NMR assignments on the Springer website states that "author retains the right to use his/her article for his/her further scientific career by including the final published journal article in other publications such as dissertations and postdoctoral qualifications provided acknowledgement is given to the original source of the publication."

Figures 1.1, 1.2, 1.3, 1.4, and 3.6 have been re-used with express permission from the Royal Society of Chemistry, Elsevier, John Wiley and Sons, The American Association for the Advancement of Science, and Nature Publishing Group, respectively.

I thank Professor Tom Steitz (Yale University) for providing the plasmid containing the Dbh gene and F. W. (Rick) Dahlquist for the gift of E. Coli strain EA1. Egest Pone prepared several of the selectively labeled or unlabeled samples used for the Dbh backbone assignment.

The UC Cancer Research Coordinating Committee-35159 provided funding for the research contained herein.

# Sean Leo Moro, Ph.D.

406 Verano Place
Irvine, CA 92617
(617) 997-7073
sean.l.moro@gmail.com

## OBJECTIVE

Structural biologist with expertise in biomolecular NMR spectroscopy and experience in a number of biophysical techniques and protein modelling/simulations. I am interested in how the dynamics of thermophilic Y-family polymerases are modulated by temperature, and the changes in dynamics that occur upon DNA and nucleotide binding. I am also interested in the structural and dynamic basis of thermostability for Dbh and investigating the mechanism of cold denaturation for Dbh.

## EDUCATION

**University of California, Irvine,** Irvine, CA **2011-2015**

*Ph.D., Molecular Biology & Biochemistry*

Dissertation Title: "Characterization of the dynamics and thermostability of a translesion DNA polymerase"

Dissertation defended on: October 26, 2015

**Brandeis University,** Waltham, MA **2005-2008**

*B.S., Biology,* Cum Laude

**Springfield Technical Community College,** Springfield, MA **2004-2005**

## RESEARCH EXPERIENCE

**University of California, Irvine**, Irvine, CA **2012-2015**

*Graduate Student Researcher*

Dr. Melanie J. Cocco laboraotry

*Graduate Rotation Student* **2011**

Dr. James S. Nowick laboratory

## TEACHING EXPERIENCE

**University of California, Irvine**, Irvine, CA **2012-2015**

*Graduate Teaching Assistant*

**Cambridge Tutoring and Test Prep**, Laguna Beach, CA **2009-2010**

*Tutor*

## PUBLICATIONS AND PRESENTATIONS

**Publication**:

$^{1"}$H, $^{13}$C, and $^{15}$N backbone resonance assignments of the full-length 40 kDa S. acidocaldarius Y-family DNA polymerase, dinB homolog", *Biomolecular NMR Assignments*, Vol. 9, Issue 2, pgs. 441-445

Published **Oct 9, 2015**. DOI: 10.1007/s12104-015-9626-y

**Manuscripts in preparation:**

1. Structural response of thermophilic Y-family polymerase Dbh to temperature change: by hydrogen deuterium exchange NMR and cold denaturation studies.
2. Picosecond-nanosecond dynamics of Y-family polymerase Dbh by nuclear spin relaxation and molecular dynamics simulation

**Presentations:**

- **Research in progress, University of California, Irvine**; Irvine, CA
  - "Backbone Dynamics of a Y-family polymerase" **2015**
  - "Backbone NMR assignment of a Y-family polymerase" **2014**
  - "Recombinant expression and Purification of Nogo-18" **2013**
- **UCI Molecular Biology and Biochemistry Retreat**; Lake Arrowhead, CA
  - "Backbone Dynamics of a Y-family polymerase" **2015**

## PROFESSIONAL EXPERIENCE

**Bausch + Lomb**, Aliso Viejo, CA **2010-2011**

*Global Product Surveillance and Safety Associate*

**Exiqon Diagnostics**, Tustin, CA (location closed Jun 2010) **2009-2010**

*Document Control Specialist*

**Shire Human Genetic Therapies**, Lexington, MA **2008-2009**

*Quality Control Coordinator*

## PROFESSIONAL AFFILIATIONS

**American Association for the Advancement of Science** **2012-2015**

*Graduate Student Member*

## AWARDS

Dean's List, Brandeis University: Fall 2005, Spring 2007, Fall 2007, Spring 2008

Dean's List, Springfield Tech CC: Fall 2004, Spring 2005

CRC Freshman Chemistry Award – Top Chemistry Student, Springfield Tech CC, 2004-2005

## ABSTRACT OF THE DISSERTATION

Characterization of the Dynamics and Thermostability of Y-family translesion DNA polymerase Dbh

By

Sean L. Moro

Doctor of Philosophy in Molecular Biology and Biochemistry

University of California, Irvine, 2015

Professor Melanie J. Cocco, Chair

The *dinB* homolog (Dbh) from the thermophilic archaeon *Sulfolobus acidocaldarius* is a member of the Y-family of translesion DNA polymerases, which are specialized to accurately replicate DNA across from a wide variety of lesions in living cells. Dbh is also a remarkably thermostable polymerase, functioning well at 80°C, the optimum growth temperature of *S. acidocaldarius*. Herein I present the study of the dynamics of apo Dbh at atomic resolution by hydrogen-deuterium exchange (HDX) NMR, NMR spin relaxation, and molecular dynamics (MD) simulations at two temperatures. In order to interpret the NMR data, it was necessary to assign the backbone resonances of Dbh. To this end, I have assigned the $^{15}N$, $^{1}H$, and $^{13}C$ backbone resonance signals at two temperatures (35°C and 50°C) for 86% of the residues of Dbh, which have been published (Moro and Cocco, 2015).

The experiments presented herein demonstrate the remarkable stability of the palm and little finger (LF) domains of Dbh, which remain rigid and well-folded at 50°C. For instance, residues in the palm and LF have protection factors greater than $10^8$ and $10^9$.

MD simulations indicate that the LF domain is free to rotate about the linker region with respect to the polymerase core. The LF domain reorientation is much faster at 50°C and can explain the stark difference in Dbh activity at low versus high temperatures. It is also possible that the reorientation of the LF allows the binding of DNA distorted by various types of lesions.

In addition, I performed biophysical experiments (circular dichroism, differential scanning calorimetry, and HDX-NMR) to investigate the possibility of cold denaturation above zero °C for Dbh. Cold denaturation is a well-established phenomenon, but there are few examples of proteins that denature above zero °C. I obtained conflicting evidence for cold denaturation, with CD and DSC indicating some structural change, but no evidence for structural change by HDX-NMR. No evidence was found for a full unfolding of the polypeptide chain of Dbh. Further investigation is merited in order to fully characterize structural changes in Dbh at low temperature.

**Reference:**

Moro, S.L., and Cocco, M.J. (2015). H, C, and N backbone resonance assignments of the full-length 40 kDa S. acidocaldarius Y-family DNA polymerase, dinB homolog. Biomol NMR Assign.

## Chapter 1: Introduction – Dynamics of a Thermostable Polymerase

Dbh (*dinB* homolog*)* is a 354-amino acid, thermostable, Y-family translesion polymerase from the thermophilic archaeon *Sulfolobus acidocaldarius*. I am interested in two aspects of the structure and dynamics that affects the biological function of Dbh and related enzymes: 1) how the dynamic properties of the enzyme relate to its thermostability and 2) how the structure and dynamics of the enzyme determine its function and biological role in translesion synthesis. This chapter reviews what has been discovered about the structural and dynamic properties of thermophilic enzymes, the biological function of translesion synthesis, and the structure and dynamics of Y-family polymerases during replication of undamaged and damaged DNA templates.

### Thermostable enzymes

*Life at extremes*

Life on earth has evolved to inhabit even the most extreme of environments, from the frozen wastes of the Antarctic ice sheet to the crushing pressure of the Challenger Deep, to brine pools, highly alkaline soda lakes, and acidic solfataras. The specialist organisms that thrive in extreme environments that are hostile to most forms of life have evolved mechanisms to carry out cellular chemistry in these environments using the same basic materials as other organisms. In particular, the cellular substructures and macromolecules of extremophile organisms must remain functional under conditions which would irreversibly degrade and denature their mesophilic counterparts. For adaptations to ionic strength and extremes of pH, extremophiles possess efficient mechanisms of maintaining homeostasis, such as pumping out protons in highly acidic

environments (Baker-Austin and Dopson, 2007), or concentrating stabilizing molecules such as glycerol in the cytoplasm to deal with high salt concentrations (Plemenitas et al., 2014). In the case of low pH, these adaptations maintain a cellular environment more similar to mesophilic organisms, although with a much increased membrane pH gradient (Baker-Austin and Dopson, 2007). However, to adapt to extremes of temperature or pressure, cellular substructures themselves, and especially proteins, have to be changed to achieve stability and activity. Since the native, folded states of proteins at mesophilic conditions are often only marginally stable, with the ΔG of the folded state only equivalent to a few hydrogen bonds, imparting thermostability to proteins is a particular challenge for these organisms. Adaptive mutations must be made to stabilize the protein fold at high temperature while preserving function and activity.

The question of how thermophilic proteins maintain their stability and activity at high temperatures has been a hotly debated question in the structural biology and biotechnology communities for many years. An active pursuit of this matter in the biotech sector has been improving the thermal stability of proteins for multitudinous applications such as the decomposition of polysaccharides for biofuel production (Bleicher et al., 2011) and as catalysts for the generation of building blocks for pharmaceutical compounds (Wieteska et al., 2015). To more efficiently engineer enzymes for these applications, many researchers have sought to understand what mutations impart stability to thermophilic enzymes in comparison to their mesophilic homologs. Unfortunately, although a number of stabilizing factors have been identified in thermophilic enzymes, these trends do not always hold for all examples and classes

of enzymes. Consequently, further defining the structural basis of protein thermostability remains an active area of research.

**Factors Contributing to Thermostability**

Some factors which have been proposed to impart thermostability to proteins include an increase in rigidity, increases in the relative composition of particular amino acids (especially charged residues) (Fukuchi and Nishikawa, 2001; Gromiha and Suresh, 2008), additional disulfide bonds (Wieteska et al., 2015), and increases of networked ion pairs (Yip et al., 1995), increases in main-chain hydrogen bonds (Sadeghi et al., 2006), and increases in hydrophobic packing due to increases in branched-chain amino acids (Gromiha et al., 1999). However, there are often exceptions to these factors, especially in the case of rigidity. The cost of increased thermostability resulting from increased rigidity is a decrease in activity at lower temperatures (Mamonova et al., 2013), although some thermophilic enzymes have been found to have similar flexibility and activity at room temperature as their mesophilic counterparts (Kamal et al., 2012). The only stabilizing property which seems to be true for most thermophilic proteins is a relative increase in salt bridges (Szilagyi and Zavodszky, 2000). It has been postulated that at higher temperatures, the desolvation cost of forming a salt bridge is far lower than at mesophilic temperatures, increasing the energy of interaction and contributing to the stability of the enzyme (Elcock, 1998). Even so, not all salt bridges provide the same degree of stabilization; in the case of aqualysin I, a thermophilic protease, Jonsdottir and co-workers found that only certain key salt bridges made a significant contribution to the thermostability of the enzyme (Jonsdottir et al., 2014).

Efforts to further improve the thermostability of enzymes have been informative regarding the properties that impart thermostability. The thermostability of *Thermococcus sp.* 1519 DNA ligase was increased by mutating two alanine residues to lysine, one serine residue to isoleucine, and one glycine to aspartate, which provided increased electrostatic interactions on the surface of the protein (Pezeshgi Modarres et al., 2015). A study by Wieteska and co-workers mutated two separate residues to cysteine to increase the stability of L-threonine aldolase from *Thermotoga maritima*; they also attempted to insert additional salt bridges to no effect (Wieteska et al., 2015). In the case of *Photinus pyralis* luciferase, mutating one residue to proline in one flexible region increased thermostability, while a proline substitution in another flexible region decreased thermostability (Yu et al., 2015). In another study, mutation of certain proline residues in one portion of *E. coli* phytase, yet also glycine residues in other regions increased its thermostability (Wu et al., 2014). While increasing rigidity would appear to be a surefire way to engineer thermostability in an enzyme, in certain cases increased flexibility is more entropy stabilizing (Karshikoff et al., 2015). It seems that there is no one simple solution to improving thermostability, as the structure and fold dictate what changes need to be made to each particular enzyme. Therefore, any attempt to confer improved thermostability to a protein should be preceded by a thorough biophysical and structural characterization in order to determine the best experimental approach.

**Dynamics, Activity, and Thermostability**

It has become increasingly clear that the various dynamic processes in proteins are intimately tied to their mechanism of function. While high-resolution crystal structures

provide detailed structural information on static conformations, they do not reveal dynamic properties, such as the timescales and degree of movement of domains. Sub-states important for the function of the protein that are not captured by crystallography can be revealed by methods that probe dynamics. In the case of thermophilic enzymes, the dynamic properties have an intimate relationship with the thermostability and the temperature at which the enzyme is most active. Although increased thermostability is often generated through increased rigidity at lower temperatures, in many cases the picture is more complex.

Many studies have been performed which investigate the dynamics of thermostable enzymes, primarily through computational methodologies, Mamonova and co-workers used an empirical method, [FoldUnfold (Galzitskaya et al., 2006)] and a molecular modelling method [MD/FIRST(Mamonova et al., 2005)], to evaluate the consequences of flexibility in thermostable proteins from many classifications. They concluded that the ion pairs in thermophilic proteins occur more often in networks, so that unbinding of one pair does not significantly affect thermostability (Mamonova et al., 2013). A molecular dynamics study evaluating specific ion pairs thermophilic adenylate kinase concluded that one particular ion pair was stable over the course of the simulation and conferred significant stability to the enzyme, while the other pair drifted apart and did not confer stability (Gromiha and Suresh, 2008). Another MD study demonstrated that the stability of CYP119, an extremely stable cytochrome P450 from *Sulfolobus acidocaldarius*, was conferred in large part by the hydrophobic packing of two key sidechains (Tyr and Leu), stabilizing a critical loop (Meharenna and Poulos, 2010). Joo and co-workers observed highly unstable regions from the MD simulation of *Bacillus circulans*, and demonstrated

that mutating polar residues in these regions to large hydrophobic side chains stabilized the enzyme without sacrificing activity (Joo et al., 2011). A network analysis and MD study comparing wild-type *Bacillus subtilis* lipase A to more thermostable mutant enzyme concluded that marginal local increases in stability have a large effect on overall stability, while leaving the overall tertiary structure unchanged (Srivastava and Sinha, 2014). Thus, improved hydrophobic packing and cooperative ion pair interactions in key locations, without necessarily altering tertiary structure, appears to confer thermostability to proteins

Experimental comparative dynamics studies on mesophilic and thermophilic homologues can provide rich information on dynamic changes that lead to thermostability when the tertiary structure is very similar. One set of thermophilic-mesophilic enzyme homologues that have been extensively studied for factors contributing to thermostability are those of ribonuclease H. The Palmer research group investigated the dynamics of RNAseH from *Thermus thermophilus* by $^{15}N$ NMR relaxation dynamics, and compared its dynamics to its mesophilic counterpart from *Escherichia coli* (Butterwick and Palmer, 2006; Stafford et al., 2013; Stafford et al., 2015). One interesting result of these studies is that the thermophilic RNAse occupies the closed conformation, which is not competent for substrate binding, while the mesophilic enzymes occupy the open, binding-competent conformation far more often (Stafford et al., 2013). A network of residues in the handle region (residues 81-101), which is important for substrate binding, was shown to have subtle changes in hydrophobic packing and dynamics that affect the population of the open, binding-competent state, thereby conferring activity to the mesophile and stability at the

expense of activity to the thermophile (Stafford et al., 2013). An extra glycine residue in the thermophile in this region was shown to be especially important in determining the conformation of this region (Butterwick and Palmer, 2006). In a subsequent study, Stafford and co-workers showed that even though the handle showed slightly increased flexibility, the handle region spent more time in a non-competent conformation for binding, while other regions were less flexible (Stafford et al., 2015).

In sum, these results suggest that the relationship of dynamics, activity, and thermostability is complex: certain regions may increase or decrease in flexibility, but cumulatively the dynamic processes increase stability and decrease activity. To understand how the conformational dynamics of thermophilic enzymes is changed at different temperatures, I am studying the conformational dynamics of Dbh polymerase from the thermophilic archaeon *Sulfolobus acidocaldarius* (optimal growth conditions, 80°C and pH 2), at two different temperatures (35°C and 50°C) by a combination of NMR relaxation dynamics and molecular dynamics simulations. Like many thermophilic enzymes, Dbh is increasingly active at higher temperatures, *in vitro* its activity is approximately 40-fold higher at 65°C than at 22°C (Potapova et al., 2002). Therefore, I have been able to observe how the dynamics of the enzyme change over a temperature range where increasing activity has been observed.

**Cold Denaturation**

It is logical and expected that proteins denature at high temperatures, since sufficient energy is available to overcome the solvation penalty for hydrophobic groups in solution, and conformational entropy is increased at higher temperature. However,

proteins can also denature at low temperatures. The thermodynamic stability of proteins is a parabolic function of temperature, possessing a point of maximum stability. This implies that proteins have two denaturation temperatures, above and below which the denatured is thermodynamically favored over the native state of the protein (Graziano, 2014). For most proteins, the cold denaturation temperature is below freezing; hence, cold denaturation has not been as extensively studied, as it requires the use of denaturants, super-cooled solutions, or high pressure to observe the cold denaturation process. Many studies have demonstrated the existence of cold denaturation using these methods such as scanning calorimetry, CD spectroscopy, and NMR spectroscopy (Azuaga et al., 1992; Griko and Privalov, 1992; Griko and Kutyshenko, 1994; Kumar et al., 2006; Pometun et al., 2006; Whitten et al., 2006; Vajpai et al., 2013). In some cases, even in the absence of denaturants cold denaturation occurs above freezing (Pastore et al., 2007; Buchner et al., 2012). The phenomenon has allowed a more nuanced understanding of thermodynamics of water and how it affects protein stability.

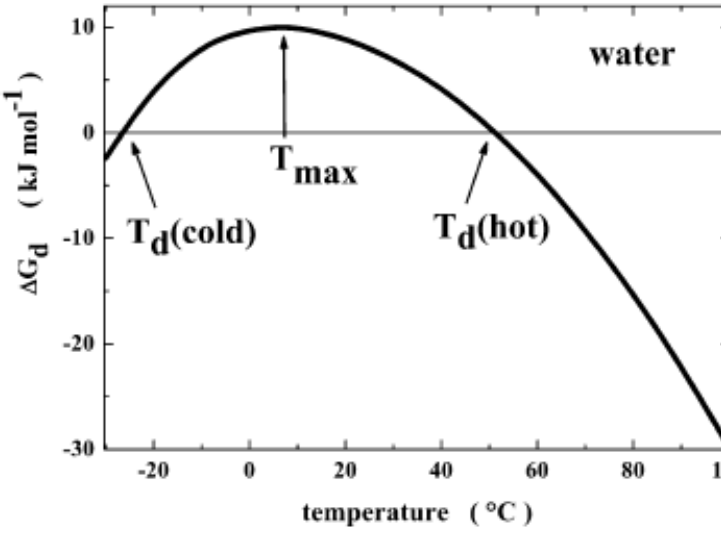



**Figure 1.1**: Stability curve for a typical protein. Maximum stability is at approximately 10°C, and the protein possesses both cold and heat denaturation temperatures. Reproduced from *Phys. Chem. Chem. Phys.,* Graziano, 2014, with permission of the PCCP Owner Societies dx.doi.org/10.1039/c4cp02729a

The main mechanism of cold denaturation is a change in the properties of water at low temperature (Davidovic et al., 2009). The entropic penalty for solvation of hydrophobic groups of proteins is reduced at lower temperatures – at a certain point the interaction of water with hydrophobic groups becomes favorable and the protein can denature (Graziano, 2012). This leads to penetration of water into the protein core, and formation of non-native tertiary contacts and disruption of intramolecular hydrogen bonding (Yang et al., 2014). Cold denaturation proceeds with a decrease in both the entropy and enthalpy of the protein system; subsequently, the hydrophobic effect is weakened (Graziano, 2014). For many proteins, cold denaturation is a slow process that proceeds through an intermediate state, then to the fully denatured state (Privalov, 1990).

For thermostable proteins, their increased thermostability can be explained as a shift in the protein stability curve (in Figure 1.1, the curve shifts to the right, resulting in a higher temperature of maximum stability and a higher heat denaturation temperature (Graziano, 2014). Whether this entails the protein stability curve is deeper, shifted to the right, or both, is a matter of active research. Characterizing the process of cold denaturation can also lead to a greater understanding of protein folding and stability. To this end, I have investigated whether Dbh cold denatures at temperatures above zero Celsius. Based on HSQC NMR spectra, it appeared that Dbh began slowly denaturing at temperatures below 20°C. Four experiments – circular dichroism spectroscopy, differential scanning calorimetry, dynamic light scattering, and hydrogen deuterium exchange NMR – were performed. However, there was no conclusive evidence that Dbh secondary structure unfolds to a random coil conformation approaching zero Celsius, although there was a gradual, reversible conformational transition by DSC and

by HSQC NMR spectra. Therefore, it appears that non-native tertiary structure gradually forms in Dbh at temperatures approaching zero Celsius.

**Translesion DNA polymerases**

In 1956, Arthur Kornberg discovered the first DNA-copying enzyme from *E. coli*, which was named DNA polymerase I (Bessman et al., 1956). Since then, many DNA polymerases from all forms of life have been discovered, classified by homology and function into six families: A, B, C, D, which are replicative polymerases, and X and Y, which are the translesion polymerases (Joyce and Benkovic, 2004).

In 1999, the *RAD30* gene was recognized as a translesion DNA polymerase, pol η, which could accurately replicate cyclobutane pyrimidine dimers (Johnson et al., 1999a; Johnson et al., 1999b). Many other translesion polymerases were discovered shortly thereafter, such as pol IV (Wagner et al., 1999) pol κ (Ogi et al., 1999; Ohashi et al., 2000; Gerlach et al., 2001), pol ι (McDonald et al., 1999; Tissier et al., 2000), and Rev1 (Nelson et al., 1996a; Lin et al., 1999; Wiltrout and Walker, 2011), and were categorized as the Y-family of DNA polymerases (Ohmori et al., 2001). Y-family polymerases carry out translesion synthesis (TLS), accurately bypassing a variety of damaged DNA templates that stall higher-fidelity replicative polymerases (Yang and Woodgate, 2007). However, Y-family polymerases are quite error-prone on undamaged templates (Yang and Woodgate, 2007).

The error-prone replication of undamaged DNA templates by Y-family polymerases is a direct consequence of their more open active sites. The expansion of the active site limits the ability of the polymerase to perform a “fidelity check” on the incoming

nucleotide. Nevertheless, the increased size of the active site allows the accommodation of bulky DNA adducts or other lesions which cause distortion of the DNA double helix (Sale et al. 2012).

In eukaryotes, TLS activity is not confined to the Y-family polymerases – several other enzymes are directly capable of performing TLS. The X-family polymerases pol β, pol λ, and pol μ primarily function in the base excision repair (BER) and non-homologous end-joining (NHEJ) pathways, yet all three are also capable of translesion synthesis across a number of lesions, such as 8-oxoguanine (pol β, pol μ) 1-$N^6$-ethenoadenine (pol μ), and 2-hydroxyadenine (pol λ) (Yamtich and Sweasy, 2010). Pol ζ, a B-family polymerase consisting of the *Rev3* polymerase and *Rev7* regulatory domain, was shown to bypass cyclobutane pyrimidine dimers (Nelson et al., 1996b). Pol ζ can also replicate across a number of other lesions and extension from lesions, and interacts with Rev1 and PCNA to regulate TLS (Makarova and Burgers, 2015). TLS is used by the organism to alleviate replication fork stalling and allow completion of replication, which avoids the consequences of improper segregation of chromosomes, such as cellular senescence and apoptosis (Sale et al. 2012). A recently discovered protein, human Primase-Polymerase (PrimPol), a member of the archaeo-eukaryotic primase superfamily, possesses both primase and translesion synthesis activities (Garcia-Gomez et al., 2013). The primase activity of PrimPol is able to restart a stalled replication fork after lesion bypass (Mouron et al., 2013). PrimPol can also replicate across 6-4 thymine-thymine photoproducts and 8-oxoguanine, and extend from cyclobutane pyrimidine dimers (Bianchi et al., 2013; Zafar et al., 2014)

## Biological Function

*Classification of Y-family polymerases*

As noted previously, Y-family polymerases allow cells to bypass a wide variety of cognate lesions in an error-free manner, while replicating undamaged DNA in an error-

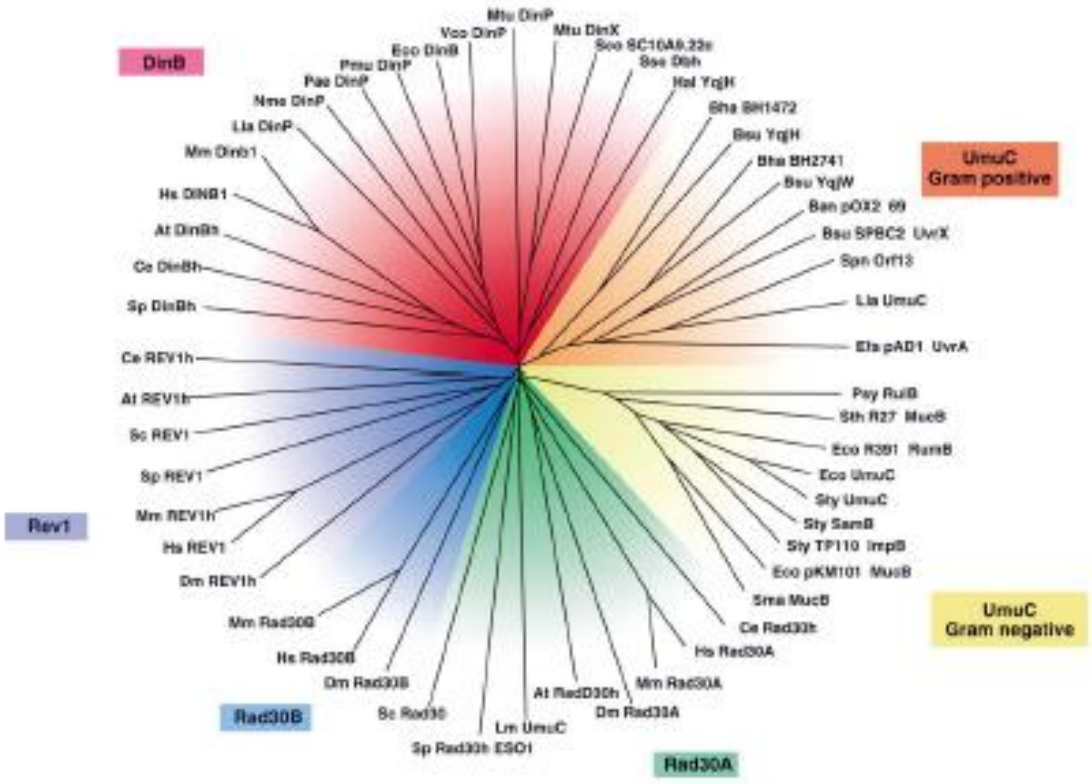



**Figure 1.2:** Phylogenetic tree of Y-family DNA polymerases, Dbh is in the *dinB* subfamily, and it is listed as "SsoDbh" in this diagram. Reprinted from *Mol. Cell.*, vol. 8, issue 1, Ohmori et. al, "The Y-family of DNA polymerases", pgs. 7-8, © 2001, with permission from Elsevier.

prone fashion. Prokaryotes usually have one or two Y-family polymerases, whereas eukaryotes have four. In general, the prokaryotic polymerases handle a greater variety of lesions, while the eukaryotic polymerases have more specialized roles. The Y-family polymerases across all organisms are classified for their homology and function (Figure 1.2). A phylogenetic analysis performed by Ohmori et al. groups the Y-family

polymerase into five subfamilies: the *dinB* subfamily (e.g., *E. coli* pol IV, Dbh, and Dpo4) found in all domains of life, the bacterial *UmuC* subfamily(e.g. *E. coli* pol V), containing separate branches for gram-negative and gram-positive bacteria, the eukaryotic *Rad30A* subfamily (pol η), the *Rad30B* subfamily (pol ι) found only in higher eukaryotes, and finally the eukaryotic Rev1 subfamily (Ohmori et al., 2001).

*Translesion Synthesis by Y-family polymerases*

Translesion synthesis is carried out with varying efficiency and fidelity by each subfamily of Y-family polymerases, and some variation within different enzymes of each subfamily. Some polymerases are specialized to handle specific types of in an error-free manner. Pol κ is specialized to bypass bulky $N^2$-guanine adducts generated by polyaromatic hydrocarbons (PAH) such as benzo[a]pyrene (Zhang et al., 2000a; Ogi et al., 2002; Zhang et al., 2002). Without pol κ, cellular bypass of these lesions is more error-prone (Avkin et al., 2004). *E. coli* pol IV swiftly and accurately bypasses $N^2$-benzo[a]pyrene diol epoxide (BPDE) adducts at stalled replication forks (Ikeda et al., 2014). Pol IV is also able to replicate DNA lesions produced by reactive oxygen species, such as thymine glycol, 8-oxoguanine, 2-oxoguanine, and 5-formyluracil; however, pol IV preferred mutagenic bypass of the 2-oxoadenine lesion with dCTP (Hori et al., 2010).

Pol η is best known for its ability to accurately bypass cyclobutane pyrimidine dimers generated by ultraviolet radiation, both *in vitro* (McDonald et al., 1999) and *in vivo* (Yoon et al., 2009). Loss of pol η leads to dramatically increased susceptibility to UV light in cells and mice (Lin et al., 2006), and is a cause of the disease xeroderma pigmentosum

variant (XPV), which is characterized by keratoses and high rates of skin cancer (McDonald et al., 1999). In XPV cells, pol ι takes over lesion bypass with the consequence of much higher rates of mutation (Lin et al., 2006). Nevertheless, even mutagenic lesion bypass appears to be favorable to the alternative – cells lacking both pol η and pol ι are even more sensitive to UV radiation, and mice develop tumors even more frequently (Dumstorf et al., 2006; Ohkumo et al., 2006).

The major role of Rev1, in contrast to the other Y-family polymerases, is not to directly catalyze a lesion bypass reaction; although in certain contexts its dCMP transferase activity does a have a role in lesion bypass (Wiltrout and Walker, 2011). Rev1 has key importance in eukaryotes as a regulatory scaffolding protein for TLS (Kosarek et al., 2008). It also has a key role in forming a complex with pol ζ in extension past many varieties of lesions, such as interstrand crosslinks (Budzowska et al., 2015).

Dpo4 and Dbh are the only TLS polymerases in their respective species of *Sulfolobus*, and share 54% overall sequence identity, yet the two differ significantly in their bypass properties (Boudsocq et al., 2004). Dpo4 has bypass properties more similar to pol η, such as the ability to bypass cyclobutane pyrimidine dimers and abasic sites, albeit less efficiently than pol η (Boudsocq et al., 2004). Dpo4 is able to efficiently bypass the 8-oxoguanine lesion, which occurs with great frequency at the optimal growth temperature of *Sulfolobus* (Rechkoblit et al., 2006). Dpo4 also can accurately bypass the N2-AAF-dG (2-acetylaminofluorene) lesion accurately (Boudsocq et al., 2001). Dpo4 can bypass at least one stereoisomer, 10*S* (+)-*trans*-*anti*-benzo[a]pyrene-$N^2$-dG, yet inserts all four dNTPs equally well (Perlow-Poehnelt et al., 2004). In contrast, the bypass properties of

Dbh resemble those of pol κ; Dbh is more likely to bypass N2-BPDE lesions and more accurately than Dpo4 (Sholder and Loechler, 2015). *In vivo,* at the optimal growth temperature of 80°C for *S. acidocaldarius*, Dbh is active in preventing transversion mutations resulting from 8-oxoguanine lesions (Sakofsky et al., 2012). Interestingly, the lesion bypass properties of the two enzymes can be switched by switching the little finger domain, or the core-LF linker region (Boudsocq et al., 2004; Wilson and Pata, 2008).The significant differences in lesion bypass activity between these closely related enzymes demonstrates that the degree and fidelity of lesion bypass is intimately tied to the structure and dynamics of the particular polymerase, and in particular the conformational dynamics of the LF domain.

*Regulation of Translesion Synthesis – Processivity Clamp Interaction*

Because of their intrinsic error rates on undamaged DNA and differential ability to bypass lesions in an error-free fashion, Y-family polymerases are usually highly regulated in all organisms. Failure to properly regulate TLS can increase mutational burdens, and decrease the ability of the organism to deal with genotoxic stress. In prokaryotes such as *E. coli*, one way to control TLS is through the direct interaction with the β-clamp, with the main replicative polymerase (pol III) and Y-family polymerases (pol IV and pol V) both able to bind directly to the clamp. Pol IV will be able to access the replication fork only when the SOS DNA damage response occurs, or when replication fork stalling allows pol IV to switch places with pol III at the primer-template junction (Wagner et al., 1999; Furukohri et al., 2008; Ikeda et al., 2014; Kath et al., 2014).

Regulation of Y-family polymerases in eukaryotes also occurs through interactions with the processivity clamp, proliferating cell nuclear antigen (PCNA), with additional scaffolding provided by Rev1. Each eukaryotic Y-family polymerase can interact directly with PCNA. Pol ι, pol η, and pol κ interact through their PCNA-interacting-protein boxes (PIPs) (Haracska et al., 2005; Ogi et al., 2005; Acharya et al., 2008). PIP-boxes have a consensus sequence of Qxxhxxaa, where h is a hydrophobic residue, a is an aromatic residue, and x is any residue (Hishiki et al., 2009). These polymerases also contain Rev1-interaction regions (RIRs) that allow them to bind to the C-terminal domain of Rev1 (D'Souza and Walker, 2006), which itself is able to bind PCNA through its BRCT domain (Pustovalova et al., 2013). The C-terminal domain of Rev1 also forms a complex with pol ζ that is critical for the proper functioning of pol ζ in TLS (Acharya et al., 2005). In this way, Rev1 can mediate interactions of other TLS polymerases with PCNA and with each other.

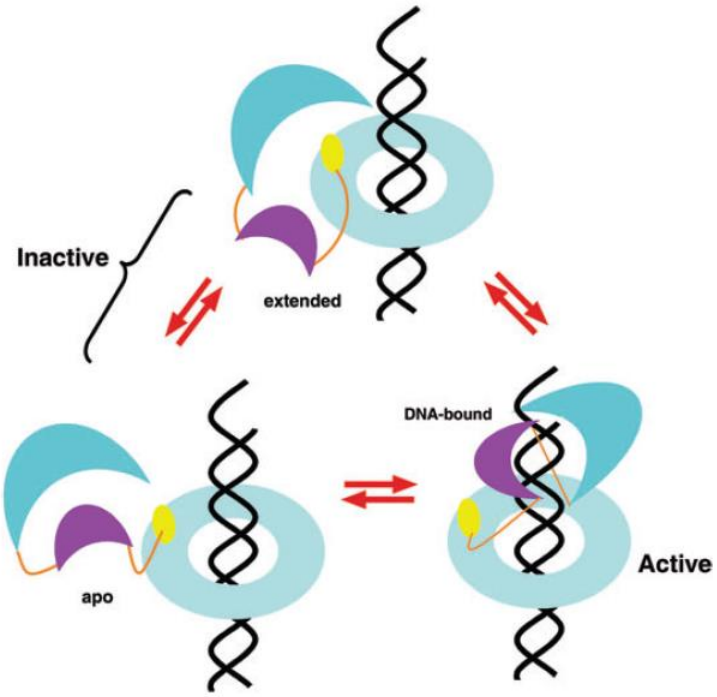



**Figure 1.3** Model of Dpo4 interaction with PCNA and dsDNA. Linker flexibility allows extension of the LF, resulting in Dpo4 being in an inactive state. Image reprinted from Xing et al., *Molecular Microbiology*, © 2009, John Wiley & Sons.

Although the interaction of Dbh with the clamp protein (archaeal PCNA) from *S. acidocaldarius* has not been studied, it is highly probable that translesion synthesis

could be regulated through interactions with the clamp. However, structural studies have been performed on Dpo4 and *S. solfataricus* PCNA. A co-crystal structure of Dpo4 with a heterodimer of *S. solfataricus* PCNA-1 and -2 demonstrates the Dpo4 possesses a PIP-box motif (AIG**L**DK**FF**) in the C-terminal tail that binds to a hydrophobic pocket on the surface of PCNA, structuring the tail into a $3_{10}$ helix; supplemental contacts are also formed between PCNA and loops in the palm, thumb, and finger domains (Xing et al., 2009). Interestingly, the LF was found in an extended conformation and rotated xxx degrees with respect to the polymerase core from the apo structure (1K1S,) and DNA-bound structures. This observation suggests the conformational flexibility of the LF domain via the linker region assists in regulation of Y-family polymerases when bound to the processivity clamp, maintaining the enzyme in an inactive state. Dbh also possesses the PIP-box motif in its C-terminal tail (KTN**L**SD**FF**), so it is quite likely that Dbh binds to *S. acidocaldarius* PCNA in a similar manner. It would be interesting to observe how the dynamics of Dbh are altered when bound to the clamp.

**Y-family polymerases and disease**

Y-family polymerases are highly regulated to ensure that they are deployed only in the appropriate context; using the wrong polymerase for a specific lesion or undamaged template can generate high levels of mutations. Loss of function of a translesion polymerase or any of the many associated regulatory mechanisms can lead to defects in handling bulky lesions and interstrand crosslinks. The resultant aberrant functionality of TLS increases organismal susceptibility to DNA damage from environmental mutagens and endogenous processes, leading to cell death or increased carcinogenesis.

Increased levels of expression of translesion polymerases in cells can drive tumorigenesis through accumulation of mutations. Elevated expression of pol ι in breast cancer cells increases mutation frequency during replication, with a concomitant reduction in mutation frequency upon immunodepletion of pol ι (Yang et al., 2004). Overexpression of pol ι also drives mutagenesis in bladder cancer (Yuan et al., 2013). Pol κ is overexpressed in various types of non-small cell lung cancer (J et al., 2001). Pol η, pol κ, and pol ι overexpression have been found in esophageal carcinomas (Zhou et al., 2012). TLS polymerases also participate in aberrant DNA re-replication, a process that results in significant genomic instability (Sekimoto et al., 2015).

Cancer cells can exploit Y-family polymerases to tolerate DNA damage caused by chemotherapeutics. One prominent example is the bypass of adducts generated by cisplatin and oxalipatin by pol η, which can insert two cytosines across from the G-Pt-G lesion (Vaisman et al., 2000; Alt et al., 2007). The active site of pol η accommodates the helix-distorting lesion well through interactions with key residues (Ummat et al., 2012). Pol ζ complements pol η in the bypass of G-Pt-G adducts by extending the primer from the lesion site (Lee et al., 2014). Pol η deficient cells are far more susceptible to cisplatin-induced DNA damage (Albertella et al., 2005). In non-small cell lung cancer, patient survival during a chemotherapeutic regimen negatively correlated with level of mRNA expression of pol η (Ceppi et al., 2009). However, phenanthriplatin, a next generation platin drug, is not efficiently bypassed by pol η and is toxic to both pol η + and pol η – cells (Gregory et al., 2014).

DNA interstrand cross-links (ICLs), which strongly block replication are also generated with high frequency by many DNA damaging agents such as cisplation, cyclophosphamide, mitomycin C, and carmustine (Sharma and Canman, 2012). Defective recruitment of Rev1-pol ζ complexes to the site of ICLs prevent proper repair and bypass of ICLs (Budzowska et al., 2015). Depletion of the catalytic subunit of pol ζ (*REV3*) in multiple human cancer cell lines increases their sensitivity to chemotherapeutics (Doles et al., 2010; Knobel et al., 2011). Depletion of Rev1 in human ovarian carcinoma cells increases the cytotoxicity of cisplatin and reduces the acquisition of resistance (Okuda et al. 2005). In aggregate, these results show the critical role of the Rev1/pol ζ in tolerance of chemotherapeutics which generate ICLs. Pol η overexpression is also implicated in tolerance of nitrogen mustard crosslinking agents in multiple cancer cells *in vivo* and *in vitro* (Tomicic et al., 2014).

Furthermore, loss of translesion synthesis leads to replication fork stalling and collapse, which can generate chromosomal abnormalities that can contribute to a cancerous phenotype (Lange et al., 2011). The overlap in function between TLS polymerases (both Y- and X-family polymerases) complicates the role of the loss of function of these enzymes in tumorigenesis. Often, one enzyme can fulfill the function of another TLS polymerases in bypassing a certain lesion, which may or may not be mutagenic. In certain cases, the consequences are clear. When pol η is not available for error-free bypass of CPD lesions, pol ι can bypass the lesion, but at the cost of much increased mutagenesis (Wang et al., 2007b).

Nevertheless, there is at least one instance in which the error-prone properties of TLS polymerases are advantageous. Extremely error-prone replication by pol ι, which preferentially mispairs G across from T (Tissier et al., 2000; Zhang et al., 2000b), and often generates other mispairs (Vaisman et al., 2001), is leveraged to diversify antibody genes in immature mammalian B-cells in a process known as somatic hypermutation (Faili et al., 2002). Pol η is also involved in generating A/T transversions in somatic hypermutation (Delbos et al., 2007). Combined with recombination and the generation of mutations across from cytosines by AID, this process generates extremely diverse sequences in the variable regions of antibodies.

**Y-family polymerase structure, function, and dynamics**

Dbh is a Y-family, TLS DNA polymerase from the thermophilic archaeon *Sulfolobus acidocaldarius* (Boudsocq et al., 2004), and shares 54% sequence identity to Dpo4 from the related archaeon *Sulfolobus solfataricus*. Members of the Y-family share canonical polymerase architecture consisting of a catalytic core composed of palm, fingers, and thumb domains; and have an additional unique C-terminal domain termed the “wrist”, “polymerase-associated domain” (PAD), or “little finger” (LF) domain. Notably, the LF domain is tethered to the catalytic core by a flexible linker and has been found to occupy a variety of conformations (Pata, 2010).

The error-prone replication of undamaged DNA templates by Y-family polymerases is a direct consequence of their more open active sites. The expansion of the active site limits the ability of the polymerase to perform a “fidelity check” on the incoming numb Nevertheless, the increased size of active site allows the accommodation of bulky DNA

adducts or other lesions which cause distortion of the DNA double-helix (Sale et al. 2012).

All DNA polymerases share a two-metal mechanism (Figure 1.4) by which they catalyze the addition of one nucleotide to the free 3'-OH of the primer strand of duplex DNA (Joyce and Benkovic, 2004). Typically, two to three carboxylate side chains, and in some cases a backbone carbonyl, are used to chelate two divalent metal cations (mostly magnesium) in the active site of the polymerase, located in the palm domain (Steitz, 1999). High-fidelity polymerases involved in replication form tighter contacts with the DNA primer/template and the incoming nucleotide, thereby discriminating between the incorrect and correct nucleotide (Joyce and Benkovic, 2004). Upon nucleotide binding, the fingers domain closes over the active site, performing a fidelity check on the nascent base pair. This conformational transition was originally thought to be rate limiting, but was shown to occur much faster than the rate of catalysis (Rothwell et al., 2005). In contrast, the fingers domains of Y-family polymerases do not close upon nucleotide binding (Pata, 2010). The rate-limiting step of all DNA polymerases is now believed to be a subtle repositioning of residues surrounding the active site, which properly align the 3'OH of the primer terminus for nucleophilic attack on the 5' α-phosphate of the incoming nucleotide (Maxwell and Suo, 2014).

The mechanism of Y-family polymerases has been investigated by both ensemble and single-molecule Förster resonance energy transfer (FRET) (Xu et al., 2009a; Maxwell et al., 2012, 2014), molecular dynamics (MD) simulations (Perlow-Poehnelt et al., 2004; Rechkoblit et al., 2006; Perlow-Poehnelt et al., 2007; Xu et al., 2007; Jia et al., 2008; Donny-Clark and Broyde, 2009; Xu et al., 2009b; Qin et al., 2013; Lior-Hoffmann et al.,

2014; Maxwell et al., 2014), and fluorescence based techniques (DeLucia et al., 2007; Wong et al., 2008). The greatest conformational rearrangement occurs on binding duplex DNA, with smaller changes occurring during nucleotide binding and repositioning for catalysis (Maxwell and Suo, 2014).

*Experimental Dynamic Studies*

Experimental dynamic studies on Y-family polymerases have largely used fluorescence-based techniques. DeLucia and co-workers studied the dynamics of nucleotide insertion by Dbh in a stopped-flow experiment using 2-aminopurine as the fluorescence reporter; they found that there are three fast conformational transitions after dNTP binding, followed by a fourth conformational transition that had a rate similar to the rate-limiting step of the reaction (DeLucia et al., 2007). They also demonstrated how Dbh can generate single-base deletions – the templating base can slip to pair with the primer terminus, unstacking the preceding template base (DeLucia et al., 2007). A study on Dpo4 dynamics monitoring tryptophan fluorescence confirmed that large conformational change in the little finger domain that is observed in crystal structures also occurs in solution, with the polymerase core remaining rigid (Wong et al., 2008). Studies monitoring tryptophan fluorescence in Dpo4 (Eoff et al., 2009) and pol κ (Zhao et al., 2014) during nucleotide incorporation suggest the rate-limiting step is a slow conformational relaxation *after* phosphodiester bond formation.

Several FRET-based studies have also been performed on Dpo4. The authors of a FRET-based study on Dpo4 concluded that translocation of the DNA happens upon correct nucleotide binding, followed by movement of the little finger domain away from

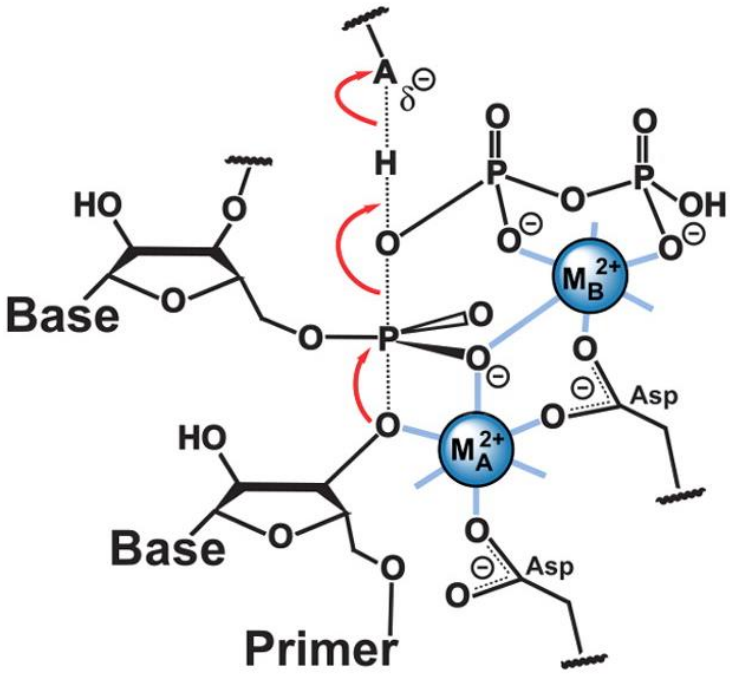



**Figure 1.4:** Two-metal mechanism in the active site of a DNA polymerase. Figure from Shechner et al., "Crystal Structure of the Catalytic Core of an RNA-Polymerase Ribozyme", *Science,* 326(5957), pgs.1271-1275, © 2009, reprinted with permission from AAAS.

the polymerase core (Xu et al., 2009a). They also speculated that the rate-limiting step was the rearrangement of the active site following larger domain movements; however, they could not observe this motion directly with the FRET technique (Xu et al., 2009a). A follow-up FRET study of Dpo4 replicating an 8-oxoguanine lesioned template show additional movements in the thumb and palm domains, the same LF domain movement away from the palm, and slower translocation from the lesion (Maxwell et al., 2012). The differential dynamics of Dpo4 on a lesion-containing template versus an undamaged template highlights how changes in TLS polymerase dynamics could lead to accommodation of lesion-containing DNA substrates. A 2013 single-molecule FRET study of Dpo4 investigating the translocation step showed the translocation motion occurs in the binary complex, but is stabilized in the insertion conformation only by the correct dNTP (Brenlla et al., 2014). One more FRET study concluded that the apo form of Dpo4 possibly varies between DNA-binding competent and non-competent states (Maxwell et al., 2014).

*Molecular Dynamics Studies*

A multitude of molecular dynamics and molecular modelling studies have been performed on Y-family polymerases, utilizing the atomic structures derived from X-ray crystallography. A combined quantum mechanics/molecular mechanics (QM/MM) study of the nucleotide transfer reaction of Dpo4 showed the process involves a water-meditated proton relay mechanism, which is stabilized by the two coordinated $Mg^{2+}$ ions (Wang et al., 2007a). A 2014 study used a two-basin structure-based model combined with replica-exchange MD simulations to investigate the dynamics of DNA binding to Dpo4. The authors postulated that Dpo4 toggles between three successive states during substrate binding (Chu et al., 2014). The greatest conformation change between the three states of Dpo4 involved the positioning of the LF domain, mediated by the flexible linker (Chu et al., 2014). When considered in light of other studies, this result suggests that the positioning of the LF is crucial for proper binding and orientation of the DNA substrate.

For dynamics of lesion bypass, the shape of the active site and the dynamics of the enzyme affect whether the lesion can be bypassed accurately, in an error-prone fashion, or at all. In the of an case $N^2$-BPDE lesion bypass by Dpo4, the spaciousness of the active site allows the lesioned base to reorient from the *anti* conformation (which favors correct base pairing) to the *syn* conformation (which favors mismatched base pairs) in the active site, allowing any dNTP to pair with the lesion and resulting in near-equal incorporation of all four dNTPs (Perlow-Poehnelt et al., 2004). Increasing the temperature from 37°C to 55°C and concordantly the dynamic motion of the polymerase increases the mismatch frequency; because the *syn* conformation of the lesion is now

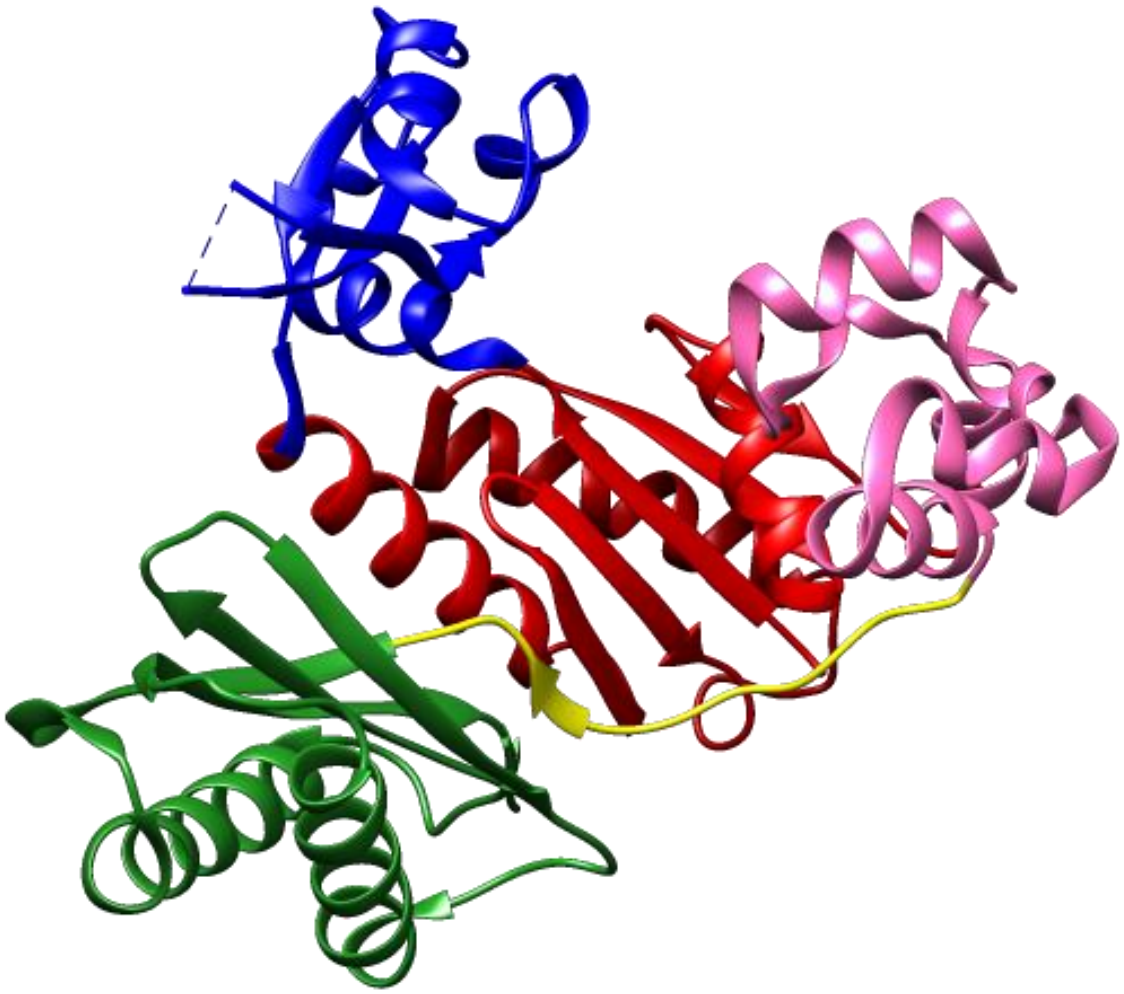


**Figure 1.5:** Crystal structure of apo Dbh (1K1S, (Silvian et al., 2001) demonstrating the characteristic Y-family polymerase fold. The polymerase core consists of the palm (red), thumb (pink), and fingers (blue) domains, and the little finger (LF) or polymerase-associated domain (PAD) (green) is unique to Y-family polymerases. The flexible linker region (yellow) can allow the LF domain to occupy many conformations with respect to the polymerase core, which can significantly influence the polymerase's activity and lesion bypass properties. This figure was generated using UCSF Chimera (Pettersen et al., 2004)

preferred in the active site (Perlow-Poehnelt et al., 2007). Another study on $N^2$-BPDE-dG lesion bypass by Dpo4 using MD suggested that a second conformation of the lesion in the polymerase represents an overlapping alternate catalytically competent active site, which may explain how Dpo4 can accommodate this lesion in multiple orientations with all four nucleotides (Xu et al., 2008). The spacious active-site of Dpo4

during $N^2$-BPDE-dG lesion bypass also explains how it can generate base deletions by allowing the templating lesion to slip into an extrahelical position, positioning the downstream undamaged base for Watson-Crick base-pairing in the active site (Xu et al., 2008). In contrast, the accurate bypass of $N^2$-AAF-dG lesions is achieved because the active site is only properly organized when the adduct is in the *anti* conformation, allowing only Watson-Crick pairing with correct dCTP (Wang and Broyde, 2006).

Pol κ can bypass $N^2$-BPDE-dG in an error-free manner due to its additional N-clasp region, which constrains the lesion in the *anti* conformation needed for Watson-Crick base pairing with dCTP (Jia et al., 2008). During pol κ bypass of an $N^2$-AAF-dG (2-acetylaminofluorene) adduct, the N-clasp still partially enforces the correct orientation of lesion for Watson-Crick pairing with dCTP; however, increased flexibility in the N-clasp allows alternate wobble base-pairing with incorrect dTTP (Lior-Hoffmann et al., 2014). The fact that pol κ can bypass $N^2$-BPDE-dG error-free and yet bypasses $N^2$-AAF-dG inaccurately – *vice versa* for Dpo4 – demonstrates that proper lesion bypass requires the proper shape and dynamics of the active site of the polymerase. Pol ι is also able to accurately bypass the $N^2$-AAF lesion (Zhang et al., 2001), due to the fact that it favors placing the lesion in the major groove side of the template, facilitating correct Watson-Crick pairing (Donny-Clark et al., 2009). Another MD study on pol ι bypass of an $N^2$-BPDE-dA lesion showed that in certain sequences contexts the *anti* conformation with Watson-Crick base-pairing is preferred, while the *syn* conformation with Hoogsteen base-pairing is preferred in others (Donny-Clark and Broyde, 2009). While the flexibility of pol ι allowed lesion rotation, pol ι incorporated correct dTTP by either WC or Hoogsteen base-pariring (Donny-Clark and Broyde, 2009). Clearly, the relationship of

the structure and dynamics to accurate lesion bypass in Y-family polymerases is complex.

In contrast, high-fidelity polymerases should be unable to properly position bulky lesions in their active sites, leading to complete blockage or very inefficient replication. A MD study on high-fidelity Bacillus fragment on the $N^2$-BPDE-dG lesion shows it can only accommodate the lesion in the *syn* conformation, allowing insertion across from it; however, the post-insertion complex is highly distorted, precluding any extension from the lesion (Xu et al., 2007). These MD studies make it clear that the position and orientation of a lesioned base in the active site of the enzyme can largely determine whether or not bypass occurs, either in an error-free or mutagenic manner.

The experimental techniques reviewed above have been able to show domain movements during the catalytic cycle during replication of undamaged and lesioned DNA templates, yet lack atomic resolution. Molecular dynamics simulations of Y-family polymerases, while very informative, have not been directly compared to an atomic-level resolution dynamics experiment. Herein, I present the study the dynamics on the *apo* form of Dbh polymerase by $^{15}$N NMR relaxation spectroscopy and hydrogen-deuterium exchange spectroscopy, providing residue-specific information through the dynamics of the amide bond. In addition, I have investigated the dynamics of Dbh through MD simulations at two temperatures, and I have compared the wild-type polymerase to a variant ($Dbh_{RKS(243\text{-}245)}$) that has similar bypass properties to Dpo4.

No residue-specific experimental dynamic information has been reported for a DNA polymerase; this work represents the first report of NMR relaxation dynamics of a DNA

polymerase. The hydrogen-deuterium exchange demonstrates the thermostability of Dbh is partially due to extremely stable hydrogen bonds in the palm and little fingers domains. Many residues in the core of these domains have no observable decay even after 2 days at 50°C, with protection factors of $10^8$ to $10^9$ and perhaps greater. These highly protected residues are rich in branched-chain amino acids, indicating that increased hydrophobic packing may be responsible for the thermostability of Dbh. In addition, the molecular dynamics simulations indicate the protein is highly rigid at 35°C, but at 50°C the LF domain can reorient freely around the linker region. Since Dbh is far more active at higher temperatures, this leads me to conclude the free reorientation of the LF is necessary for polymerase function, including binding and proper positioning of substrates for catalysis. The MD simulations also indicate the cores of the domains of Dbh are very rigid at 35°C, and even quite rigid at 50°C, indicating that rigidity is crucial to the stability of Dbh at high temperature.

## References:


Acharya, N., Haracska, L., Johnson, R.E., Unk, I., Prakash, S., and Prakash, L. (2005). Complex formation of yeast Rev1 and Rev7 proteins: a novel role for the polymerase-associated domain. Mol Cell Biol 25, 9734-9740.

Acharya, N., Yoon, J.H., Gali, H., Unk, I., Haracska, L., Johnson, R.E., Hurwitz, J., Prakash, L., and Prakash, S. (2008). Roles of PCNA-binding and ubiquitin-binding domains in human DNA polymerase eta in translesion DNA synthesis. Proc Natl Acad Sci U S A 105, 17724-17729.

Albertella, M.R., Green, C.M., Lehmann, A.R., and O'Connor, M.J. (2005). A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. Cancer Res 65, 9799-9806.

Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.P., and Carell, T. (2007). Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. Science 318, 967-970.

Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E.C., and Livneh, Z. (2004). Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa. J Biol Chem 279, 53298-53305.

Azuaga, A.I., Galisteo, M.L., Mayorga, O.L., Cortijo, M., and Mateo, P.L. (1992). Heat and cold denaturation of beta-lactoglobulin B. FEBS Lett 309, 258-260.

Baker-Austin, C., and Dopson, M. (2007). Life in acid: pH homeostasis in acidophiles. In Trends Microbiol (England), pp. 165-171.

Bessman, M.J., Kornberg, A., Lehman, I.R., and Simms, E.S. (1956). Enzymic synthesis of deoxyribonucleic acid. Biochim Biophys Acta 21, 197-198.

Bianchi, J., Rudd, S.G., Jozwiakowski, S.K., Bailey, L.J., Soura, V., Taylor, E., Stevanovic, I., Green, A.J., Stracker, T.H., Lindsay, H.D.*, et al.* (2013). PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. Mol Cell 52, 566-573.

Bleicher, L., Prates, E.T., Gomes, T.C., Silveira, R.L., Nascimento, A.S., Rojas, A.L., Golubev, A., Martinez, L., Skaf, M.S., and Polikarpov, I. (2011). Molecular basis of the thermostability and thermophilicity of laminarinases: X-ray structure of the hyperthermostable laminarinase from Rhodothermus marinus and molecular dynamics simulations. J Phys Chem B 115, 7940-7949.

Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001). Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol? In Nucleic Acids Res, pp. 4607-4616.

Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W., and Woodgate, R. (2004). Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. J Biol Chem 279, 32932-32940.

Brenlla, A., Markiewicz, R.P., Rueda, D., and Romano, L.J. (2014). Nucleotide selection by the Y-family DNA polymerase Dpo4 involves template translocation and misalignment. Nucleic Acids Res 42, 2555-2563.

Buchner, G.S., Shih, N., Reece, A.E., Niebling, S., and Kubelka, J. (2012). Unusual cold denaturation of a small protein domain. Biochemistry 51, 6496-6498.

Budzowska, M., Graham, T.G., Sobeck, A., Waga, S., and Walter, J.C. (2015). Regulation of the Rev1-pol zeta complex during bypass of a DNA interstrand cross-link. Embo j.

Butterwick, J.A., and Palmer, A.G., 3rd (2006). An inserted Gly residue fine tunes dynamics between mesophilic and thermophilic ribonucleases H. In Protein Sci (United States), pp. 2697-2707.

Ceppi, P., Novello, S., Cambieri, A., Longo, M., Monica, V., Lo Iacono, M., Giaj-Levra, M., Saviozzi, S., Volante, M., Papotti, M.*, et al.* (2009). Polymerase eta mRNA expression predicts survival of non-small cell lung cancer patients treated with platinum-based chemotherapy. Clin Cancer Res 15, 1039-1045.

Chu, X., Liu, F., Maxwell, B.A., Wang, Y., Suo, Z., Wang, H., Han, W., and Wang, J. (2014). Dynamic Conformational Change Regulates the Protein-DNA Recognition: An Investigation on Binding of a Y-Family Polymerase to Its Target DNA. PLoS Comput Biol 10, e1003804.

D'Souza, S., and Walker, G.C. (2006). Novel role for the C terminus of Saccharomyces cerevisiae Rev1 in mediating protein-protein interactions. Mol Cell Biol 26, 8173-8182.

Davidovic, M., Mattea, C., Qvist, J., and Halle, B. (2009). Protein cold denaturation as seen from the solvent. J Am Chem Soc 131, 1025-1036.

Delbos, F., Aoufouchi, S., Faili, A., Weill, J.C., and Reynaud, C.A. (2007). DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. J Exp Med 204, 17-23.

DeLucia, A.M., Grindley, N.D., and Joyce, C.M. (2007). Conformational changes during normal and error-prone incorporation of nucleotides by a Y-family DNA polymerase detected by 2-aminopurine fluorescence. Biochemistry 46, 10790-10803.

Doles, J., Oliver, T.G., Cameron, E.R., Hsu, G., Jacks, T., Walker, G.C., and Hemann, M.T. (2010). Suppression of Rev3, the catalytic subunit of Pol{zeta}, sensitizes drug-resistant lung tumors to chemotherapy. Proc Natl Acad Sci U S A 107, 20786-20791.

Donny-Clark, K., and Broyde, S. (2009). Influence of local sequence context on damaged base conformation in human DNA polymerase iota: molecular dynamics studies of nucleotide incorporation opposite a benzo[a]pyrene-derived adenine lesion. Nucleic Acids Res 37, 7095-7109.

Donny-Clark, K., Shapiro, R., and Broyde, S. (2009). Accommodation of an N-(deoxyguanosin-8-yl)-2-acetylaminofluorene adduct in the active site of human DNA polymerase iota: Hoogsteen or Watson-Crick base pairing? Biochemistry 48, 7-18.

Dumstorf, C.A., Clark, A.B., Lin, Q., Kissling, G.E., Yuan, T., Kucherlapati, R., McGregor, W.G., and Kunkel, T.A. (2006). Participation of mouse DNA polymerase iota in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. Proc Natl Acad Sci U S A 103, 18083-18088.

Elcock, A.H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. In J Mol Biol (England, 1998 Academic Press.), pp. 489-502.

Eoff, R.L., Sanchez-Ponce, R., and Guengerich, F.P. (2009). Conformational changes during nucleotide selection by Sulfolobus solfataricus DNA polymerase Dpo4. J Biol Chem 284, 21090-21099.

Faili, A., Aoufouchi, S., Flatter, E., Gueranger, Q., Reynaud, C.A., and Weill, J.C. (2002). Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase iota. Nature 419, 944-947.

Fukuchi, S., and Nishikawa, K. (2001). Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria. J Mol Biol 309, 835-843.

Furukohri, A., Goodman, M.F., and Maki, H. (2008). A dynamic polymerase exchange with Escherichia coli DNA polymerase IV replacing DNA polymerase III on the sliding clamp. J Biol Chem 283, 11260-11269.

Galzitskaya, O.V., Garbuzynskiy, S.O., and Lobanov, M.Y. (2006). FoldUnfold: web server for the prediction of disordered regions in protein chain. In Bioinformatics (England), pp. 2948-2949.

Garcia-Gomez, S., Reyes, A., Martinez-Jimenez, M.I., Chocron, E.S., Mouron, S., Terrados, G., Powell, C., Salido, E., Mendez, J., Holt, I.J.*, et al.* (2013). PrimPol, an archaic primase/polymerase operating in human cells. Mol Cell 52, 541-553.

Gerlach, V.L., Feaver, W.J., Fischhaber, P.L., and Friedberg, E.C. (2001). Purification and characterization of pol kappa, a DNA polymerase encoded by the human DINB1 gene. J Biol Chem 276, 92-98.

Graziano, G. (2014). On the mechanism of cold denaturation. Phys Chem Chem Phys 16, 21755-21767.

Gregory, M.T., Park, G.Y., Johnstone, T.C., Lee, Y.S., Yang, W., and Lippard, S.J. (2014). Structural and mechanistic studies of polymerase eta bypass of phenanthriplatin DNA damage. Proc Natl Acad Sci U S A 111, 9133-9138.

Griko , Y.V., and Kutyshenko, V.P. (1994). Differences in the processes of beta-lactoglobulin cold and heat denaturations. Biophys J 67, 356-363.

Griko, Y.V., and Privalov, P.L. (1992). Calorimetric study of the heat and cold denaturation of beta-lactoglobulin. Biochemistry 31, 8810-8815.

Gromiha, M.M., Oobatake, M., and Sarai, A. (1999). Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. Biophys Chem 82, 51-67.

Gromiha, M.M., and Suresh, M.X. (2008). Discrimination of mesophilic and thermophilic proteins using machine learning algorithms. Proteins 70, 1274-1279.

Haracska, L., Acharya, N., Unk, I., Johnson, R.E., Hurwitz, J., Prakash, L., and Prakash, S. (2005). A single domain in human DNA polymerase iota mediates interaction with PCNA: implications for translesion DNA synthesis. Mol Cell Biol 25, 1183-1190.

Hishiki, A., Hashimoto, H., Hanafusa, T., Kamei, K., Ohashi, E., Shimizu, T., Ohmori, H., and Sato, M. (2009). Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen. J Biol Chem 284, 10552-10560.

Hori, M., Yonekura, S., Nohmi, T., Gruz, P., Sugiyama, H., Yonei, S., and Zhang-Akiyama, Q.M. (2010). Error-Prone Translesion DNA Synthesis by Escherichia coli DNA Polymerase IV (DinB) on Templates Containing 1,2-dihydro-2-oxoadenine. J Nucleic Acids 2010, 807579.

Ikeda, M., Furukohri, A., Philippin, G., Loechler, E., Akiyama, M.T., Katayama, T., Fuchs, R.P., and Maki, H. (2014). DNA polymerase IV mediates efficient and quick recovery of replication forks stalled at N2-dG adducts. Nucleic Acids Res 42, 8461-8472.

J, O.W., Kawamura, K., Tada, Y., Ohmori, H., Kimura, H., Sakiyama, S., and Tagawa, M. (2001). DNA polymerase kappa, implicated in spontaneous and DNA damage-induced mutagenesis, is overexpressed in lung cancer. Cancer Res 61, 5366-5369.

Jia, L., Geacintov, N.E., and Broyde, S. (2008). The N-clasp of human DNA polymerase kappa promotes blockage or error-free bypass of adenine- or guanine-benzo[a]pyrenyl lesions. Nucleic Acids Res 36, 6571-6584.

Johnson, R.E., Kondratick, C.M., Prakash, S., and Prakash, L. (1999a). hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 285, 263-265.

Johnson, R.E., Prakash, S., and Prakash, L. (1999b). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 283, 1001-1004.

Jonsdottir, L.B., Ellertsson, B.O., Invernizzi, G., Magnusdottir, M., Thorbjarnardottir, S.H., Papaleo, E., and Kristjansson, M.M. (2014). The role of salt bridges on the temperature adaptation of aqualysin I, a thermostable subtilisin-like proteinase. Biochim Biophys Acta 1844, 2174-2181.

Joo, J.C., Pack, S.P., Kim, Y.H., and Yoo, Y.J. (2011). Thermostabilization of Bacillus circulans xylanase: computational optimization of unstable residues based on thermal fluctuation analysis. J Biotechnol 151, 56-65.

Joyce, C.M., and Benkovic, S.J. (2004). DNA polymerase fidelity: kinetics, structure, and checkpoints. Biochemistry 43, 14317-14324.

Kamal, M.Z., Mohammad, T.A., Krishnamoorthy, G., and Rao, N.M. (2012). Role of active site rigidity in activity: MD simulation and fluorescence study on a lipase mutant. In PLoS One (United States), p. e35188.

Karshikoff, A., Nilsson, L., and Ladenstein, R. (2015). Rigidity versus flexibility: the dilemma of understanding protein thermal stability. Febs j.

Kath, J.E., Jergic, S., Heltzel, J.M., Jacob, D.T., Dixon, N.E., Sutton, M.D., Walker, G.C., and Loparo, J.J. (2014). Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. Proc Natl Acad Sci U S A 111, 7647-7652.

Knobel, P.A., Kotov, I.N., Felley-Bosco, E., Stahel, R.A., and Marti, T.M. (2011). Inhibition of REV3 expression induces persistent DNA damage and growth arrest in cancer cells. Neoplasia 13, 961-970.

Kosarek, J.N., Woodruff, R.V., Rivera-Begeman, A., Guo, C., D'Souza, S., Koonin, E.V., Walker, G.C., and Friedberg, E.C. (2008). Comparative analysis of in vivo interactions between Rev1 protein and other Y-family DNA polymerases in animals and yeasts. DNA Repair (Amst) 7, 439-451.

Kumar, R., Prabhu, N.P., Rao, D.K., and Bhuyan, A.K. (2006). The alkali molten globule state of horse ferricytochrome c: observation of cold denaturation. J Mol Biol 364, 483-495.

Lange, S.S., Takata, K., and Wood, R.D. (2011). DNA polymerases and cancer. Nat Rev Cancer 11, 96-110.

Lee, Y.S., Gregory, M.T., and Yang, W. (2014). Human Pol zeta purified with accessory subunits is active in translesion DNA synthesis and complements Pol eta in cisplatin bypass. Proc Natl Acad Sci U S A 111, 2954-2959.

Lin, Q., Clark, A.B., McCulloch, S.D., Yuan, T., Bronson, R.T., Kunkel, T.A., and Kucherlapati, R. (2006). Increased susceptibility to UV-induced skin carcinogenesis in polymerase eta-deficient mice. Cancer Res 66, 87-94.

Lin, W., Xin, H., Zhang, Y., Wu, X., Yuan, F., and Wang, Z. (1999). The human REV1 gene codes for a DNA template-dependent dCMP transferase. Nucleic Acids Res 27, 4468-4475.

Lior-Hoffmann, L., Ding, S., Geacintov, N.E., Zhang, Y., and Broyde, S. (2014). Structural and Dynamic Characterization of Polymerase kappa's Minor Groove Lesion Processing Reveals How Adduct Topology Impacts Fidelity. Biochemistry 53, 5683-5691.

Makarova, A.V., and Burgers, P.M. (2015). Eukaryotic DNA polymerase zeta. DNA Repair (Amst) 29, 47-55.

Mamonova, T., Hespenheide, B., Straub, R., Thorpe, M.F., and Kurnikova, M. (2005). Protein flexibility using constraints from molecular dynamics simulations. In Phys Biol (England), pp. S137-147.

Mamonova, T.B., Glyakina, A.V., Galzitskaya, O.V., and Kurnikova, M.G. (2013). Stability and rigidity/flexibility-two sides of the same coin? Biochim Biophys Acta 1834, 854-866.

Maxwell, B.A., and Suo, Z. (2014). Recent insight into the kinetic mechanisms and conformational dynamics of Y-Family DNA polymerases. Biochemistry 53, 2804-2814.

Maxwell, B.A., Xu, C., and Suo, Z. (2012). DNA lesion alters global conformational dynamics of Y-family DNA polymerase during catalysis. J Biol Chem 287, 13040-13047.

Maxwell, B.A., Xu, C., and Suo, Z. (2014). Conformational dynamics of a Y-family DNA polymerase during substrate binding and catalysis as revealed by interdomain Forster resonance energy transfer. Biochemistry 53, 1768-1778.

McDonald, J.P., Rapic-Otrin, V., Epstein, J.A., Broughton, B.C., Wang, X., Lehmann, A.R., Wolgemuth, D.J., and Woodgate, R. (1999). Novel human and mouse homologs of Saccharomyces cerevisiae DNA polymerase eta. Genomics 60, 20-30.

Meharenna, Y.T., and Poulos, T.L. (2010). Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450. Biochemistry 49, 6680-6686.

Mouron, S., Rodriguez-Acebes, S., Martinez-Jimenez, M.I., Garcia-Gomez, S., Chocron, S., Blanco, L., and Mendez, J. (2013). Repriming of DNA synthesis at stalled replication forks by human PrimPol. Nat Struct Mol Biol 20, 1383-1389.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996a). Deoxycytidyl transferase activity of yeast REV1 protein. Nature 382, 729-731.

Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996b). Thymine-thymine dimer bypass by yeast DNA polymerase zeta. Science 272, 1646-1649.

Ogi, T., Kannouche, P., and Lehmann, A.R. (2005). Localisation of human Y-family DNA polymerase kappa: relationship to PCNA foci. J Cell Sci 118, 129-136.

Ogi, T., Kato, T., Jr., Kato, T., and Ohmori, H. (1999). Mutation enhancement by DINB1, a mammalian homologue of the Escherichia coli mutagenesis protein dinB. Genes Cells 4, 607-618.

Ogi, T., Shinkai, Y., Tanaka, K., and Ohmori, H. (2002). Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. Proc Natl Acad Sci U S A 99, 15548-15553.

Ohashi, E., Bebenek, K., Matsuda, T., Feaver, W.J., Gerlach, V.L., Friedberg, E.C., Ohmori, H., and Kunkel, T.A. (2000). Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. J Biol Chem 275, 39678-39684.

Ohkumo, T., Kondo, Y., Yokoi, M., Tsukamoto, T., Yamada, A., Sugimoto, T., Kanao, R., Higashi, Y., Kondoh, H., Tatematsu, M.*, et al.* (2006). UV-B radiation induces epithelial tumors in mice lacking DNA polymerase eta and mesenchymal tumors in mice deficient for DNA polymerase iota. Mol Cell Biol 26, 7696-7706.

Ohmori, H., Friedberg, E.C., Fuchs, R.P., Goodman, M.F., Hanaoka, F., Hinkle, D., Kunkel, T.A., Lawrence, C.W., Livneh, Z., Nohmi, T.*, et al.* (2001). The Y-family of DNA polymerases. Mol Cell 8, 7-8.

Pastore, A., Martin, S.R., Politou, A., Kondapalli, K.C., Stemmler, T., and Temussi, P.A. (2007). Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. J Am Chem Soc 129, 5374-5375.

Pata, J.D. (2010). Structural diversity of the Y-family DNA polymerases. Biochim Biophys Acta 1804, 1124-1135.

Perlow-Poehnelt, R.A., Likhterov, I., Scicchitano, D.A., Geacintov, N.E., and Broyde, S. (2004). The spacious active site of a Y-family DNA polymerase facilitates promiscuous nucleotide incorporation opposite a bulky carcinogen-DNA adduct: elucidating the structure-function relationship through experimental and computational approaches. J Biol Chem 279, 36951-36961.

Perlow-Poehnelt, R.A., Likhterov, I., Wang, L., Scicchitano, D.A., Geacintov, N.E., and Broyde, S. (2007). Increased flexibility enhances misincorporation: temperature effects on nucleotide incorporation opposite a bulky carcinogen-DNA adduct by a Y-family DNA polymerase. J Biol Chem 282, 1397-1408.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605-1612.

Pezeshgi Modarres, H., Dorokhov, B.D., Popov, V.O., Ravin, N.V., Skryabin, K.G., and Dal Peraro, M. (2015). Understanding and Engineering Thermostability in DNA Ligase from Thermococcus sp. 1519. Biochemistry.

Plemenitas, A., Lenassi, M., Konte, T., Kejzar, A., Zajc, J., Gostincar, C., and Gunde-Cimerman, N. (2014). Adaptation to high salt concentrations in halotolerant/halophilic fungi: a molecular perspective. Front Microbiol 5, 199.

Pometun, M.S., Peterson, R.W., Babu, C.R., and Wand, A.J. (2006). Cold denaturation of encapsulated ubiquitin. J Am Chem Soc 128, 10652-10653.

Potapova, O., Grindley, N.D., and Joyce, C.M. (2002). The mutational specificity of the Dbh lesion bypass polymerase and its implications. J Biol Chem 277, 28157-28166.

Privalov, P.L. (1990). Cold denaturation of proteins. Crit Rev Biochem Mol Biol 25, 281-305.

Pustovalova, Y., Maciejewski, M.W., and Korzhnev, D.M. (2013). NMR mapping of PCNA interaction with translesion synthesis DNA polymerase Rev1 mediated by Rev1-BRCT domain. J Mol Biol 425, 3091-3105.

Qin, Y., Yang, Y., Zhang, L., Fowler, J.D., Qiu, W., Wang, L., Suo, Z., and Zhong, D. (2013). Direct probing of solvent accessibility and mobility at the binding interface of polymerase (Dpo4)-DNA complex. J Phys Chem A 117, 13926-13934.

Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N.E., and Patel, D.J. (2006). Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion. PLoS Biol 4, e11.

Rothwell, P.J., Mitaksov, V., and Waksman, G. (2005). Motions of the fingers subdomain of klentaq1 are fast and not rate limiting: implications for the molecular basis of fidelity in DNA polymerases. Mol Cell 19, 345-355.

Sadeghi, M., Naderi-Manesh, H., Zarrabi, M., and Ranjbar, B. (2006). Effective factors in thermostability of thermophilic proteins. In Biophys Chem (Netherlands), pp. 256-270.

Sakofsky, C.J., Foster, P.L., and Grogan, D.W. (2012). Roles of the Y-family DNA polymerase Dbh in accurate replication of the Sulfolobus genome at high temperature. In DNA Repair (Amst) (Netherlands, 2012 Elsevier B.V), pp. 391-400.

Sale, J.E., Lehmann, A.R., and Woodgate, R. (2012). Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat Rev Mol Cell Biol 13, 141-152.

Sekimoto, T., Oda, T., Kurashima, K., Hanaoka, F., and Yamashita, T. (2015). Both high-fidelity replicative and low-fidelity Y-family polymerases are involved in DNA rereplication. Mol Cell Biol 35, 699-715.

Sharma, S., and Canman, C.E. (2012). REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. Environ Mol Mutagen 53, 725-740.

Shechner, D.M., Grant, R.A., Bagby, S.C., Koldobskaya, Y., Piccirilli, J.A., and Bartel, D.P. (2009). Crystal structure of the catalytic core of an RNA-polymerase ribozyme. Science 326, 1271-1275.

Sholder, G., and Loechler, E.L. (2015). A method to accurately quantitate intensities of (32)P-DNA bands when multiple bands appear in a single lane of a gel is used to study dNTP insertion opposite a benzo[a]pyrene-dG adduct by Sulfolobus DNA polymerases Dpo4 and Dbh. DNA Repair (Amst) 25, 97-103.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8, 984-989.

Srivastava, A., and Sinha, S. (2014). Thermostability of in vitro evolved Bacillus subtilis lipase A: a network and dynamics perspective. PLoS One 9, e102856.

Stafford, K.A., Robustelli, P., and Palmer, A.G., 3rd (2013). Thermal adaptation of conformational dynamics in ribonuclease H. PLoS Comput Biol 9, e1003218.

Stafford, K.A., Trbovic, N., Butterwick, J.A., Abel, R., Friesner, R.A., and Palmer, A.G., 3rd (2015). Conformational preferences underlying reduced activity of a thermophilic ribonuclease H. J Mol Biol 427, 853-866.

Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. J Biol Chem 274, 17395-17398.

Szilagyi, A., and Zavodszky, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. In Structure (England), pp. 493-504.

Tissier, A., McDonald, J.P., Frank, E.G., and Woodgate, R. (2000). poliota, a remarkably error-prone human DNA polymerase. Genes Dev 14, 1642-1650.

Ummat, A., Rechkoblit, O., Jain, R., Roy Choudhury, J., Johnson, R.E., Silverstein, T.D., Buku, A., Lone, S., Prakash, L., Prakash, S.*, et al.* (2012). Structural basis for cisplatin DNA damage tolerance by human polymerase eta during cancer chemotherapy. Nat Struct Mol Biol 19, 628-632.

Vaisman, A., Masutani, C., Hanaoka, F., and Chaney, S.G. (2000). Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase eta. Biochemistry 39, 4575-4580.

Vaisman, A., Tissier, A., Frank, E.G., Goodman, M.F., and Woodgate, R. (2001). Human DNA polymerase iota promiscuous mismatch extension. J Biol Chem 276, 30615-30622.

Vajpai, N., Nisius, L., Wiktor, M., and Grzesiek, S. (2013). High-pressure NMR reveals close similarity between cold and alcohol protein denaturation in ubiquitin. Proc Natl Acad Sci U S A 110, E368-376.

Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P., and Nohmi, T. (1999). The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. Mol Cell 4, 281-286.

Wang, L., and Broyde, S. (2006). A new anti conformation for N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG) allows Watson-Crick pairing in the Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4). Nucleic Acids Res 34, 785-795.

Wang, L., Yu, X., Hu, P., Broyde, S., and Zhang, Y. (2007a). A water-mediated and substrate-assisted catalytic mechanism for Sulfolobus solfataricus DNA polymerase IV. J Am Chem Soc 129, 4731-4737.

Wang, Y., Woodgate, R., McManus, T.P., Mead, S., McCormick, J.J., and Maher, V.M. (2007b). Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase eta, DNA polymerase iota causes the very high frequency and unique spectrum of UV-induced mutations. Cancer Res 67, 3018-3026.

Whitten, S.T., Kurtz, A.J., Pometun, M.S., Wand, A.J., and Hilser, V.J. (2006). Revealing the nature of the native state ensemble through cold denaturation. Biochemistry 45, 10163-10174.

Wieteska, L., Ionov, M., Szemraj, J., Feller, C., Kolinski, A., and Gront, D. (2015). Improving thermal stability of thermophilic l-threonine aldolase from Thermotoga maritima. J Biotechnol 199, 69-76.

Wilson, R.C., and Pata, J.D. (2008). Structural insights into the generation of single-base deletions by the Y family DNA polymerase dbh. Mol Cell 29, 767-779.

Wiltrout, M.E., and Walker, G.C. (2011). The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. Genetics 187, 21-35.

Wong, J.H., Fiala, K.A., Suo, Z., and Ling, H. (2008). Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. J Mol Biol 379, 317-330.

Wu, T.H., Chen, C.C., Cheng, Y.S., Ko, T.P., Lin, C.Y., Lai, H.L., Huang, T.Y., Liu, J.R., and Guo, R.T. (2014). Improving specific activity and thermostability of Escherichia coli phytase by structure-based rational design. J Biotechnol 175, 1-6.

Xing, G., Kirouac, K., Shin, Y.J., Bell, S.D., and Ling, H. (2009). Structural insight into recruitment of translesion DNA polymerase Dpo4 to sliding clamp PCNA. Mol Microbiol 71, 678-691.

Xu, C., Maxwell, B.A., Brown, J.A., Zhang, L., and Suo, Z. (2009a). Global conformational dynamics of a Y-family DNA polymerase during catalysis. PLoS Biol 7, e1000225.

Xu, P., Oum, L., Beese, L.S., Geacintov, N.E., and Broyde, S. (2007). Following an environmental carcinogen N2-dG adduct through replication: elucidating blockage and bypass in a high-fidelity DNA polymerase. Nucleic Acids Res 35, 4275-4288.

Xu, P., Oum, L., Geacintov, N.E., and Broyde, S. (2008). Nucleotide selectivity opposite a benzo[a]pyrene-derived N2-dG adduct in a Y-family DNA polymerase: a 5'-slippage mechanism. Biochemistry 47, 2701-2709.

Xu, P., Oum, L., Lee, Y.C., Geacintov, N.E., and Broyde, S. (2009b). Visualizing sequence-governed nucleotide selectivities and mutagenic consequences through a replicative cycle: processing of a bulky carcinogen N2-dG lesion in a Y-family DNA polymerase. Biochemistry 48, 4677-4690.

Yamtich, J., and Sweasy, J.B. (2010). DNA polymerase family X: function, structure, and cellular roles. Biochim Biophys Acta 1804, 1136-1150.

Yang, C., Jang, S., and Pak, Y. (2014). A fully atomistic computer simulation study of cold denaturation of a beta-hairpin. Nat Commun 5, 5773.

Yang, J., Chen, Z., Liu, Y., Hickey, R.J., and Malkas, L.H. (2004). Altered DNA polymerase iota expression in breast cancer cells leads to a reduction in DNA replication fidelity and a higher rate of mutagenesis. Cancer Res 64, 5597-5607.

Yang, W., and Woodgate, R. (2007). What a difference a decade makes: insights into translesion DNA synthesis. Proc Natl Acad Sci U S A 104, 15591-15598.

Yip, K.S., Stillman, T.J., Britton, K.L., Artymiuk, P.J., Baker, P.J., Sedelnikova, S.E., Engel, P.C., Pasquo, A., Chiaraluce, R., Consalvi, V.*, et al.* (1995). The structure of Pyrococcus furiosus glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. Structure 3, 1147-1158.

Yoon, J.H., Prakash, L., and Prakash, S. (2009). Highly error-free role of DNA polymerase eta in the replicative bypass of UV-induced pyrimidine dimers in mouse and human cells. Proc Natl Acad Sci U S A 106, 18219-18224.

Yu, H., Zhao, Y., Guo, C., Gan, Y., and Huang, H. (2015). The role of proline substitutions within flexible regions on thermostability of luciferase. Biochim Biophys Acta 1854, 65-72.

Yuan, F., Xu, Z., Yang, M., Wei, Q., Zhang, Y., Yu, J., Zhi, Y., Liu, Y., Chen, Z., and Yang, J. (2013). Overexpressed DNA polymerase iota regulated by JNK/c-Jun contributes to hypermutagenesis in bladder cancer. PLoS One 8, e69317.

Zafar, M.K., Ketkar, A., Lodeiro, M.F., Cameron, C.E., and Eoff, R.L. (2014). Kinetic analysis of human PrimPol DNA polymerase activity reveals a generally error-prone enzyme capable of accurately bypassing 7,8-dihydro-8-oxo-2'-deoxyguanosine. Biochemistry 53, 6584-6594.

Zhang, Y., Wu, X., Guo, D., Rechkoblit, O., and Wang, Z. (2002). Activities of human DNA polymerase kappa in response to the major benzo[a]pyrene DNA adduct: error-free lesion bypass and extension synthesis from opposite the lesion. DNA Repair (Amst) 1, 559-569.

Zhang, Y., Yuan, F., Wu, X., Taylor, J.S., and Wang, Z. (2001). Response of human DNA polymerase ι to DNA lesions. In Nucleic Acids Res, pp. 928-935.

Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J.S., Geacintov, N.E., and Wang, Z. (2000a). Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro. Nucleic Acids Res 28, 4138-4146.

Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000b). Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. Mol Cell Biol 20, 7099-7108.

Zhao, L., Pence, M.G., Eoff, R.L., Yuan, S., Fercu, C.A., and Guengerich, F.P. (2014). Elucidation of kinetic mechanisms of human translesion DNA polymerase kappa using tryptophan mutants. Febs j 281, 4394-4410.

Zhou, J., Zhang, S., Xie, L., Liu, P., Xie, F., Wu, J., Cao, J., and Ding, W.Q. (2012). Overexpression of DNA polymerase iota (Poliota) in esophageal squamous cell carcinoma. Cancer Sci 103, 1574-1579.

## Chapter 2 – $^{1}H$, $^{15}N$, and $^{13}C$ backbone assignments of *dinB* homolog (Dbh)

**Published in *Biomolecular NMR Assignments*, 2015, 9(2): 441-445, with kind permission from Springer Science, dx.doi.org/10.1007/s12104-015-9626-y**

### Abstract

The *dinB* homolog (Dbh) is a member of the Y-family of translesion DNA polymerases, which are specialized to accurately replicate DNA across from a wide variety of lesions in living cells. Lesioned bases block the progression of high-fidelity polymerases and cause detrimental replication fork stalling; Y-family polymerases can bypass these lesions. The active site of the translesion synthesis polymerase is more open than that of a replicative polymerase; consequently Dbh polymerizes with low fidelity. Bypass polymerases also have low processivity. Short extension past the lesion allows the high-fidelity polymerase to switch back onto the site of replication. Dbh and the other Y-family polymerases have been used as structural models to investigate the mechanisms of DNA polymerization and lesion bypass. Many high-resolution crystal structures of Y-family polymerases have been reported. NMR dynamics studies can complement these structures by providing a measure of protein motions. I have assigned the $^{15}N$, $^{1}H$, and $^{13}C$ backbone resonance signals at two temperatures (35°C and 50°C) for *Sulfolobus acidocaldarius* Dbh polymerase. Backbone resonance assignments have been obtained for 86% of the residues. The polymerase active site is assigned as well as the majority of residues in each of the four domains.

### Rationale and Strategy of Experiments

NMR relaxation spectroscopy provides atomic level resolution concerning the dynamics of proteins. However, in order to interpret NMR spectra, it is necessary to assign the individual resonances in the HSQC spectra to their corresponding atoms in the protein. To accomplish this goal, series of 2D and 3D spectra must be obtained to establish connectivity between sets of resonances. First, sets of sequential resonances are identified. Since the primary sequence of the protein is already known, myriad methods can be used to determine individual amino acid residue identity within the sequential segment. Most amino acid residue types can be determined, and eventually a long enough segment with one or more known amino acid residues is obtained. By a process of elimination, it can be determined unambiguously that the sequential segment is equivalent to a particular portion of the primary sequence. When this condition is fulfilled, it can be said that that portion of the polypeptide chain is now “assigned”.

The approach to assigning individual atoms to resonances in the NMR spectra of a polypeptide depends on several parameters. The size of the polypeptide correlates with how many resonances are visible, with increasing polypeptide size there is more chance of crowding of signals in the NMR spectra. Additionally, with increasing molecular weight the rotational correlation time (or $\tau_m$) of the protein increases and the transverse relaxation time ($T_2$) decreases, leading to loss of signal strength and providing a further complication (Frueh, 2014). Well-folded proteins with stable secondary structure show the greatest level of spectral dispersion, but if the protein has disordered regions or repeats, or it is over 300 or so residues, the spectra can become extremely crowded (Frueh, 2014). Proteins with high levels of conformational exchange or solvent-exposed surface area will have signals be lost due to exchange broadening

and exchange with solvent, respectively (Pervushin, 2001). Given these caveats, the spectroscopist must consider the particular challenges of assigning the protein in question, and choose the appropriate experiments which are essential for assignment.

At a minimum, the spectroscopist will seek to assign as many of the polypeptide "backbone" $^{1}H$, $^{15}N$, and $^{13}C$ resonances as possible, and if possible, the side chain resonances as well. The first step in assigning the backbone resonances is collecting a $^{15}N$-$^{1}H$ 2D HSQC (Heteronuclear Spectroscopy, Quantum Correlated, which detects the cross-peaks for every N-H bond in the protein: the backbone amide NH bonds, the $NH_2$ bonds of asparagine and glutamine side chains, and the $\varepsilon_1$-NH bond of the tryptophan side chain (Bodenhausen and Ruben, 1980). Lysine ε-NH and arginine ε-NH resonances are in a different frequency range from the backbone amides; hence, they are not within the frequency range of the typical HSQC experiment. The HSQC is a relatively fast experiment, which allows the spectroscopist to determine if the protein is suitable for assignment. If the signals are well dispersed, indicative of a stable protein fold, and the signal-to-noise ratio is sufficient, more time-consuming and costly 3D experiments, which are necessary for assignment, can be performed. If the protein is above a certain molecular weight, the TROSY (Transverse Relaxation-Optimized SpectroscopY, (Pervushin et al., 1997) method can be used to optimize the signal-to-noise ratio, mitigating the signal loss due to decreased transverse relaxation time.

The necessary component of any NMR assignment strategy is to determine sequential backbone resonances. The most common approach is to use 3D ($^{1}H$, $^{15}N$, $^{13}C$) experiments to connect a particular NH resonance with its own backbone carbon resonances (Cα and CO, of the "i" residue) with the backbone carbon resonances of

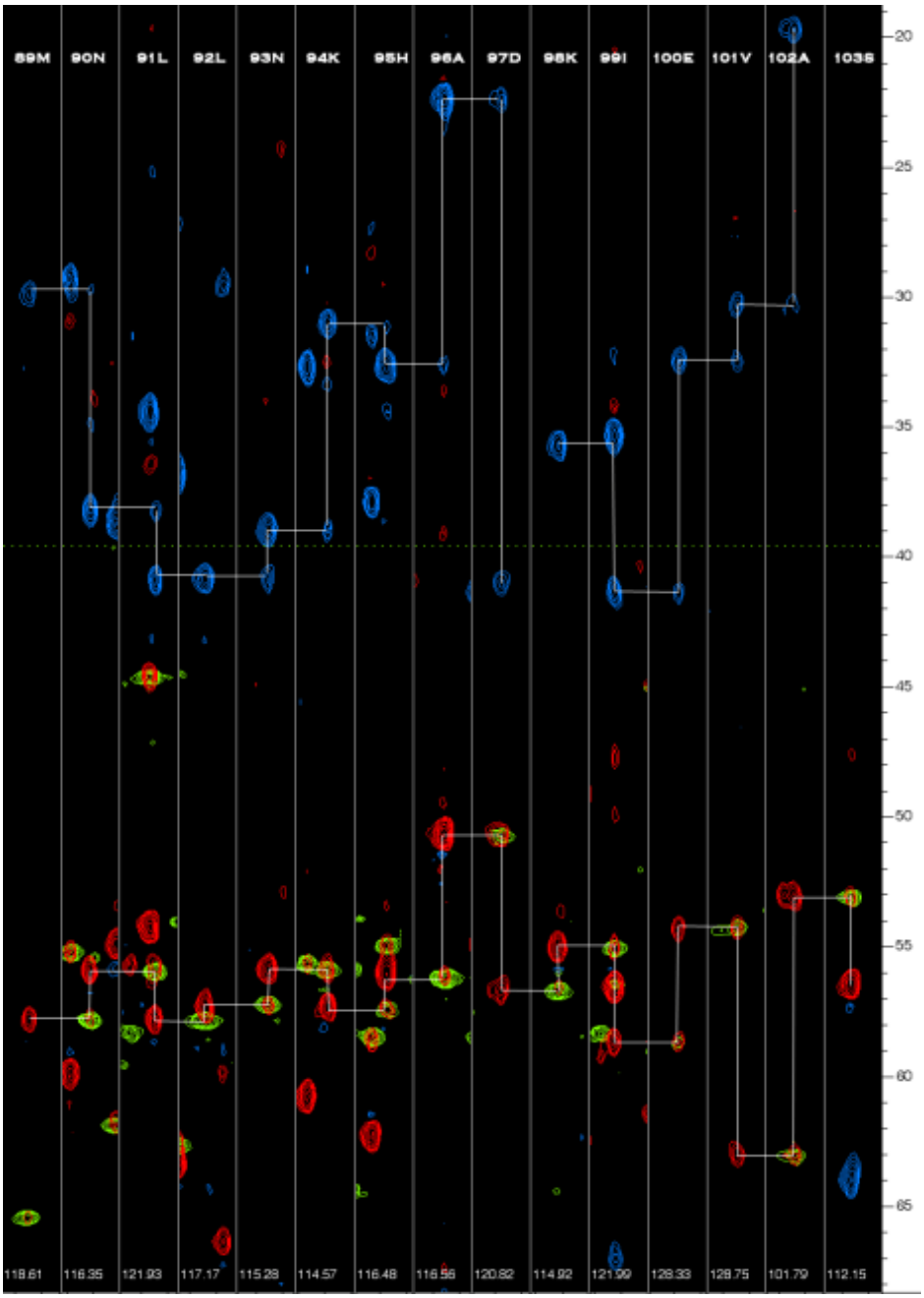
adjacent residues [the i-1 and/or i+1 residues (Ikura et al., 1990; Leopold et al., 1994)]. A spectrum that detects only the carbons from adjacent residues through the “i” NH can be overlaid with a spectrum that detects the carbons from both adjacent residues and the carbons from the “i” residue. The adjacent residue NH position can be found in the 3D data through matching the i-1 resonance of the current NH with the “i” resonance of the adjacent residue. Typically, connectivity is determined for at least the Cα and CO carbons, as well as the Cβ carbons. A representative strip plot displaying an HNCACB spectrum (detecting $C\alpha_i$, $C\beta_i$, $C\alpha_{i-1}$, and $C\beta_{i-1}$ through $NH_i$) overlaid with an HNCOCA spectrum (detecting only $C\alpha_{i-1}$ through $NH_i$) visually demonstrates how the process of determining sequential residues is achieved (Figure 2.1). The connectivity through the CO atoms, combined with connectivity through Cα and Cβ atoms, provides strong evidence that the resonances detected belong to sequential residues. Nevertheless, for a protein of any significant size, it is inevitable there will be some overlap in a significant portion of the spectrum, which is the case for Dbh. In this case, it can be extremely helpful to determine which resonances belong to a particular amino acid type through the preparation of selectively-unlabeled HSQCs.

There are few amino acid residues which can be confidently determined by the carbon shifts alone: glycine, which has Cα shifts less than 47ppm, alanine, with Cβ shifts less than 25 ppm, and serine or threonine, which have Cβ shifts greater than 60 ppm (BMRB database, http://www.bmrb.wisc.edu/ref_info/statsel.htm, (Ulrich et al., 2008). Outside of these residues, the Cα and Cβ shifts of the other amino acids show a significant amount of overlap. This makes it impossible to unambiguously determine the amino acid type from Cα and/or Cβ carbon shifts alone. The amino acid type can be narrowed down to a

few probable residue types using the Cα and Cβ, but to be confident of the amino acid, a selectively unlabeled or labeled sample should be used.

To selectively unlabel or label samples, the protein is recombinantly expressed in *E. coli* growing in minimal media using an NMR-active isotope in the essential nitrogen or carbon source (most often $^{15}N$ $NH_3$ or $^{13}C$ glucose), then an excess of the amino acid that is to be delabeled is added (Jaipuria et al., 2012). To selectively label with certain amino acid, the protein is recombinantly expressed in natural isotope abundance carbon and nitrogen source, then supplemented with an isotopically-enriched carbon or nitrogen source. In practice, this generates a 2D HSQC spectrum or 3D spectrum that is either missing only signals from that particular residue or containing only resonances for selectively labeled residues. In this way, certain NH, or CH resonances can be unambiguously determined to be that particular amino acid residue type. With enough amino acid types known in a particular section of sequential resonances, that region can then be assigned to a particular portion of the protein's primary sequence.

Another NMR experiment that can provide further clues for NMR assignment is a $^{15}N$ or $^{13}C$-edited $^{1}H$-$^{1}H$ NOESY (Nuclear Overhauser Effect SpectroscopY) experiment (Marion et al., 1989; Zuiderweg and Fesik, 1989). This experiment uses the Nuclear Overhauser Effect to detect through space pairs of protons that are near each other. The effect is weak and falls off to the sixth power of the distance between the two atoms, becoming undetectable around 6 Å (Bax, 1989). The intensity of the NOE signal can be used to determine a more precise distance between the protons. For larger proteins, the 2D $^{1}H$-$^{1}H$ NOESY is far too crowded to interpret signals. Ergo, the




## Figure 2.1: Representative "Backbone Walk" Strip Plot

A series of "strips" generated from overlaid 3D HNCACB ($C\alpha_i$ and $C\alpha_{i-1}$, in red; $C\beta_i$ and $C\beta_{i-1}$, in blue) and 3D HNCOCA ($C\alpha_{i-1}$ only, in yellow) spectra, depicting the sequential connecitivities for Dbh residues 89 through 103. The $C\alpha_{i-1}$ peaks overlaid from the HNCOCA and HNCACB spectra indicate the value for the preceding residue, which is matched by the HNCACB $C\alpha_i$ peak in the preceding strip. For the Cβ peaks, the $C\beta_{i-1}$peak is much weaker than the $C\beta_{i-}$. The $^{13}C$ dimension is displayed along the *y*-axis, the $^{1}H$ dimension is along the *x*-axis (units are not displayed due to space constraints, and the depth in the third dimension, $^{15}N$, is indicated by the white number at the bottom of the graph. The white lines connecting peaks in adjacent strips indicate the connectivities between each residue.

experiment is “edited” through NH amide or the Ha proton, the N or C dimension is used as the third dimension to space out the signals and detect the NOE between the amide or alpha proton and any proton within roughly 6 Å (Marion et al., 1989; Zuiderweg and Fesik, 1989).

The most useful NOEs for backbone assignment of a protein are the regular patterns of backbone-to-backbone NOEs that occur in regions of secondary structure (Torchia et al., 1989). Alpha helices, beta-sheets, and beta-turns have unique NOE patterns that correspond to secondary structures (Englander and Wand, 1987). Therefore, when that particular pattern for a residue is discerned, the spectroscopist can be confident it is involved in that type of secondary structure. When the crystal structure is known, atoms that are not involved in that type of secondary structure can be eliminated as possible assignments. The positon of the backbone NOE off of the diagonal can also provide evidence whether atoms are sequential, as the chemical shift of the (N)H to (N)H signal must be in accordance with the sequential assignments in other 3D spectra. Many NOEs are detected from the protons in the residue’s own side chain, as well as those of adjacent and nearby side chains in the tertiary structure. Fortunately, side chain protons are predominantly aliphatic and less than 4 ppm, and do not obscure the strong signals between adjacent (N)H to (N)H or (Ca)H to (N)H atoms, which are greater than 4 ppm (Reid et al., 1997).

To assign the backbone resonances of Dbh, I first started by establishing the resonances which were unambiguously sequential in the 3D-HNCACB, HN(CO)CA, HNCO, and HN(CA)CO data, that is, there were no other possible connections to make between that segment of atoms. Then, I looked for segments more than one of the

residue types was known, and then matched that segment to a portion of the primary sequence. Next, I confirmed the assignment with the 3D-NOESY data. After I eliminated these sequences from the search, I was able to determine the identity of the remaining segments. I repeated this process until I assigned as many of the backbone atoms as possible.

**Materials and Methods**

*Protein Expression and Purification*

The Dbh gene was incorporated into the vector pKKT7-H (a derivative of pKK233, Promega) containing an N-terminal $His_6$ tag (MHHHHHHLVPRGM). Quick-change mutagenesis (Stratagene) was used to change Cys31 to Ser (hereafter referred to as C31S-Dbh) to eliminate potential formation of disulfide bonds. Transfected *E. coli* BL21 cells were grown in 1L Neidhart's minimal media (Neidhardt et al., 1974) at 37°C containing 1g $^{15}N$ ammonium chloride ($^{15}N$-labeled samples), or 1g $^{15}N$ ammonium chloride, 3g $^{13}C$ glucose, and 80% $D_2O$ ($^{2}H$, $^{15}N$, $^{13}C$-labeled samples) to ~1.0 OD; expression was induced by the addition of 1mM IPTG. Protein was expressed for 5 hours; subsequently, the cells were harvested by centrifugation and frozen at -80°C. Dbh or C31S-Dbh were purified from cell lysate by Ni-NTA affinity chromatography under native conditions, and then dialyzed into buffer (20mM HEPES, 100mM NaCl, 50μM EDTA, 50μM $NaN_3$, pH 7.5) at 4°C, then one change of buffer without EDTA (20mM HEPES, 100mM NaCl, 50μM $NaN_3$, pH 7.5). To prepare the NMR samples, Dbh or C31S-Dbh protein was concentrated to at least 0.5mM, and transferred into a Shigemi tube. $D_2O$ was added to the sample for a final concentration of 10% v/v. Since polymerase enzymes use aspartic acid side chains to coordinate $Mg^{2+}$ at the active site,

we were particularly interested in assigning the Asp groups. $^{15}$N-HSQC spectra of selectively $^{15}$N labeled (Asp) and un-labeled (Asn, Arg, Gly, Lys, His, Met, Ser) samples were used to confirm amino acid identity within the sequence. Selectively labeled $^{15}$N-Asp C31S-Dbh was prepared using the *E. coli* auxotroph strain EA1, which is unable to convert Asp to Asn (Muchmore et al., 1989), and by supplementing 1L of Neidhart's media with 100 mg $^{15}$N-Asp. Selectively unlabeled Asn, Arg, Gly, Lys, His, Met, Ser C31S-Dbh samples were prepared using 1g $^{15}$N ammonium chloride per L of Neidhart's minimal media, BL21 cells, and by supplementing with 0.5g of each $^{14}$N amino acid separately. Since the HNCO experiment on C31S-Dbh was by far the most sensitive, an HNCO spectrum of a selective $^{13}$C'-Leu, fully $^{15}$N enriched sample was used to confirm amide resonances preceded by leucine residues. $^{13}$C'-Leu was incorporated by expressing the protein in BL21 cells and adding 150 mg of $^{13}$C'-labeled Leu to 1 L of Neidhart's medium containing 1g $^{15}$N ammonium chloride.

*NMR Experiments*

NMR data were acquired at 35°C and 50°C on a Varian INOVA 800 MHz NMR spectrometer equipped with a 5-mm triple resonance *xyz*-gradient probe. The chemical shifts were referenced using 2,2-dimethyl-2-silapentane-sulfonic acid (DSS). All spectra were processed using NMRPipe/NMRDraw (Delaglio et al., 1995) and analyzed using CcpNmr Analysis (Vranken et al., 2005). A set of 3D triple resonance experiments, including HNCO, HN(CA)CO, HN(CO)CA, and HNCACB were carried out using TROSY (Pervushin et al., 1997) for the sequential backbone resonance assignment (Kay, 1997).

In addition, $^{15}$N-edited NOESY-HSQC spectra were also used to confirm resonance assignments.

**Results**

Figure 2.2 displays the $^{15}$N-HSQC of the full length C31S-Dbh protein (354 amino acids) with an N-terminal hexahistidine tag. Complete or partial backbone resonance assignments have been obtained for 86% (306 of 354) residues in C31S-Dbh, and 81% of amide resonances (276 of 339 non-proline residues). Twelve additional peaks were visible in the $^{15}$N-HSQC spectrum and the 3D spectra; I was unable to find connectivities for these resonances. In addition, four resonances in the $^{15}$N-HSQC did not have visible corresponding resonances in the 3D data or in the $^{15}$N-edited NOESY spectrum. Figure 2.3 presents the C31S-Dbh amino acid sequence with assigned residues indicated, and Figure 2.4 displays the assigned backbone portions of Dbh mapped onto the crystal structure (PDB entry 3BQ0 – Wilson and Pata). Residues 36-38 and the C-terminus (residues 345-354) were disordered in the crystal structures of Dbh [PDB entries 1K1S/1K1Q (Silvian et al., 2001) and 3BQ0 (Wilson and Pata, 2008)]; unfortunately, I was unable to assign the backbone resonances of the regions that were not visible in the crystal structures. If I omit the 13 residues (36-38; 345-354) that are disordered in the crystal structures of Dbh, I can account for 89% of residues with at least one assigned backbone resonance. Likewise, most of the linker region (residues 232-245) between the LF and thumb domains was not assigned. The linker and disordered regions from the crystal structure are flexible and solvent exposed. Hence, the signals may be missing due to intrinsic exchange with the solvent. In conclusion, for

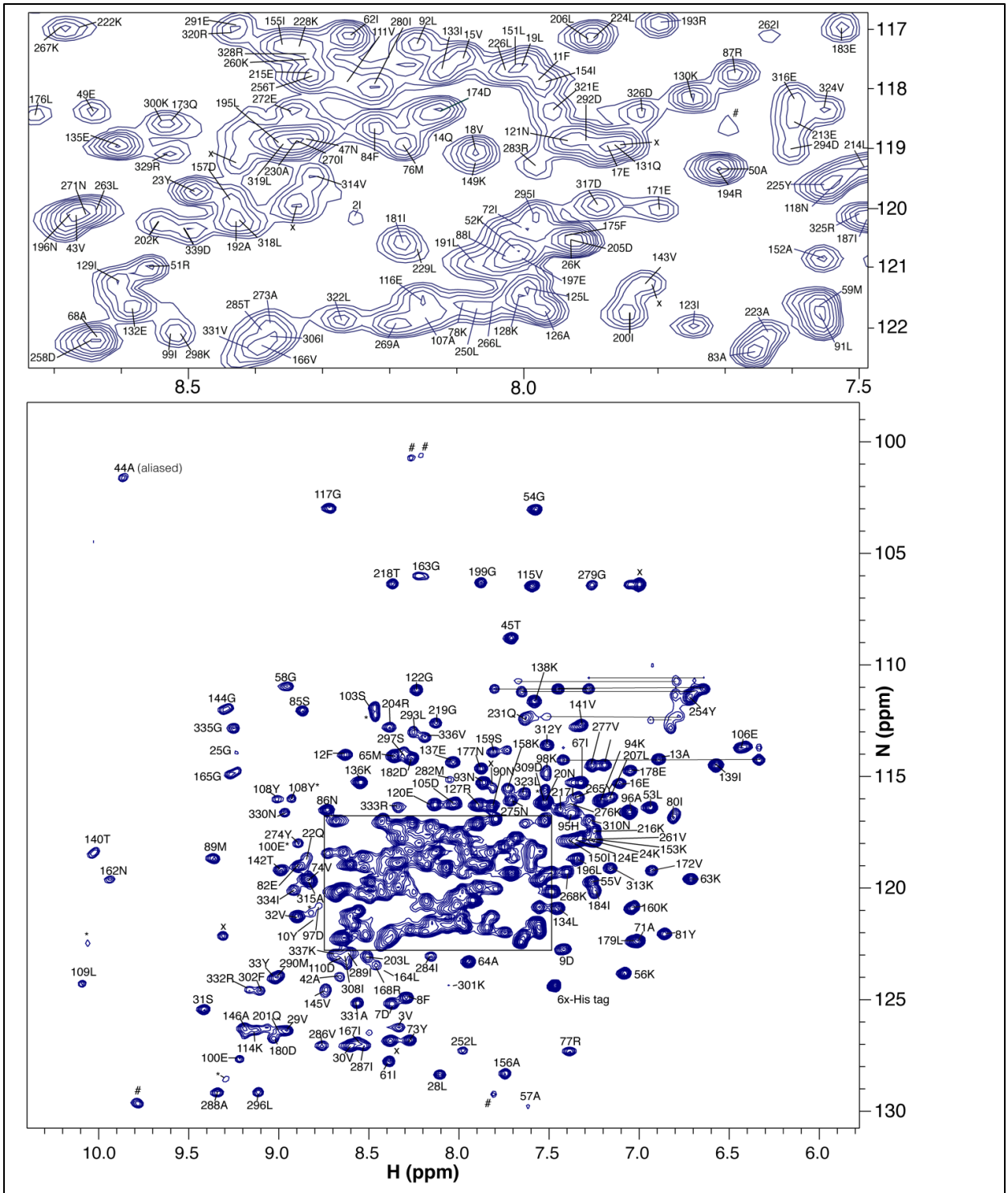



**Figure 2.2: $^{15}$N-$^{1}$H HSQC Map of Dbh assignments.** 2D $^{15}$N-HSQC spectra of C31S-Dbh, recorded at 800MHz and 50°C (323K). The one-letter amino acid code and the residue number indicate resonance assignments. Unassigned resonances are indicated by an "x" symbol. Additional conformations of a particular residue are indicated by an asterisk "*". Peaks in the HSQC which did not have visible corresponding peaks in the 3D spectra are marked by a hash "#". The crowded central region of the spectra is displayed in the insert for clarity. 44A is aliased in the $^{15}$N dimension; its true $^{15}$N shift is 133.701 ppm.

portions of C31S-Dbh where we would expect rigid structure to enable us to detect NMR signals, I have assigned 94% of the protein.

A representative strip plot of sequential residues is displayed in Figure 2.1, showing overlaid 3D HNCACB and 3D HN(CO)CA spectra for residues 89-103. By performing a "backbone walk" through the white lines indicated in the figure, I was able to determine these residues were sequential. From there, I determined the amino acid types. For example, 96A and 102A, with Cβ shifts of <24 ppm, and S103, with a Cβ shift of ~64ppm, are perfect examples of residues whose type can be determined from their $^{13}C$ chemical shifts. Other residue types were determined from selectively labeled and de-labeled spectra: 89M, 90N, 93N, 94K, H95H, 97D, and 98K. In addition, $^{13}C$'-Leu /$^{15}N$ HNCO data had CO*i-1* signals for 93N and 92L, indicating that these residues were preceded in the sequence by leucine residues. After establishing that these residues were sequential, and having determined the type of many of the amino acids, I concluded with confidence that this segment corresponded to residues 89-103.

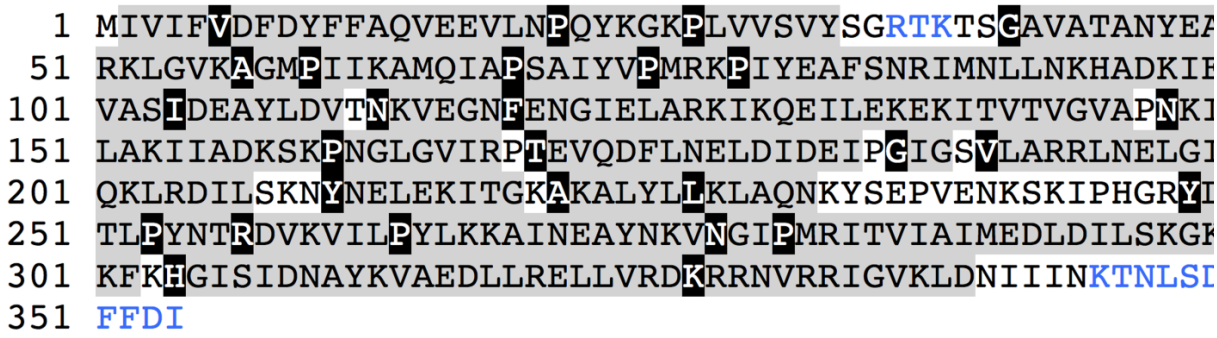



**Figure 2.3: Completeness of Assignment – Dbh Sequence.** C31S-Dbh amino acid sequence with assignments indicated. Grey shade is used to indicate residues with at least one backbone atom assigned. Black background with white lettering indicates residues where only the $^{13}C'$ or $^{13}C\alpha$ were found but not the NH. White background represents residues that have not been assigned. In blue lettering, amino acids 36-38 and 345-354 are too disordered in crystal structures to be detected [PDB entries 1K1S/1K1Q (Silvian et al., 2001) and 3BQ0 (Wilson and Pata, 2008)]. These regions are also not detected by NMR.

A small part of the structure appears to have an alternate conformation; several residues have two corresponding peaks in the $^{15}N$-HSQC spectrum and 3D data (97D, 98K, 100E, 101V, 102A, 103S, 108Y, 109L, and 110D) consistent with slow chemical exchange. These residues are located in the β-sheet structure of the active site palm domain, surrounding the metal ion coordinating residues 105D and 106E. The mutation of residue 31 from Cys to Ser does not appear to affect the structure. For example, the crystal structures of apo C31S-Dbh [PDB entries 1K1S/1KIQ (Silvian et al. 2001)] superimpose well with ligand-bound forms of WT Dbh [PDB entries 3BQ0, 3BQ1, and 3BQ2 (Wilson and Pata, 2008)]. In addition, this mutation did not significantly affect the NMR spectrum, as the $^{15}N$-HSQC of WT Dbh and C31S-Dbh overlay extremely well (data not shown). Only 11 peaks (9D, 10Y, 12F, 31S, 32V, 45T, 56K, 64A, 77R, 140T, and 301K) were observed to have shifted by any appreciable amount (>0.05ppm for $^{1}H$ or >0.2ppm for $^{15}N$) upon mutation. Unsurprisingly, the peaks for residue 31 and adjacent 32V are shifted, and all but two (140T, 301K) of the remaining shifted peaks are located in the same domain (finger) as residue 31.

The backbone resonance assignments of Dpo4 catalytic core (Ma et al., 2010) and LF domain (Ma et al., 2011) at 50°C have been published. Given the homology between

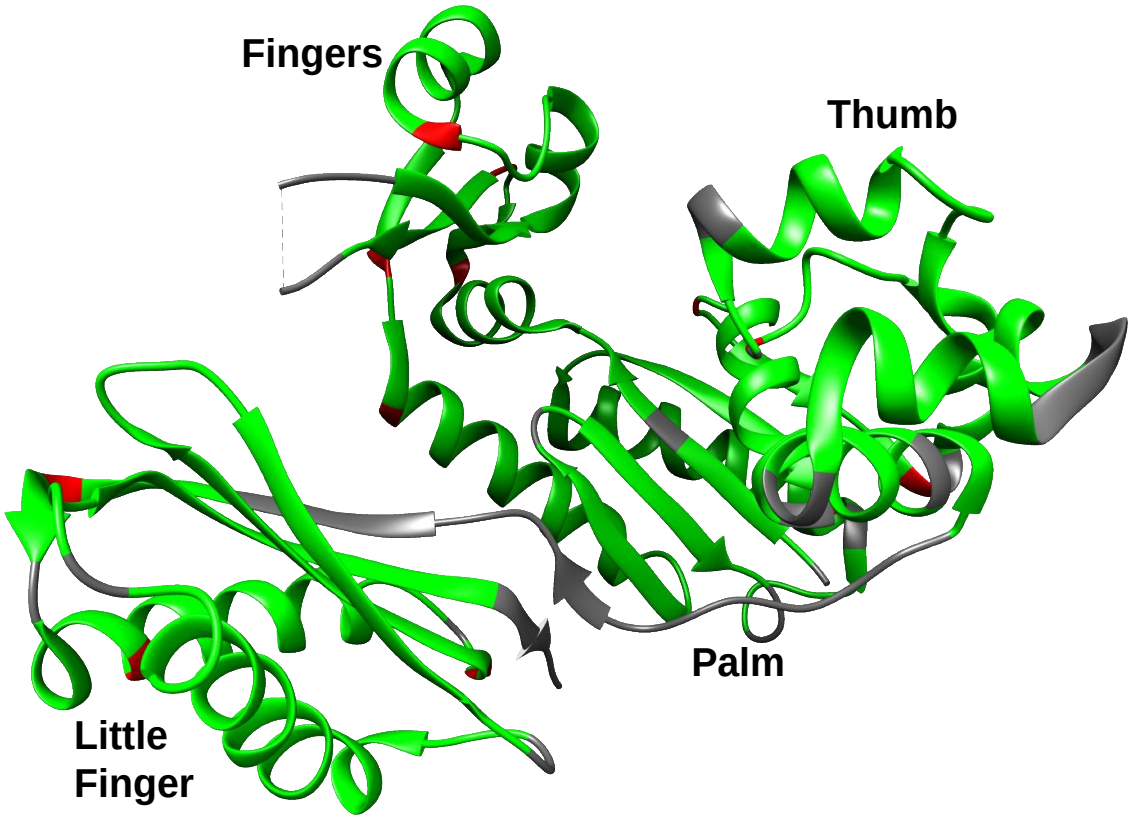



**Figure 2.4: Completeness of Assignment – Dbh Structure**. Completeness of Dbh assignments mapped onto the crystal structure of Dbh (PDB entry 3BQ0). Assigned residues are indicated in green, unassigned residues are indicated in gray, and proline residues are indicated in red.

the two proteins (54% sequence identity) and similar tertiary structure, some of the resonances of the two proteins would be expected to have similar chemical shifts. It should be noted that all of the assignments of C31S-Dbh were completed independently using only my own data; the Dpo4 assignments were compared to those of C31S-Dbh after I completed my assignments. The mutually assigned amide peak positions of

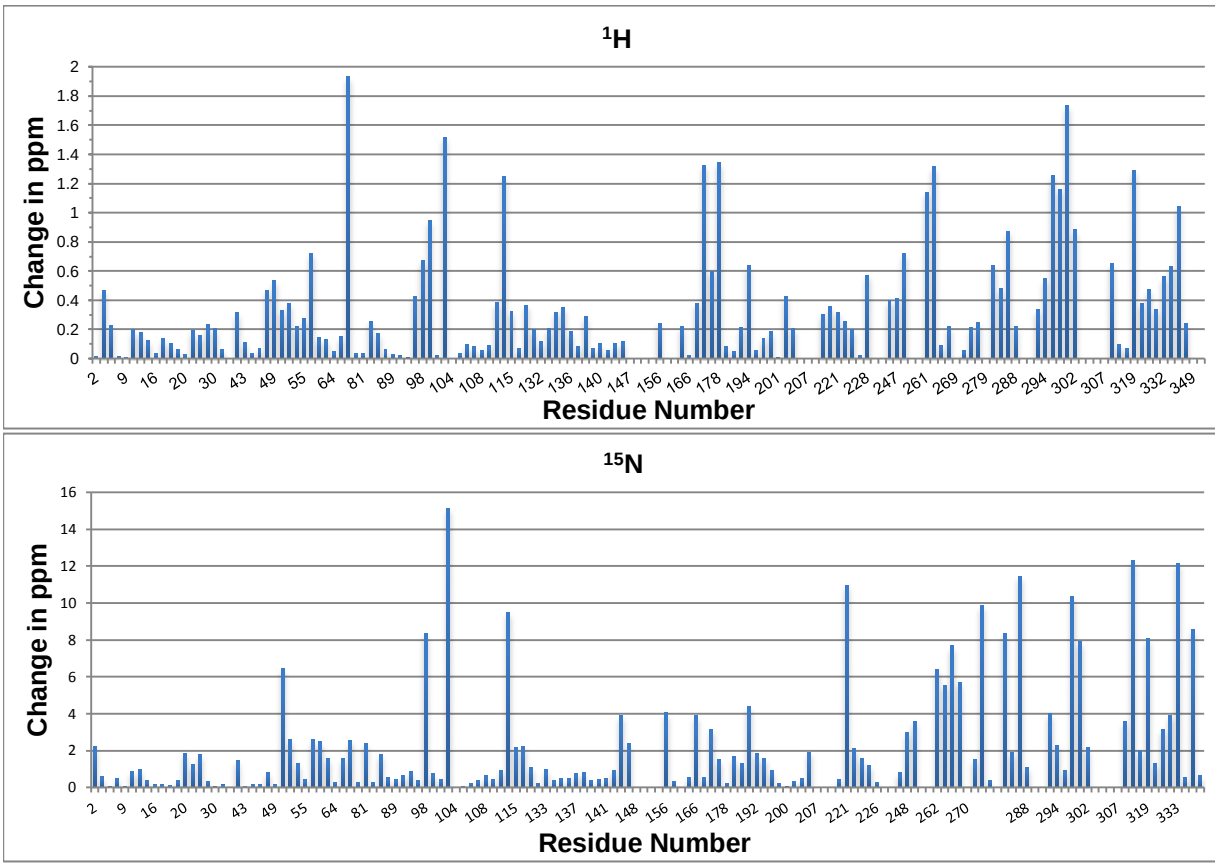



**Figure 2.5: Comparison of Dbh and Dpo4 chemical shifts for identical residues.** Comparison of $^{1}$H and $^{15}$N chemical shifts from identical residues from Dbh and Dpo4 is displayed as the difference of the chemical shift values.

identical residues do not correlate very well, with only 62 of 135 (46%) available $^{1}$H shifts within 0.2 ppm and 54 of 135 (40%) available $^{15}$N shifts within 0.8 ppm (Figure 2.5). However, the nearest neighbors of a particular residue can have a significant effect on the amide chemical shift, even if the residues are identical. For instance, a neighboring isoleucine residue on the C-terminal side of an amide could influence the $^{1}$H shift downfield by ~0.2 ppm (Schwarzinger et al., 2001) and the $^{15}$N shift downfield by almost 5 ppm (Braun et al., 1994; Schwarzinger et al., 2001). Eliminating identical residues from the comparison that do not also have identical neighbors, the correlation between the two sets of shifts is improved: 30 of 54 (56%) $^{1}$H shifts within 0.2 ppm and

(57%) 31 of 54 $^{15}$N shifts within 0.8ppm. All but one of the $^{1}$H shifts and one of the $^{15}$N shifts in the preceding comparison are found in the polymerase core: 29 of 45 (64%) $^{1}$H shifts within 0.2 ppm and 30 of 45 (67%) $^{15}$N shifts within 0.8ppm. This is not surprising, since the polymerase core between the two proteins shares 59% sequence identity, while the LF domains of the two proteins only have 41% sequence identity. Even though the assignments of Dpo4 were completed on the polymerase core and the LF as separate constructs, the core and LF domains appear to fold independently into roughly the same native structure as found in the full-length protein based on chemical shifts.

**Conclusions**

I have obtained complete or partial backbone resonance assignments for 86% of residues in Dbh. Accounting for regions where NMR signals may not have been detected, I have assigned backbone resonances from 94% of residues in Dbh. A comparison of the Dbh chemical shifts with published Dpo4 chemical shifts revealed moderate correlation between shifts of identical residues. The backbone resonance assignments have been published in the journal *Biomolecular NMR Assignments* (Moro and Cocco, 2015). The chemical shifts of C31S-Dbh polymerase at 308K and 323K (35°C and 50°C) have been deposited in the BioMagResBank database under accession number 26564 [http://www.bmrb.wisc.edu, (Ulrich et al., 2008)]. The backbone assignments have been used for relaxation dynamics studies of Dbh, and the published crystal structures of Dbh have been used to interpret NMR relaxation data through molecular modeling and molecular dynamics simulations

## References:


Bax, A. (1989). Two-dimensional NMR and protein structure. Annu Rev Biochem 58, 223-256.

Bodenhausen, G., and Ruben, D.J. (1980). Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. Chemical Physics Letters 69, 185-189.

Braun, D., Wider, G., and Wüthrich, K. (1994). Sequence-Corrected 15N "Random Coil" Chemical Shifts. J Am Chem Soc 116, 8466-8469.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Englander, S.W., and Wand, A.J. (1987). Main-chain-directed strategy for the assignment of 1H NMR spectra of proteins. Biochemistry 26, 5953-5958.

Frueh, D.P. (2014). Practical aspects of NMR signal assignment in larger and challenging proteins. Prog Nucl Magn Reson Spectrosc 78, 47-75.

Ikura, M., Kay, L.E., and Bax, A. (1990). A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry 29, 4659-4667.

Jaipuria, G., Krishnarjuna, B., Mondal, S., Dubey, A., and Atreya, H.S. (2012). Amino acid selective labeling and unlabeling for protein resonance assignments. Adv Exp Med Biol 992, 95-118.

Kay, L.E. (1997). NMR methods for the study of protein structure and dynamics. Biochem Cell Biol 75, 1-15.

Leopold, M.F., Urbauer, J.L., and Wand, A.J. (1994). Resonance assignment strategies for the analysis of NMR spectra of proteins. Mol Biotechnol 2, 61-93.

Ma, D., Fowler, J.D., and Suo, Z. (2011). Backbone assignment of the little finger domain of a Y-family DNA polymerase. Biomol NMR Assign 5, 195-198.

Ma, D., Fowler, J.D., Yuan, C., and Suo, Z. (2010). Backbone assignment of the catalytic core of a Y-family DNA polymerase. Biomol NMR Assign 4, 207-209.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. (1989). Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. Biochemistry 28, 6150-6156.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E., and Dahlquist, F.W. (1989). Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol 177, 44-73.

Neidhardt, F.C., Bloch, P.L., and Smith, D.F. (1974). Culture medium for enterobacteria. J Bacteriol 119, 736-747.

Pervushin, K. (2001). The use of TROSY for detection and suppression of conformational exchange NMR line broadening in biological macromolecules. J Biomol NMR 20, 275-285.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94, 12366-12371.

Reid, D.G., MacLachlan, L.K., Edwards, A.J., Hubbard, J.A., and Sweeney, P.J. (1997). Introduction to the NMR of proteins. Methods Mol Biol 60, 1-28.

Schwarzinger, S., Kroon, G.J., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. In J Am Chem Soc (United States), pp. 2970-2978.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8, 984-989.

Torchia, D.A., Sparks, S.W., and Bax, A. (1989). Staphylococcal nuclease: sequential assignments and solution structure. Biochemistry 28, 5509-5524.

Ulrich, E.L., Akutsu, H., Doreleijers, J.F., Harano, Y., Ioannidis, Y.E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z.*, et al.* (2008). BioMagResBank. Nucleic Acids Res 36, D402-408.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59, 687-696.

Wilson, R.C., and Pata, J.D. (2008). Structural insights into the generation of single-base deletions by the Y family DNA polymerase dbh. Mol Cell 29, 767-779.

Zuiderweg, E.R., and Fesik, S.W. (1989). Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. Biochemistry 28, 2387-2391.

## Chapter 3 – Dynamics of Dbh investigated by Hydrogen-Deuterium Exchange (HDX) NMR spectroscopy

### Rationale and Strategy

Hydrogen-deuterium exchange (HDX) NMR spectroscopy is a technique used to measure the dynamics, stability, and folding of proteins (Krishna et al., 2004). The amide hydrogen of polypeptide chains can undergo exchange with the hydrogens of bulk water. The utility of this exchange in probing the dynamics of protein structure has long been recognized; if the protein is lyophilized, then re-suspended in $D_2O$, the protons will exchange with deuterons at a certain rate for given solvent conditions (Hvidt and Linderstrøm-Lang, 1955). For an amide exposed to solution, this exchange is fast, usually on the order of seconds. The exchange can be detected as a gain in mass (HDX-MS) or as a decay in signal (HDX-NMR). The amide protons that are participating in secondary structure or otherwise shielded from the solvent will undergo exchange more slowly than if they were freely exposed to solvent. This reduction in the rate of exchange is referred to as the “protection factor” (Raschke and Marqusee, 1998).

The protection factor reflects the local structure and dynamics of a polypeptide. If a local region of a polypeptide is involved in secondary structure or otherwise shielded from the aqueous solvent, the structure will need to undergo an "opening' event in order for exchange of the amide proton to occur (Raschke and Marqusee, 1998). This opening is dependent on the degree of order in the local structure. Dynamic regions have a higher probability that an opening event will occur; conversely, more rigid and stable structures have a low probability of opening (Maity et al., 2003). For example, in the interior of β-

sheets in stably folded protein domains, the amide hydrogens may not exchange for many days. However, if the exchange is extremely slow, it may not be expedient to run the experiment for the necessary time to observe sufficient signal decay to extract a rate measurement. The rate of opening can be determined from the rate of exchange when normalized for the intrinsic exchange rate of the random coil polypeptide (Bai et al., 1995).

After the opening event occurs, the reverse process is the rate of closing, whereby the structure reforms and the amide hydrogen is again protected from exchange with solvent deuterium. The actual probability of exchange occurring is dependent on the actual time spent in the open state (Bai et al., 1995). Hence, the equilibrium constant is determined by the rate of closing ($k_{cl}$) divided by the rate of opening ($k_{op}$), as defined in Eq. 1 (Hvidt and Linderstrøm-Lang, 1955), assuming that the system is under what is called the EX2 exchange limit defined where $k_{cl} \gg k_{op}$.

$$(\mathbf{1})\ \boldsymbol{k_{ex} = (k_{op}/k_{cl})k_{rc} = K_{op}k_{rc}}$$

For the determination of the observed hydrogen-deuterium exchange rate, a series of 2D $^{1}$H, $^{15}$N HSQC spectra are taken at regular periodic intervals. Next, the signal decay for each amide is measured over time and a single-order exponential decay is used to extract the rate. Unfortunately, during the dead time of a conventional HDX-NMR experiment, many solvent-exposed amide hydrogens exchange very rapidly. Hence, these signals are not observed at the first time point and an exact rate cannot be calculated, although an upper bound for the rate can be established. Likewise, well-protected amide hydrogens may take weeks to fully exchange, and hence significant

decay may not occur during the time course of the experiment. As such, only a lower bound can be determined.

To determine the protection factor $(PF = k_{ex}/k_{rc})$, the exchange of the intrinsic exchange of the amide hydrogen – that is, if it was in a random coil conformation – must first be calculated. To calculate the intrinsic exchange rate of any particular amide hydrogen, the rate must be corrected for pH, temperature, and nearest neighbor effects. Englander and co-workers performed an extensive study for the effect of each nearest neighbor to provide an approximate correction factor on the exchange rate; these rates are sufficient to calculate the protection factor (Molday et al., 1972; Bai et al., 1993). The calculation of the intrinsic exchange rate will be explained in greater detail in the methods section of this chapter.

In this way, hydrogen-deuterium is a sensitive technique for the backbone dynamics of proteins from seconds to days. Magnetization transfer of solvent saturation can be used to monitor very rapidly exchange amides to access sub-second time scales (Krishna et al., 1979). Conventional NMR-HDX can be used to observe the sample for as long as the spectroscopist deems necessary and informative. The time range available for investigation nicely complements other experiments which probe millisecond and faster motions, such as nuclear spin relaxation and relaxation dispersion measurements.

Provided that the system is in the EX2 exchange limit, thermodynamic parameters can be extracted from the exchange constants (Eq. 2). Within the EX2 limit, the free energy of exchange is equivalent to the equilibrium constant of structural opening (Bai et al., 1994; Bai et al., 1995).

$$\textbf{(2)}\ \Delta G_{HX} = -RTlnK_{op} = -RTln\, k_{ex}/k_{rc}$$

Since $k_{rc}$ can be calculated, the free energy can be calculated from the measured rate constant. The temperature dependence below the thermal melting point of the free energy can give the entropy ($d\Delta G_{HX}/dT = \Delta S_{HX}$) and hence, the enthalpy of the exchange reaction. One example where this approach has been applied was the study of thermodynamic stability of hyperthermophilic rubredoxin (Hiller et al., 1997; LeMaster et al., 2005). Of particular interest in the case of Dbh is the thermodynamic basis of its stability. A higher exchange rate at a larger temperature is expected, but the high stability of the core of a thermostable protein has been demonstrated in the following data. The portions of the protein involved in the core interactions of each domain, especially in the central β-sheet/α-helix bundle of the palm and little finger domains, are remarkably stable.

I have obtained HDX measurements at 35°C and 50°C, to investigate the backbone dynamics of Dbh. From this information, I can see the modulation of dynamics in Dbh upon heating. Core residues in the palm and little finger domain, especially those in the interior of the β-sheets and interior face of α-helices, are key in maintaining the stability of Dbh; with the palm providing a stable platform for the active site residues. In addition, the thumb and finger domains are more flexible, with only a handful of residues protected at 35°C and almost none at 50°C. These results highlight the importance of maintaining the active site structure and shape of LF domain for the function of Dbh.

## Materials and Methods

Samples of 100% $^{15}$N-labeled Dbh containing 0.5mM or higher concentration were lyophilized, then re-suspended in 300µL of $D_2O$, pD 7.5, which was pre-warmed to the experimental temperature, and transferred to a Shigemi NMR tube. The sample was immediately placed in an 800MHz Varian Inova NMR spectrometer, containing an *xyz* triple resonance probe, equilibrated at 35 or 50°C. After shimming and tuning the magnet, the acquisition of the first $^{15}$N-$^{1}$H TROSY-HSQC (Pervushin et al., 1997) spectrum was started approximately 15 minutes after the re-suspension of lyophilized protein. Additional $^{15}$N-$^{1}$H HSQC spectra were taken sequentially every 147 minutes and 13 seconds (2.454 hours), except for the last five spectra at 50°C, for which additional scans were taken (2x for the 17$^{th}$ through 20$^{th}$ spectra, 4x for the 21$^{st}$ spectrum) to improve signal-to-noise ratio. A total of 20 spectra were collected at 35°C, and 21 spectra at 50°C. Due to the additional length of acquisition in the last five spectra at 50°C, the acquisition of the 18$^{th}$, 19$^{th}$, 20$^{th}$ and 21$^{st}$ spectra was started 4.907 hours after the start of the previous spectra. Representative spectra for various time points at both temperatures are shown in Figure 3.1a and b. The data was processed using NMRPipe (Delaglio et al., 1995), and visualized using CcPNMR Analysis (Vranken et al., 2005). The peak intensity was plotted as a function of time, and fit to a single order exponential-decay function ($I(t) = I_0 \times e^{-kt}$) to extract the exchange rate. Representative rate fits to the experimental data are displayed in Figure 3.2. To obtain the protection factor the hypothetical exchange rate for the amide proton in a random coil conformation was calculated, corrected for the effect of side chain identity to the left and right of the amide proton (Molday et al., 1972; Bai et al., 1993).

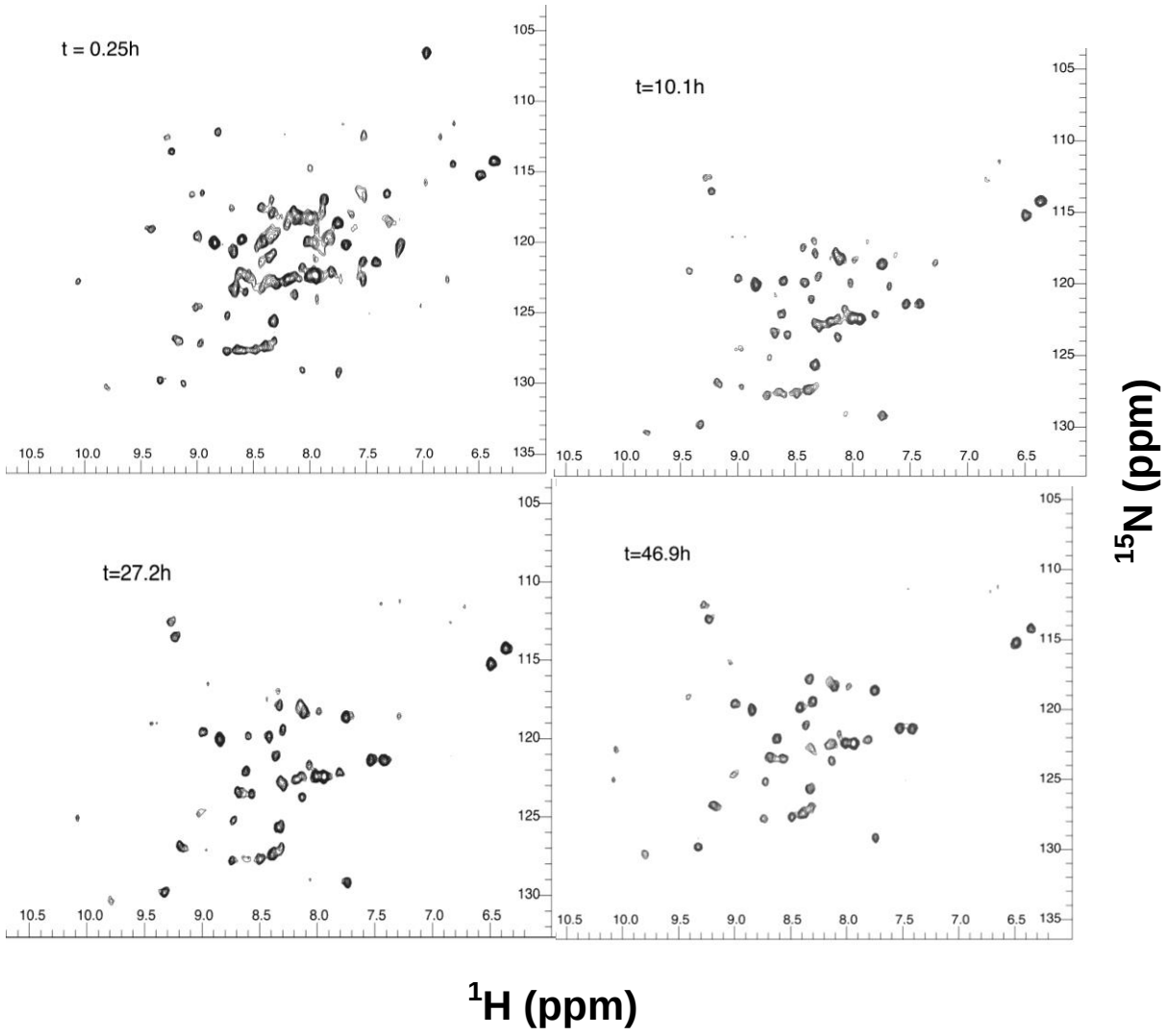



**Figure 3.1a:** $^{15}$N-$^{1}$H TROSY-HSQC HDX spectra of Dbh at 35°C at selected time points, demonstrating decay of signal over the course of the experiment.

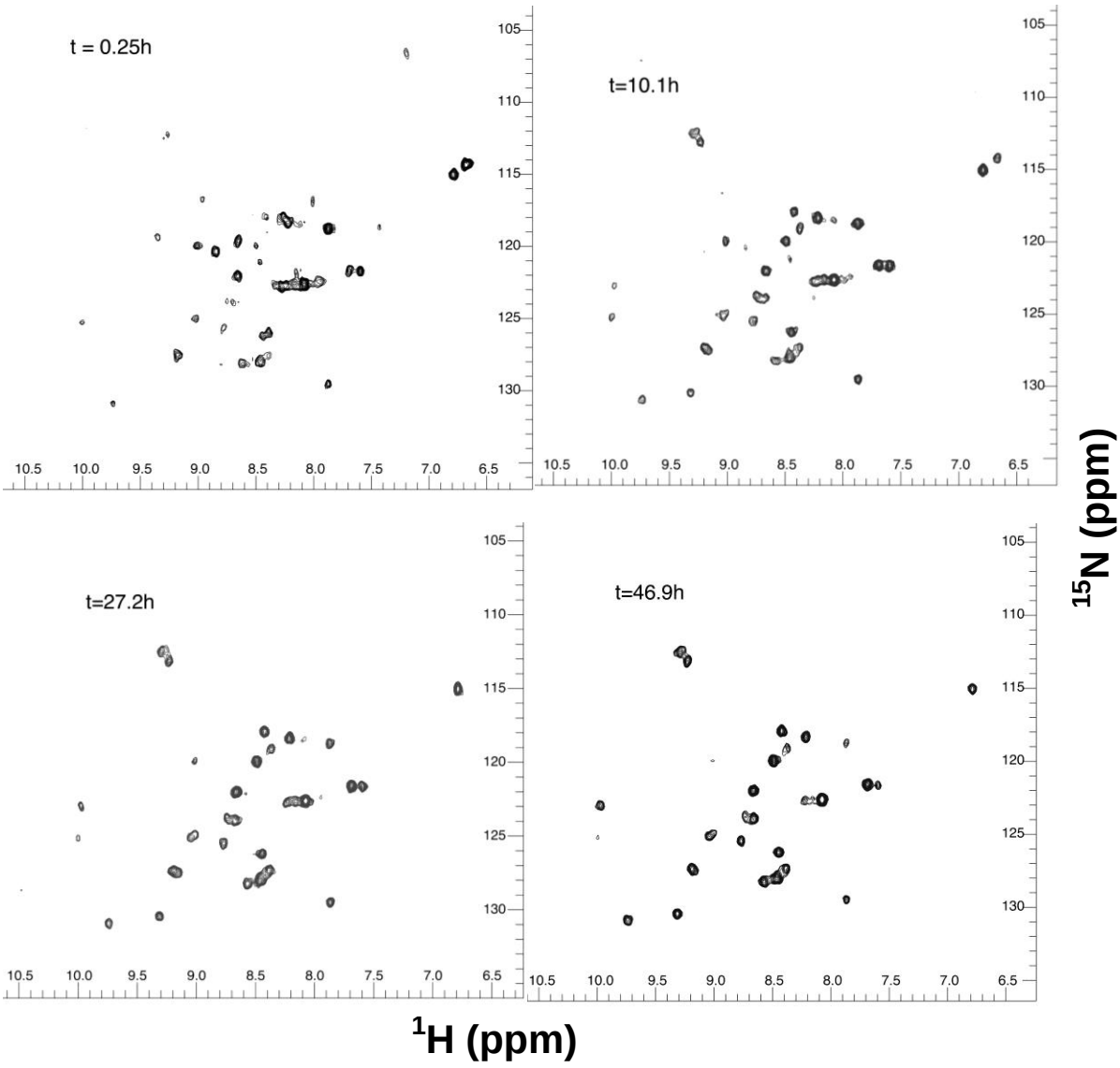



**Figure 3.1b:** $^{15}$N-$^{1}$H TROSY-HSQC HDX spectra of Dbh at 50°C at selected time points, demonstrating decay of signal over the course of the experiment.

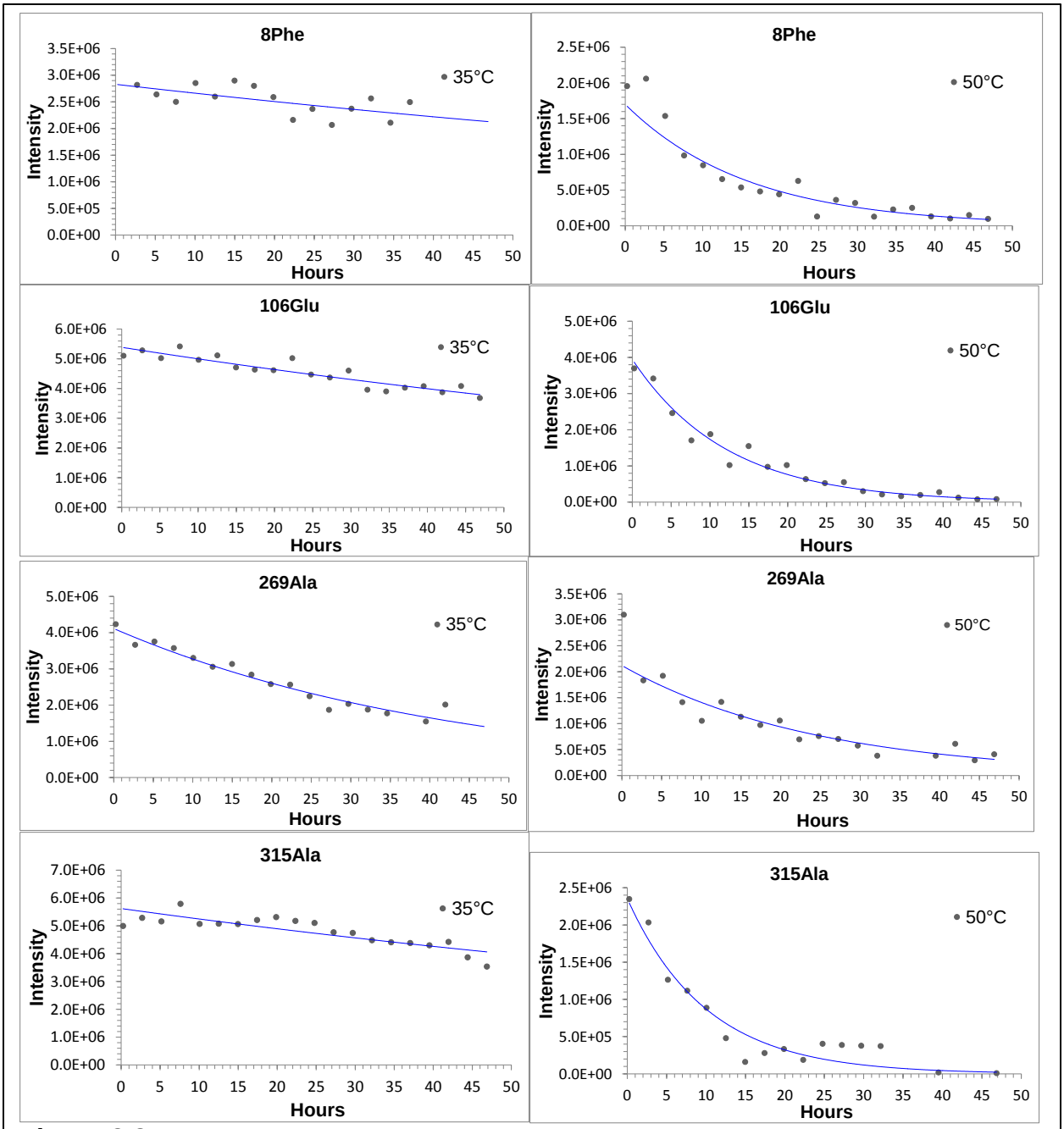



**Figure 3.2:** Representative decay rate fits for selected amides from HSQC-HDX spectra for Dbh at 35°C and 50°C. The rates were determined from the following equation: $I(t) = I_0 \times e^{-kt}$.

The protection factor was calculated using Eq. 2, as the ratio of the observed rate and the hypothetical random coil rate for the amide proton (Bai et al., 1995).

$$(3)\boldsymbol{P} = \boldsymbol{K_{rc}}/\boldsymbol{K_{prot}}$$

**Calculation of Random Coil Rate and Free Energy**

The intrinsic H-D exchange rates for random coil polypeptides are known, and can be calculated for any amide hydrogen for a given pD, temperature, and local sequence context (Molday et al., 1972; Bai et al., 1993). The exchange can be expressed as a sum of the rates of the water, base, and acid catalyzed reactions, as given in Eq. 3. Each of these rates can be calculated individually for a given pD and every possible sequence context, by using a reference rate for the reaction at 20°C, and correcting for effects of the side chains to the immediate left and right (Eqs. 4 and 5a, 5b, and 5c) (Bai et al., 1993). The side chain to the "left" is the residue's own side chain, and the side chain to the right, as the preceding residue's side chain, which is displayed for clarity in Figure 3.5. Eq. 6 correct for the increase in rate due to the temperature, with 14, 17, and 19 kcal/mol used as the values for the activation energy of the acid, base, and water catalyzed reactions, respectively. The increase in $k_{rc}$ from 20°C to 35°C is 3.23x, 4.14x, and 4.89x for the acid, base, and water-catalyzed reactions, respectively. The increase in $k_{rc}$ from 20°C to 50°C is 9.88x, 16.1x, and 22.4x for the acid, base, and water-catalyzed reactions, respectively. The calculated random coil rates were divided by the experimental rates to obtain the protection factors.

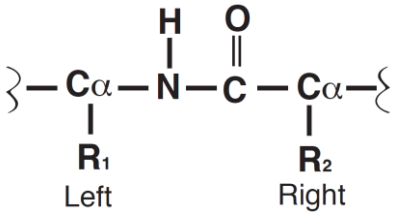



**Figure 3.3: L vs. R for $k_{rc}$ calculation**

$$.(4) k_{ex} = k_A 10^{-pD} + k_B 10^{[pD-pK_D]} + k_W$$

$$(5)\ k_{rc} = k(\text{acid}) + k(\text{base}) + k(\text{water}) = k_{A,ref}(A_L \times A_R)[D^+] + k_{B,ref}(B_L \times B_R)[OD^-] + k_{W,ref}(B_L \times B_R)$$

$$(6a) \log k(\text{acid}) = \log k_{A,ref} + \log A_L + \log A_R - pD$$

$$(6b) \log k(\text{base}) = \log k_{B,ref} + \log A_L + \log A_R - pOD$$

$$(6c) \log k(\text{water}) = \log k_{W,ref} + \log A_L + \log A_R$$

$$(7) k_{rc}(T) = k_{rc}(293) e^{\left(\frac{-E_a[1/T - 1/293]}{R}\right)}$$

The free energy of the hydrogen exchange reaction, measuring the equilibrium of local structural opening, was calculated according to Eq. 2.

**Results and Discussion**

$^{15}$N-$^{1}$H HSQC spectra of selected time points for each temperature are displayed in Figure 3.5. Signals for 95 residues were resolved at 35°C, and for 50 residues at 50°C. Of these, sufficient decay was observed to enable calculation of a rate for 61 residues at 35°C and 27 residues at 50°C. There were a few signals which could not be resolved due to spectral overlap. Many of the residues in Dbh undergo exchange quickly enough at 35°C and 50°C so that no signal is detected in the first spectra. As expected, residues that are solvent-exposed or not involved in secondary structure undergo rapid exchange. The lower limit for the rate of fast-exchanging residues was calculated as >0.2 min$^{-1}$, for which the signal to decay to 1/20$^{th}$ of its original value, the ratio, by the start of acquisition of the first spectra 15 minutes. For residues which exchange fully by the start of the second time point (159.2 minutes), the exchange is approximately

between 0.018 $min^{-1}$ and 0.2 $min^{-1}$. Since the base-catalyzed reaction dominates at pD 7.5, the most base-activating dipeptide present in Dbh (Asn-NH-*cis*-Pro) defines the maximum protection factor possible for the fast exchanging residues (Bai et al., 1993) for the lower limit of the exchange rate. The $k_{rc}$ for an Asn-NH-Ser dipeptide is $3.3 \times 10^4$ $min^{-1}$ at 35°C and $1.3 \times 10^5$ $min^{-1}$ at 50°C, giving a maximum possible protection factor of $1.7 \times 10^5$ at 35°C and $6.4 \times 10^5$ for 50°C. For exchange by the second time point, the maximum possible protection factor is $1.8 \times 10^6$ at 35°C, and $7.1 \times 10^6$. However, most dipeptide combinations in Dbh will have a much slower $k_{rc}$, and hence lower maximum protection factors, which are mostly in the range of $10^3$-$10^4$ at 35°C and $10^3$-$10^6$ at 50°C. Protection factors for each detected amide group, along with estimated lower (highly protected residues) and upper (fast exchanging residues) limits for assigned residues whose rates could not be determined at that temperature, are displayed in Figure 3.4. These protection factors have also been mapped to the crystal structure of apo Dbh [1K1S, (Silvian et al., 2001)] in Figure 3.5. For the residues for which no decay was detected, assuming that 5% or less of initial peak intensity decayed by the end of the experiment, the minimum protection factors for possible for the most blocking dipeptide (Ile-NH-Ile) $>1.7 \times 10^7$ at 35°C, and $>9.2 \times 10^7$ at 50°C.

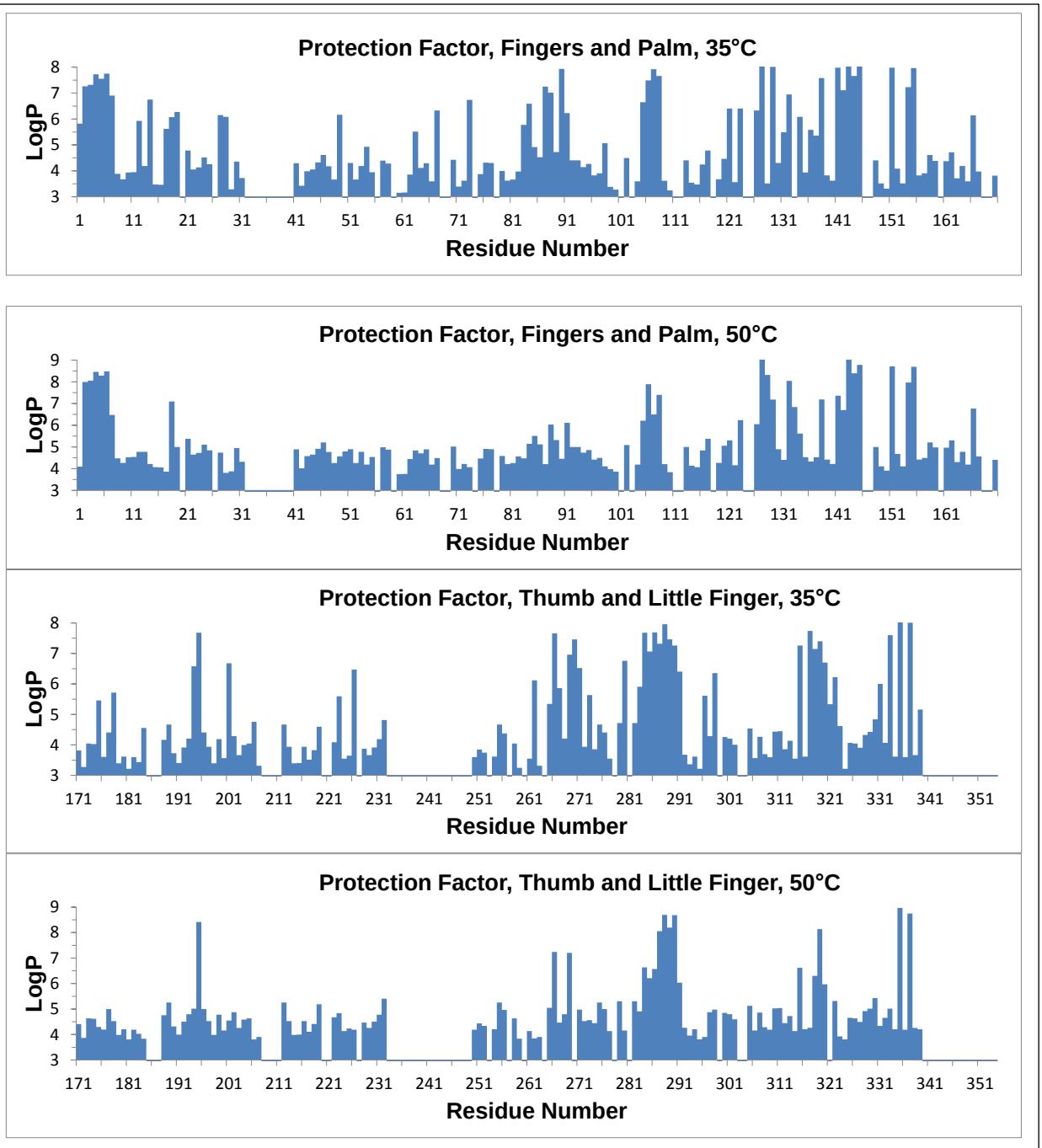



**Figure 3.4**: Protection factors for Dbh at 35°C and 50°C. Residues for which a value has been calculated, an upper limit determined (exchange before first spectrum), or a lower limit determined (no decay detected by the end of experiment, ~69h at 50°C). Unassigned residues or residues that could not be resolved due to spectral overlap are not given a value.

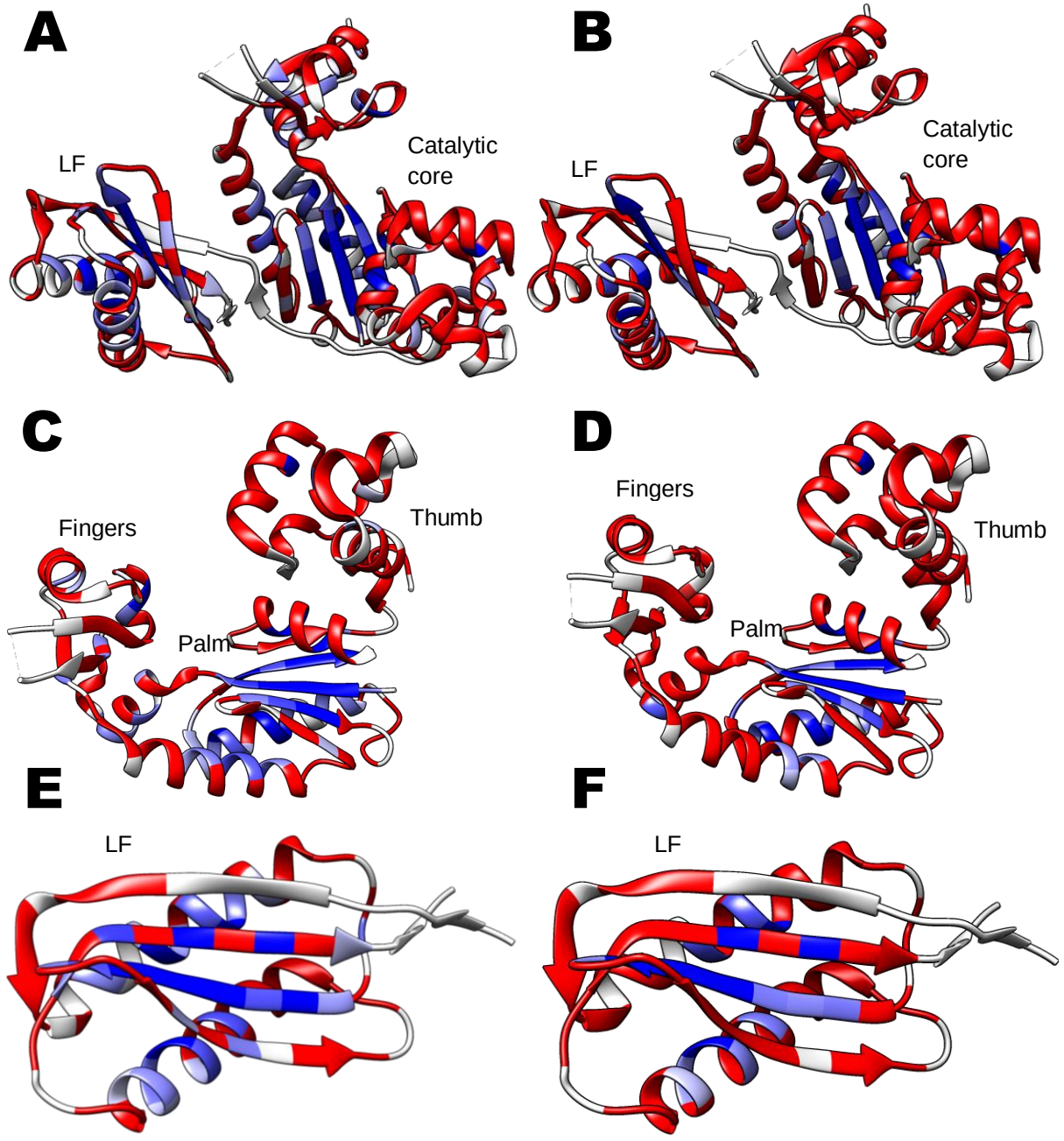



**Figure 3.5:** Hydrogen-deuterium exchange of Dbh at 35°C (left-side images A, C, and E) and 50°C (right-side images B, D, and F). Panels A and B show a view of the whole protein, C and D show the catalytic core (residues 1-232) and E and F show the LF domain (residues 240-344). Protection from exchange is indicated in blue, with darker shades indicating greater protection fast exchange is indicated in red, and white indicating residues that are unresolved due to spectral overlap, or unassigned. PDB entry 1K1S (Silvian et al., 2001) and UCSF Chimera (Pettersen et al., 2004) were used to generate the images.

Most of the protected residues in Dbh are located in the palm and little finger domains (78 of 95 at 35°C, and 49 of 50 at 50°C), especially residues on the interior of the β-sheets and the face of the α-helices packed against the central β-sheets (Figure 3.5).

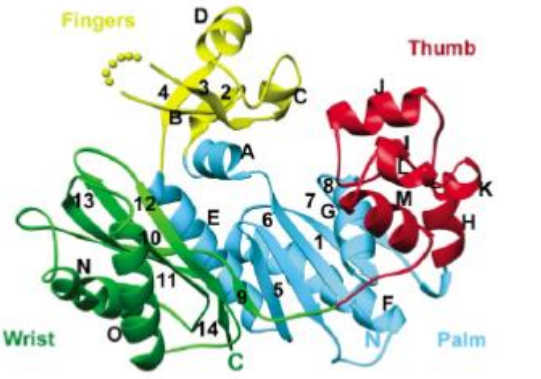



**Figure 3.6:** Numbering of secondary structure elements in apo Dbh. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, Silvian et. al, © 2001

Twenty of these residues in the interior of β-sheets and on the inward facing side of α-helices are so well protected from exchange that no discernible decay is even at 50°C (3V,7D,128K,129I, 133I, 144G, 145V, 146A, 152A, 155I, 156A, 195L, 270I, 287I, 288A, 289I, 290M, 319L, 335G, and 337K), with protection factors $\geq 10^8$ (assuming 5% or less decay of the original intensity has occurred by the end of experiment at 69 hours, which gives a minimum $k_{ex}$ of 1.24 x $10^{-5}$). The calculated $\Delta G_{HX}$ for the residues in the palm and LF whose rates could be measured, averaged across both temperatures, is 9.24 ±2.33 kcal/mol and 8.96±2.93 kcal/mol, respectively. The average minimum $\Delta G_{HX}$ for the stable residues at 50°C, assuming 5% or less decay occurred, is ≥12.5kcal/mol. This indicates the cores of the palm and little finger domains are particularly rigid and stable.

However, the smaller thumb and finger domains have fewer (17) residues that are protected from exchange at 35°C (28L, 30Vl, 50A, 51R, 55V, 64A, 68A, 74V, 175F, 178E, 184I, 194R, 195L, 201Q, 206I, 223A, 226L), and only one (195L, located in an α-helix and buried in the interior of the thumb domain) at 50°C. The α-helices in the thumb and fingers domains are shorter than those of the palm and LF domains, and are

frequently broken by proline residues. Proline is the most rigid amino acid and can confer rigidity, but experiments examining the effects of addition of proline residues demonstrate that they do not confer stability in all cases (Wang et al., 2014; Yu et al., 2015). These prolines (21P, 27P, 60P, 69P and 75P in the fingers, and 185P in the thumb), likely constrain the .However, *E. coli* pol IV, a mesophilic homolog of Dbh, also has proline residues in roughly similar locations, so this may only to restrict the volume of the domain while maintaining a stable fold. The smaller volume of thumb and fingers domains of Dbh and other Y-family polymerases increases the volume of their active sites, enabling them to accommodate bulky lesions. Therefore, the increased hydrogen exchange observed for the palm and fingers domains is likely due to their reduced volume.

In the palm, metal-ion coordinating residues 106E and 7D are protected from exchange, along with surrounding residues 8F, 107A, 108Y, 6V, and 5F. The β-sheet forms a rigid platform for the active site of the enzyme, ensuring that these residues remain in a proper orientation for catalysis. The third metal-ion coordinating residue 105D is located in a β-turn, and hence exchanges before the first time point. 103S NH is on the edge β-strand β5, right before the β-turn containing 104I and 105D. 103S NH forms a hydrogen bond with 106E CO and exchanges quickly even at 35°C. This may indicate structural fraying at the end of the edge strand and consequent propagation of motion through the β-turn. Motion of the backbone in this location could subtly affect the positioning of the metal-ion coordinating side chains. The peaks for 103S, 108Y, 109L, 110D, 97D, and 98K are doubled in 2D and 3D spectra of Dbh (Moro and Cocco, 2015), indicating slow exchange between two stable conformations. Subtle repositioning of active site residues

is thought to be the rate-limiting step for catalysis; however, any possible motion of these residues would also have to be observed for the ligand bound states of Dbh.

Other thermostable proteins have rigid central β-sheet architecture which is resistant to exchange, as was demonstrated recently for a thermostable chimeric avidin (Tossavainen et al., 2014). Increased β-sheet structure often confers additional thermostability, provided that the hydrophobic interactions in the core of the protein can be increased (Yang et al., 2012; Niu et al., 2015). Many of the slowly exchanging residues in Dbh are branched-chain hydrophobic amino acids (I, L, and V), clusters of which can confer increased thermostability to a protein (Gangadhara et al., 2013). Comparing the amino acid composition of Dbh to mesophilic homolog *E. coli* pol IV to Dbh, Dbh has 17 more isoleucine, 4 less leucine, and 4 more valine residues. In particular, certain clusters of hydrophobic residues in the interior of the palm and LF are very resistant to exchange. In the palm, 129I, 133I, 6V, 4I, 109L, 145V, 143V, and 141V, all of whose amides are protected from exchange (expect for non-H-bonded 141V NH), form a cluster on the interior of the domain, with many of the side chain atoms within 4.5Å of each other. In place of 143V and 141V, *E. coli* pol IV has two alanine residues, and the equivalent to 133I (130I in pol IV) is located away from the cluster. The LF also has branched chain amino acids clustered in the central spine of the domain (323L, 322L, 318L, 319L, 286V, 336V, 270I, 266L, 338L, 334I, and 284I), Many of these residues are protected from exchange at 50°C, except for 338L, 336V, and 334I, which face the β10 strand contiguous with the very flexible linker region. The structural opening of the β10 strand, which has no protected residues likely facilitates fast exchange of the 338L, 336V, and 334I. In summary, the large of amount of

branched chain hydrophobic residues likely results in more efficient packing of the core of the palm and LF domains of Dbh is likely a key contributor to the thermostability of Dbh.

Salt-bridge and hydrogen-bonding networks can also impart thermostability to proteins (Mamonova et al., 2013; Jonsdottir et al., 2014; Makshakova et al., 2015). Examining the most protected residues in the *apo* Dbh crystal structure that are charged (Silvian et al., 2001) reveals that the some of them are involved in salt bridges. 128K and 124E, with protection factors of $>10^8$ and $1.7 \times 10^6$ at 50°C, form a salt bridge on the exterior of the F-helix. Further down the F-helix, the 135E side chain (PF at 50°C, $4.2 \times 10^5$) forms an H-bond with 131Q $\varepsilon$-$NH_2$, and 130K (PF at 50°C, $1.5 \times 10^7$) H-bonds with 162N CO and $\gamma$-CO, anchoring the G/β8 loop to the F-helix. The equivalent residue in pol IV 127R is too far away from the G/β8 loop to H-bond with the equivalent 159D backbone CO. The additional stabilization provided by side chain salt-bridges and hydrogen-bonding can rigidify the local structure, lowering the rate of transient opening reactions that allow hydrogen exchange. Further experiments to study the effect of salt-bridge and hydrogen-bonding disrupting mutations would be needed to confirm each residues' contribution to the thermostability of Dbh.

## Conclusions

The HDX experiments detailed above indicate the most stable domains of Dbh are the palm and LF domains. Many of the residues that are protected from exchange at 50°C are clustered together to form key sites of stability through hydrophobic interactions or networks of salt bridges. Maintaining the proper fold of the palm domain is likely crucial

for preserving the activity at high temperature. It is also important to maintain the shape of the LF as it is necessary for proper DNA binding and positioning, and residues important for substrate binding must be properly oriented in the major groove of DNA. Future HDX experiments can be performed on Dbh to investigate the stability and dynamics of the ligand bound states. Additionally, it would be informative to determine the residues which were stable even at 50°C also show protection at 65°C and 80°C, and therefore crucial to the stability of the Dbh Since many thermophilic enzymes show similar flexibility to the their mesophilic counterparts at their respective physiological temperatures, Given that Dbh is active *in vivo* at 80°C in *S. acidocaldarius* (Sakofsky et al., 2012), I hypothesize that residues in the interior of the β-sheets will be protected enough from exchange to detected by conventional HDX-NMR even at 80°C.

## References


Bai, Y., Englander, J.J., Mayne, L., Milne, J.S., and Englander, S.W. (1995). Thermodynamic parameters from hydrogen exchange measurements. Methods Enzymol 259, 344-356.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. Proteins 17, 75-86.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1994). Protein stability parameters measured by hydrogen exchange. Proteins 20, 4-14.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Gangadhara, B.N., Laine, J.M., Kathuria, S.V., Massi, F., and Matthews, C.R. (2013). Clusters of branched aliphatic side chains serve as cores of stability in the native state of the HisF TIM barrel protein. J Mol Biol 425, 1065-1081.

Hiller, R., Zhou, Z.H., Adams, M.W., and Englander, S.W. (1997). Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 degrees C. Proc Natl Acad Sci U S A 94, 11329-11332.

Hvidt, A., and Linderstrøm-Lang, K. (1955). Exchange of hydrogen atoms in insulin with deuterium in aqueous solution. Biochim Biophys Acta 14, 574-575.

Jonsdottir, L.B., Ellertsson, B.O., Invernizzi, G., Magnusdottir, M., Thorbjarnardottir, S.H., Papaleo, E., and Kristjansson, M.M. (2014). The role of salt bridges on the temperature adaptation of aqualysin I, a thermostable subtilisin-like proteinase. Biochim Biophys Acta 1844, 2174-2181.

Krishna, M.M., Hoang, L., Lin, Y., and Englander, S.W. (2004). Hydrogen exchange methods to study protein folding. Methods 34, 51-64.

Krishna, N.R., Huang, D.H., Glickson, J.D., Rowan, R., and Walter, R. (1979). Amide hydrogen exchange rates of peptides in H2O solution by 1H nuclear magnetic resonance transfer of solvent saturation method. Conformations of oxytocin and lysine vasopressin in aqueous solution. Biophys J 26, 345-366.

LeMaster, D.M., Tang, J., Paredes, D.I., and Hernandez, G. (2005). Enhanced thermal stability achieved without increased conformational rigidity at physiological temperatures: spatial propagation of differential flexibility in rubredoxin hybrids. Proteins 61, 608-616.

Maity, H., Lim, W.K., Rumbley, J.N., and Englander, S.W. (2003). Protein hydrogen exchange mechanism: local fluctuations. Protein Sci 12, 153-160.

Makshakova, O.N., Semenyuk, P.I., Kuravsky, M.L., Ermakova, E.A., Zuev, Y.F., and Muronetz, V.I. (2015). Structural basis for regulation of stability and activity in glyceraldehyde-3-phosphate dehydrogenases. Differential scanning calorimetry and molecular dynamics. J Struct Biol 190, 224-235.

Mamonova, T.B., Glyakina, A.V., Galzitskaya, O.V., and Kurnikova, M.G. (2013). Stability and rigidity/flexibility-two sides of the same coin? Biochim Biophys Acta 1834, 854-866.

Molday, R.S., Englander, S.W., and Kallen, R.G. (1972). Primary structure effects on peptide group hydrogen exchange. Biochemistry 11, 150-158.

Moro, S.L., and Cocco, M.J. (2015). H, C, and N backbone resonance assignments of the full-length 40 kDa S. acidocaldarius Y-family DNA polymerase, dinB homolog. Biomol NMR Assign.

Niu, C., Zhu, L., Zhu, P., and Li, Q. (2015). Lysine-Based Site-Directed Mutagenesis Increased Rigid beta-Sheet Structure and Thermostability of Mesophilic 1,3-1,4-beta-Glucanase. J Agric Food Chem 63, 5249-5256.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94, 12366-12371.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605-1612.

Raschke, T.M., and Marqusee, S. (1998). Hydrogen exchange studies of protein structure. Curr Opin Biotechnol 9, 80-86.

Sakofsky, C.J., Foster, P.L., and Grogan, D.W. (2012). Roles of the Y-family DNA polymerase Dbh in accurate replication of the Sulfolobus genome at high temperature. In DNA Repair (Amst) (Netherlands, 2012 Elsevier B.V), pp. 391-400.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8, 984-989.

Tossavainen, H., Kukkurainen, S., Maatta, J.A., Kahkonen, N., Pihlajamaa, T., Hytonen, V.P., Kulomaa, M.S., and Permi, P. (2014). Chimeric Avidin--NMR structure and dynamics of a 56 kDa homotetrameric thermostable protein. PLoS One 9, e100564.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59, 687-696.

Wang, K., Luo, H., Tian, J., Turunen, O., Huang, H., Shi, P., Hua, H., Wang, C., Wang, S., and Yao, B. (2014). Thermostability improvement of a streptomyces xylanase by introducing proline and glutamic acid residues. Appl Environ Microbiol 80, 2158-2165.

Yang, T.C., Legault, S., Kayiranga, E.A., Kumaran, J., Ishikawa, K., and Sung, W.L. (2012). The N-Terminal β-Sheet of the Hyperthermophilic Endoglucanase from Pyrococcus horikoshii Is Critical for Thermostability. In Appl Environ Microbiol, pp. 3059-3067.

Yu, H., Zhao, Y., Guo, C., Gan, Y., and Huang, H. (2015). The role of proline substitutions within flexible regions on thermostability of luciferase. Biochim Biophys Acta 1854, 65-72.

## Chapter 4 - $^{15}$N NMR Relaxation Spectroscopy of Dbh and Estimation of Global Rotational Correlation Time

### Rationale and Strategy

The principle of nuclear spin relaxation has been known to physicists for over half a century. First defined by Bloch, Hansen, and Packard in 1946 and expanded by Bloembergen, Purcell, and Pound in 1948, relaxation processes that occur in NMR-active nuclei in an external magnetic field return perturbed nuclear spins to an equilibrium population (Bloch, 1946; Bloch et al., 1946; Bloembergen et al., 1948). The rate of these relaxation processes is dependent on the motion of the molecule in solution; Brownian diffusional motion and internal motion of the molecule itself. In the case of protein, there are many internal motions of the molecule that enhance spin relaxation; therefore, spin relaxation rates contain dynamic information about the protein. Since the function of many proteins is dependent on internal dynamics, NMR spin relaxation can provide crucial information regarding the mechanism of function as well as protein stability. Solution NMR is unique in that provides quantitative atomic resolution dynamic data on proteins. Specifically, nuclear spin relaxation is sensitive to motions occurring on the picosecond to nanosecond time scale (Reddy and Rainey, 2010). These motions include bond vibration and librations, loop motions, side chain rotations, and smaller amplitude motions of the backbone (Morin, 2011).

*Spin Relaxation and Protein Dynamics*

There are two essential processes – longitudinal ($T_1$) and transverse relaxation ($T_2$) – which describe the decay of an NMR signal during an NMR experiment. Longitudinal

relaxation, describes the return of the population of nuclear spins to equilibrium in the *z* direction (direction of the static magnetic field) after perturbation from a radiofrequency pulse, and is typically on the order of seconds (Bloembergen, 1948). Longitudinal relaxation arises from the interactions of the nuclear spins with the surrounding environment; hence, it is also called spin-lattice relaxation. In contrast, transverse or spin-spin relaxation is the result of phase decoherence of the nuclear spins in the *xy*-plane, which leads to decay of the NMR signal. The decoherence, or dephasing of transverse magnetization as it precesses about the static magnetic field is due to the interactions of the local magnetic fields generated by each spin. The differences in local field cause the spins to precess at slightly different frequencies, generating the phase difference that leads to signal decay (Abragam, 1961). Transverse relaxation is enhanced with increasing molecular weight, which complicates the study of larger proteins by NMR (Chang et al., 2007). Another phenomenon which is sensitive to molecular motion is the heteronuclear Nuclear Overhauser Effect (hnNOE). The hnNOE involves cross-relaxation induced dipolar coupling between a proton and attached heteronucleus, which is modulated by motions of the bond vector (Morin, 2011).

These relaxation mechanisms are all sensitive to global and internal motion of the protein molecule in solution. Thus, the three essential experiments in any nuclear spin relaxation study are measuring the $T_1$ (longitudinal relaxation time), $T_2$ (transverse relaxation time) and the value of the heteronuclear $^{15}N$-$^{1}H$ NOE (Reddy and Rainey, 2010). Each of these parameters are measured separately; then, they are typically analyzed together to arrive at a model of motion for the bond vector. For the $T_1$ and $T_2$

experiments, peaks from 2D spectra taken at increasing delay times are fitted to an exponential decay equation to obtain the $R_1$ and $R_2$ rates (Morin, 2011). The $^{15}N$-$^{1}H$ NOE is measured as the ratio of peak amplitudes from spectra with and without proton saturation (Morin, 2011). There are a number of methodologies to interpret spin relaxation data; the most frequently used method is model-free analysis.

Lipari and Szabo developed the method termed "model-free" analysis (so named because no model is required *a priori* for the analysis) to extract two variables from spin relaxation data, the order parameter ($S^2$) and the effective internal correlation time ($\tau_e$) (Lipari and Szabo, 1982). "Model-free" analysis is so termed because the interpretation of the data does not depend on having a pre-existing model to describe the motion of the protein. Lipari and Szabo released that the correlation function describing the motion of a given bond vector could be distilled into a single generalized order parameter. $S^2$ describes the degree of restriction of motion of the bond vector, with an associated effective correlation time for this motion (Lipari and Szabo, 1982). Hence, it was possible to determine the time scale and degree of the motion for each detectable bond vector in the protein. However, in some cases, an accurate description of the motion of the bond vector by these two parameters is not possible. Therefore, Clore and co-workers expanded model-free analysis to include fast and slow motion order parameters with associated internal correlation times, that when combined yield the overall order parameter. One additional parameter is necessary to account for relaxation due to “chemical exchange” processes ($R_{ex}$) consisting of exchange between conformations on the µs-ms time scale, but other NMR experiments are necessary to

quantify these motions. To obtain the model-free analysis parameters, the spectral density function and model-free equations are fitted to the relaxation data through $\chi^2$ minimization (Lipari and Szabo, 1982; Clore et al., 1990). Development of software, such as Modelfree (Palmer III et al., 1991; Mandel et al., 1995), DASHA (Orekhov et al., 1995), and Relax (d'Auvergne and Gooley, 2008b, a), designed to streamline the model-free analysis of nuclear spin relaxation data has facilitated many studies on protein dynamics.

In model-free analysis, the overall tumbling time of the protein molecule in solution must be estimated to provide a starting point for estimating diffusion tensor of the protein molecule (Chang et al., 2007; Morin, 2011) The tumbling time can be estimated from the average $T_1/T_2$ of backbone amides, or estimated using hydrodynamic theory (Kay et al., 1989). The diffusion tensor can be isotropic (spherical) or anisotropic (an oblate or prolate spheroid, or ellipsoid), which (Lipari and Szabo, 1982). The estimate of diffusion tensor is evaluated with each of the model-free parameter sets until the best global fit of the data is achieved, which is called the diffusion-seeded paradigm (Lipari and Szabo, 1982; Clore et al., 1990). d'Auvergne and Gooley have developed an alternative approach to model-free analysis, whereby the diffusion tensor estimate is not necessary, but this requires the acquisition of relaxation at two separate static field strengths (d'Auvergne and Gooley, 2008a, b). This requirement doubles the amount of data needed to be collected and analyzed, and requires access to two high-field NMR magnets equipped for biomolecular NMR. Since we have obtained NMR relaxation data

at one field strength (800MHz), an estimate of the diffusion tensor must be calculated from the rotational correlation time.

Another reason the generalized order parameter is also particularly useful for the study of protein dynamics is that it can also be obtained from molecular dynamics simulation (Prompers and Bruschweiler, 2002; Fisette et al., 2012). The internal reorientation of the bond vector of the course of the simulation can be calculated from the trajectories directly and compared to experimental order parameters (Prompers and Bruschweiler, 2002). Unfortunately, I have not yet been able to process accurate heteronuclear NOE data on Dbh for the order parameter calculation. However, I have been able to fit $T_1$ and $T_2$ relaxation data on Dbh at both 35°C and 50°C. Since $T_1$ and $T_2$ values can also be calculated from the MD trajectories, I have directly compared the experimental $T_1$ and $T_2$ values for the purpose of evaluating the utility of the MD-derived order parameters.

**Materials and Methods**

30µL of 100% $D_2O$ 300 µL of samples of 100% $^{15}N$-labeled Dbh (prepared by V.V. Vu and D. Ji) containing 0.5mM or higher concentration in a buffer containing 50mM sodium phosphate, 100mM $K_2SO_4$, 50µm EDTA at pH 7.5, which were then transferred into a Shigemi NMR tube. The samples were placed in an 800MHz Varian Inova NMR spectrometer, containing an *xyz* triple resonance probe, equilibrated at 35 or 50°C. Eight 2D-$^{15}N$,$^{1}H$ TROSY-HSQC $T_1$ spectra (Zhu et al., 2000) at each temperature were collected for the $T_1$ experiment, with delay times of 10, 50, 100, 200, 400, 600, 800, and 1300ms. Six 2D-$^{15}N$,$^{1}H$ TROSY-HSQC $T_2$ spectra (Zhu et al., 2000) at each temperature were collected for the $T_2$ experiment, with delay times of 10, 30, 50, 70, 90,

and 110ms. Representative spectra for various time delays at both $T_1$ and $T_2$ experiments are shown in Figure 4.1a and b ($T_1$), and Figure 4.2a and b ($T_2$). The data was processed using NMRPipe (Delaglio et al., 1995), and visualized using CcPNMR Analysis (Vranken et al., 2005).

Twice as many scans were taken for last two spectra of the $T_1$ and the $T_2$ experiments; therefore, the data was multiplied by ½ during processing. The peak intensity was plotted as a function of delay time using CcPNMR, and fit to a single order exponential-decay function $(I(t) = I_0 \times e^{-kt})$ to extract the respective relaxation rates. Representative rate fits to the experimental data are displayed in Figure 4.3.

**Results and Discussion:**

$T_1$ and $T_2$ relaxation times have been determined for Dbh at 35°C and 50°C, and are displayed in Figure 4.4. Heteronuclear $^{15}$N-$^{1}$H NOEs measurements have been collected on Dbh. The heteronuclear NOE value is defined as the ratio between spectra with and without proton saturation($NOE_{NH} = I_{sat}/I_{unsat}$). However, the value of the ratio for many of the peaks is greater than 1. Since proton saturation should always result in a decrease of the signal In addition, the distribution of the values of the heteronuclear NOE is rather random and not consistent in regions of well-ordered secondary structure. For well-ordered regions in proteins, the value of heteronuclear NOE is usually around 0.7-0.8, with a tight distribution within contiguous regions of secondary structure (Kay et al., 1989; Sahu et al., 2000; Inman et al., 2001; Theret et al., 2001; Metcalfe et al., 2004; Goel et al., 2010). Therefore, the heteronuclear NOE experiment will have to be reanalyzed in order to calculate order parameters. Without the heteronuclear NOEs,

there is not enough information for the calculation, as the third parameter is necessary. The following analysis will focus on the information extractable from the $T_1$ and $T_2$ experiments, such as the estimate of tumbling time of the molecule in solution, and comparisons to the $T_1$ and $T_2$ value estimates from the molecular dynamics simulations of Dbh.

As expected, the average $T_2$ time for all amides is increased at 50°C, (average $T_2$ = 43.6 ± 3.0 ms versus 30.0 ± 1.8 ms at 35°C), reflecting increased tumbling time of the molecule. The $T_1$ values do not change as much with temperature; the average value at 35°C is 954 ± 136 ms and the average value at 50°C is 967 ± 102 ms. Since the tumbling time increases with temperature and the $T_1/T_2$ ratio is an estimator of the tumbling time, it makes logical sense that one of the values would be more sensitive to the increase in temperature. The $T_1/T_2$ ratio can provide a rough estimate of the correlation time, or tumbling time ($\tau_c$), in solution at a given temperature. The $T_1/T_2$ ratios for individual amides at each temperature are displayed in Figure 4.5. Including all of the measured amides, the $T_1/T_2$ ratio for Dbh at 35°C is 33.5 ± 5.4, and at 50°C 23.9 ± 2.9. The increase in the ratio with temperature is in line with on the dependence of the rotation correlation time with temperature (Garcia de la Torre et al., 2000). However, not all residues should be used, as slow internal motion of the amide bond vector results in a decrease in $T_2$.

Slow internal motion for particular residues in the μs/ms range contributes to $T_2$ relaxation, shortening the $T_2$ and resulting in an underestimation of the $T_1/T_2$ ratio (Yao et al., 1998). Failing to get an accurate estimate of the rotational correlation time of the molecule can ultimately prevent the fitting of the relaxation data for many residues in the

protein. Therefore, a number of methods have been devised for more accurate estimation, such as trimming ratios that are more than one standard deviation from the average, or selecting a particular minimum heteronuclear NOE cutoff value (Yao et al., 1998). The heteronuclear NOE is more sensitive to internal motion than the global rotational motion and hence is a good predictor of internal motion (Farrow et al., 1994). I cannot use the heteronuclear NOE ratio as a predictor of internal motion since I do not have reliable heteronuclear NOE data. Acquisition of NMR relaxation data at two static magnetic fields can also circumvent this issue. In this case, no estimation of the correlation time is necessary to begin model-free analysis, since six parameters ($T_1$, $T_2$, NOE at each field) are available to fit relaxation data (d'Auvergne and Gooley, 2008a, b).However, this requires access to two high-field spectrometers and exquisite temperature calibration to achieve the exact same temperature in both instruments. Furthermore, assuming the diffusion tensor to be axially symmetric is generally sufficient for accurate analysis of relaxation data for many proteins, and more extensive modelling is not necessary (Ryabov et al., 2006).

Since I do not have heteronuclear NOE data to identify residues undergoing internal motion, I have chosen to only use $T_1/T_2$ ratios from rigid positions involved in secondary structure in the protein. Residues involved in secondary structure should have fast effective internal correlation times (<100ps), since they are mostly restricted to vibrational motion. Excluding residues in loops or turns that are more likely to undergo slower motions that contribute to the $T_2$ decay for that residue prevents underestimation of the global tumbling time (Yao et al., 1998). After trimming $T_1/T_2$ ratios from residues not involved in secondary structure from the average of all $T_1/T_2$ ratios, the average ratio

becomes 32.5 ± 5.0 for 35°C and 23.4 ± 2.6 50°C. As expected, the trimmed average ratio is higher than the total average ratio, indicating that ratio of all spins may be underestimating the correlation time, due to conformational exchange of mobile residues on a similar time scale to global rotational diffusion. Therefore, an initial estimate of the rotational correlational time for Dbh at each of these temperatures could be based on the trimmed average ratios. However, after heteronuclear NOE data is available, it can be used to further distinguish residues undergoing slow exchange and refine the estimate of the correlation time.

The trimmed average $T_1/T_2$ ratio can be used to derive an estimate of the correlation time as seen in Eq. 1, where $v_N$ is the Larmor frequency of the $^{15}$N nucleus (Kay et al., 1989). However; using the frequency of the $^{15}$N nucleus in a 800MHz spectrometer (-81.09MHz), this equation yields a correlation time of 14.0 ± 2.2 ns at 35°C and 11.8 ± 1.4 ns at 50°C. However, this is a correlation time that would be more in line with a 20-25 kDa protein than Dbh. Since Eq. 1 was derived from data collected at 500MHz, and since $T_1$ changes more with static field strength than $T_2$ for slow-tumbling molecules such as proteins, this may not work as well for data collected at 800MHz. If the frequency of the $^{15}$N nucleus in 500MHz spectrometer is used for Eq. 1 (-50.68MHz), the correlation time result is 22.5 ± 3.5 ns at 35°C, and 18.9 ± 2.3 ns at 50°C. The decreases from in tumbling time from 35°C and 50°C for either value of $v_N$ are lower than would be expected (10.2 ± 1.2 ns, $v_N$ = -81.09MHz; 16.7 ± 2.0 ns, $v_N$ = -50.68MHz), so the equation may only work for data acquired at lower temperature.

$$(1)\ \tau_c \approx \frac{1}{4\pi v_N}\sqrt{6\frac{T_1}{T_2} - 7}$$

The value calculated from Eq. 1 using $v_N$ = -50.68MHz at 35°C (22.5 ± 3.5 ns) is closer to the predicted value from correlation time estimator based molecular weight [19.26 ns for a 41kDa at 35°C (Anthis, 2015)]. For 50°C, the estimator yields a $\tau_c$ 13.96 ns, closer to the expected (16.7 ± 2.0 ns) decrease from the value determined from Eq. 1 using $v_N$ = -50.68MHz at 35°C. The estimated values were determined from the following equation ($\tau_c$ = 0.0005998 x MW + 0.1674; MW in Daltons and $\tau_c$ in ns), which is based on linear fit of an experimentally determined data set of 20 proteins. However, the correlation time estimator is based on mostly smaller proteins (the largest is 21.9 kDa) than Dbh, and hence may not work well for proteins over 25 kDa (Northeast Structural Genomics Consortium, http://www.nmr2.buffalo.edu/nesg.wiki). Dbh is also Y-shaped; therefore, I would expect it to have a larger hydrodynamic radius and hence a longer tumbling time than an approximately spherical protein of similar molecular weight. Since the calculation from Eq. 1 using $v_N$ = -50.68MHz at 35°C of 22.5 ± 3.5 ns (extrapolated to 50°C, 16.7 ± 2.0 ns), are close to the estimated values based on other relaxation experiments, I would use these values for the initial estimate of the diffusion tensor for model-free analysis. If these values do not result in proper fitting of the relaxation data for a significant portion of the residues in Dbh, a relaxation interference experiment using the [15N,1H]-TRACT pulse sequence can be performed to directly determine the rotational correlation time for Dbh. The estimation of the rotational correlation time is necessary to obtain order parameters and effective internal correlation times for each measurable amide in Dbh. These parameters can then be compared to the order parameters obtained from molecular dynamics simulations. Once heteronuclear NOE data can be collected, then model-free analysis can be performed, and the

quantification of important areas of backbone flexibility and rigidity of Dbh can determined.

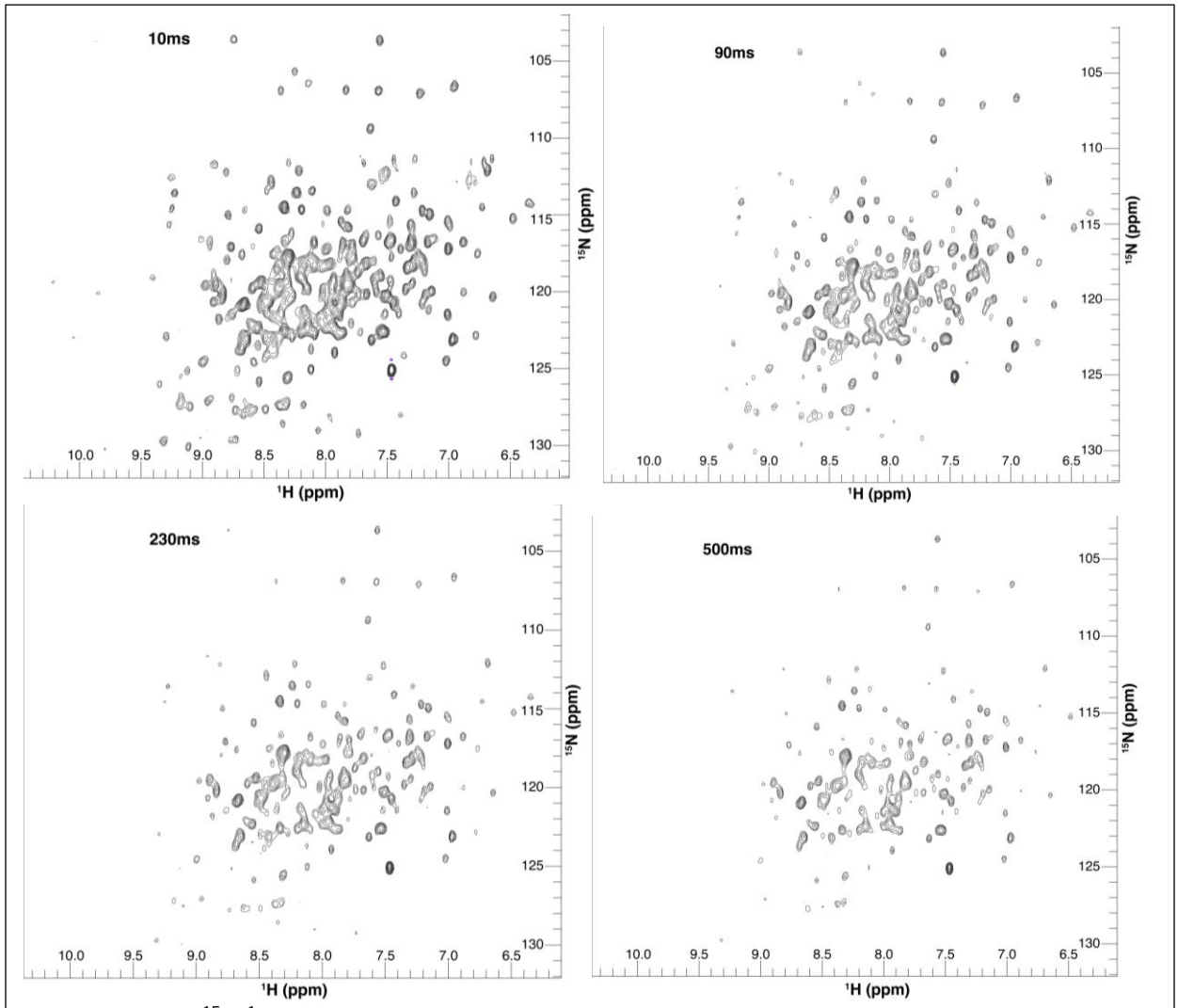



**Figure 4.1a:** $^{15}$N-$^{1}$H TROSY-HSQC T1 spectra Dbh at 35°C, with delay times as shown. The signal intensity decreases with increasing delay times.

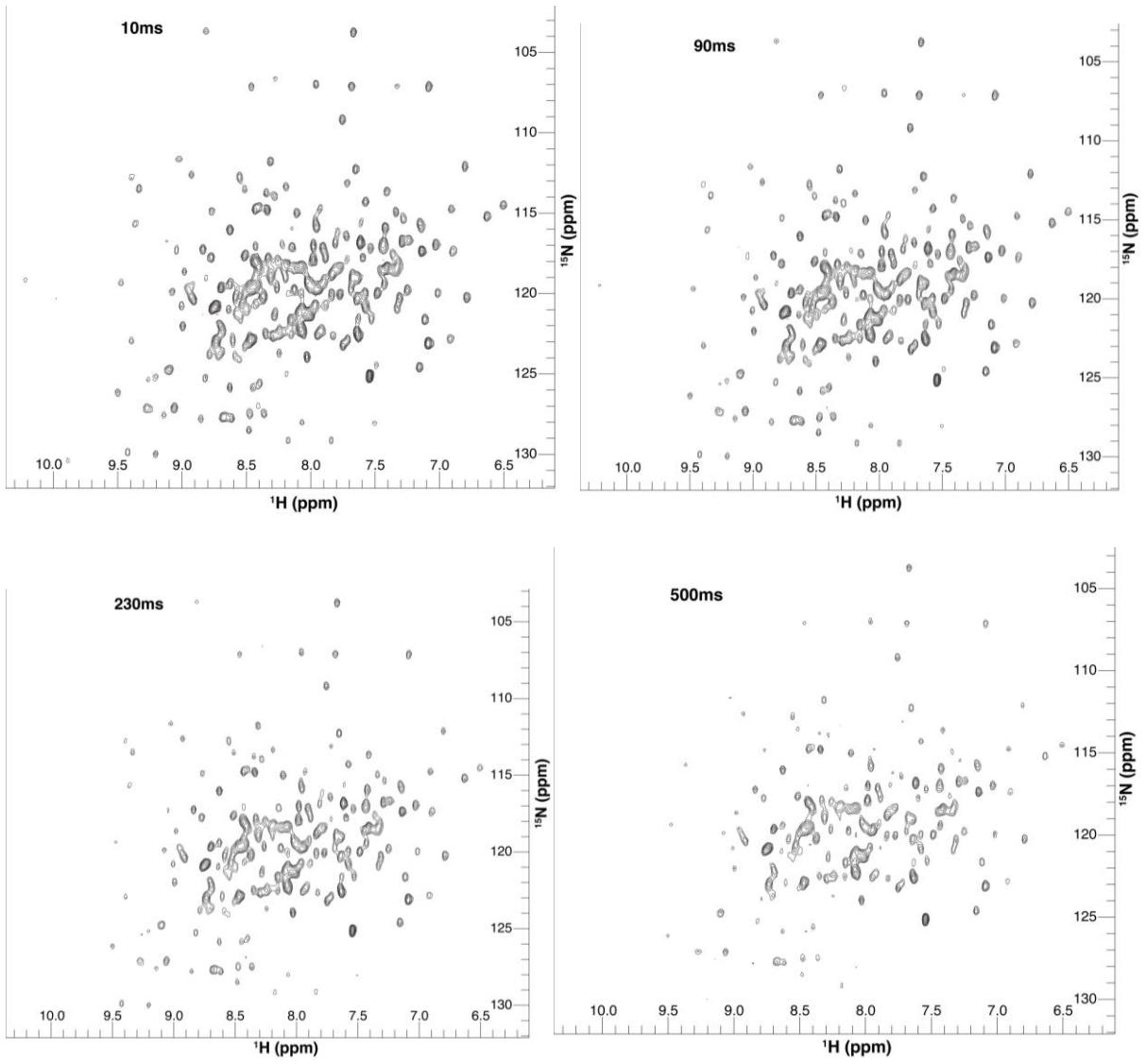



**Figure 4.1b:** $^{15}$N-$^{1}$H TROSY-HSQC T1 spectra Dbh at 50°C, with delay times as shown. The signal intensity decreases with increasing delay time.

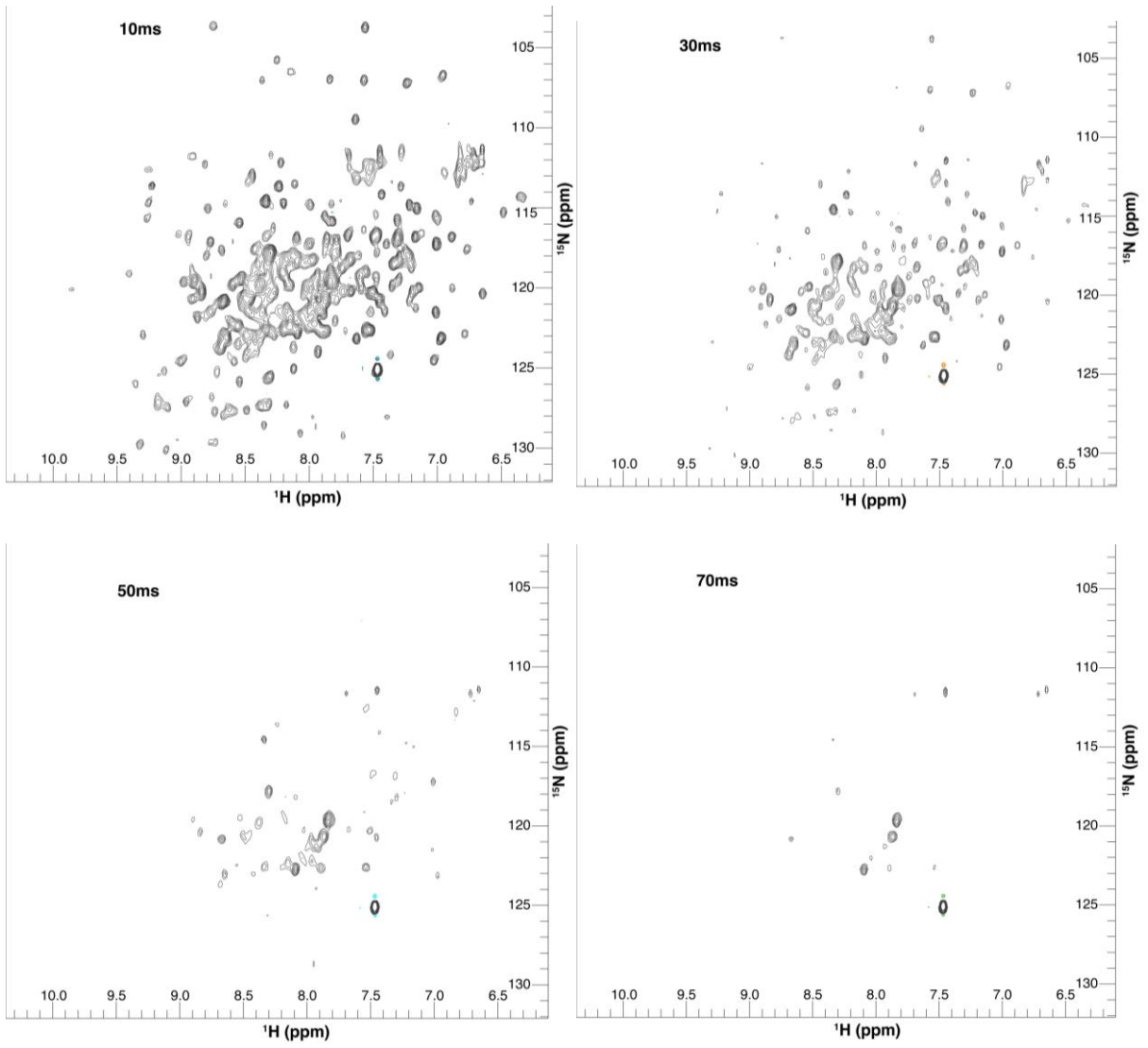



**Figure 4.2a:** $^{15}N$-$^{1}H$ TROSY-HSQC T2 spectra Dbh at 35°C, with delay times as shown. The signal intensity decreases with increasing delay times. By 70ms, most of the signals have decayed into the background noise.

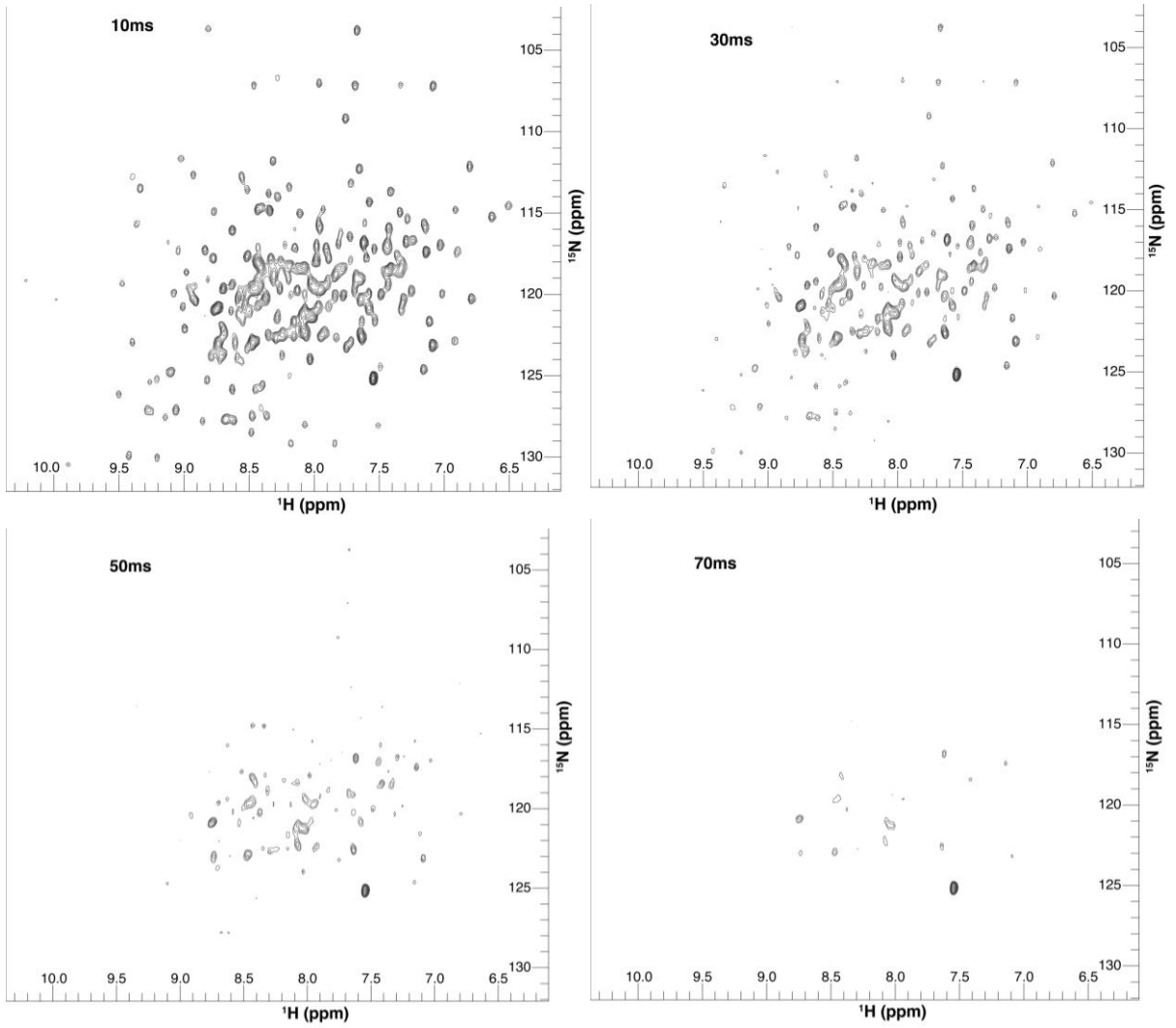



**Figure 4.2b:** $^{15}$N-$^{1}$H TROSY-HSQC T2 spectra Dbh at 50°C, with delay times as shown. The signal intensity decreases with increasing delay times. By 70ms, most of the signals have decayed into the background noise.

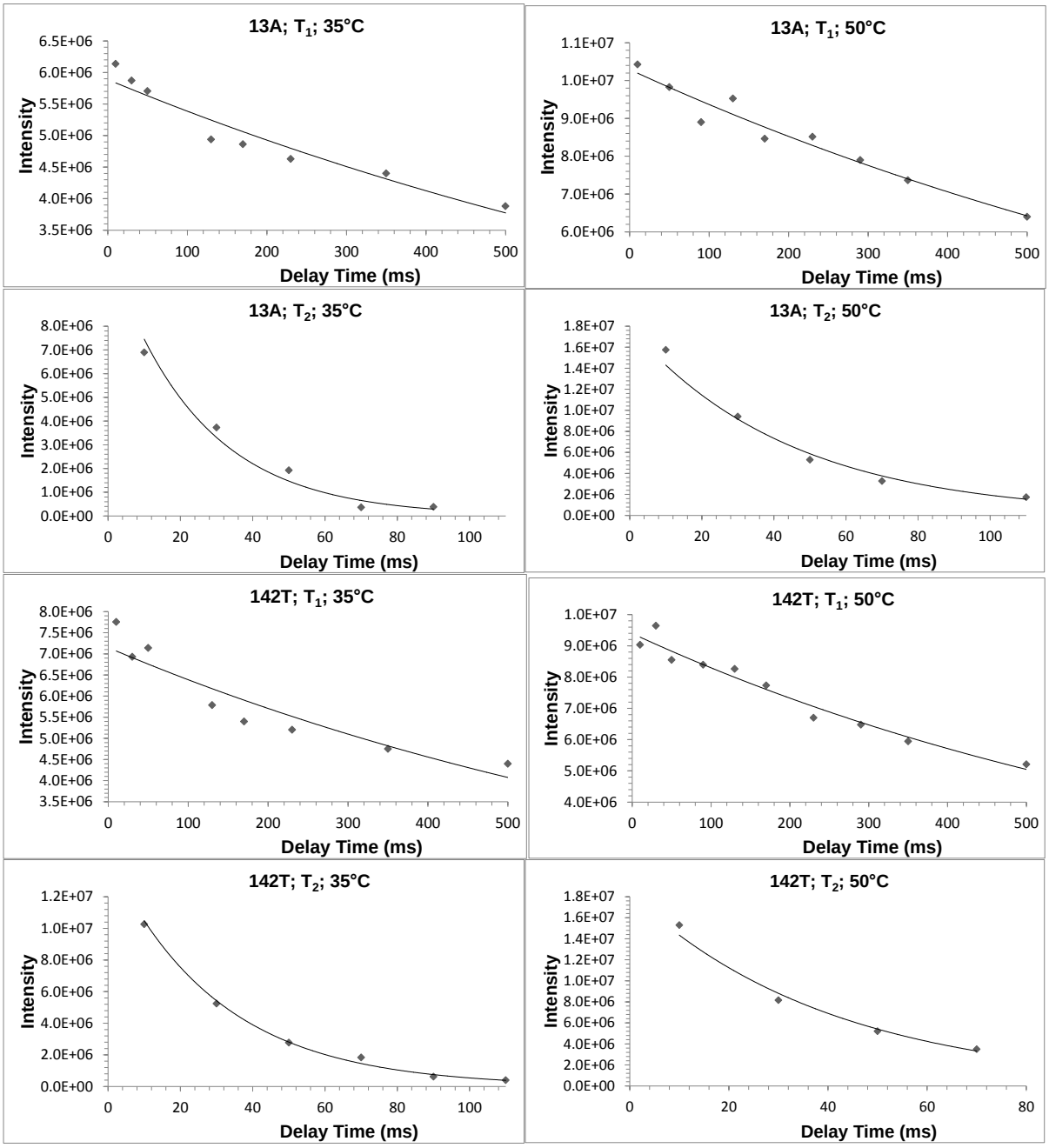



**Figure 4.3:** Representative rate fits for $T_1$ and $T_2$ data for Dbh residues 13Ala and 142Thr, at 35°C and 50°C. 13Ala is located in an $\alpha$-helix in the fingers domain, and 142Thr is located in central β-sheet of the palm domain. The decay rate is much faster for $T_2$ than $T_1$ relaxation.

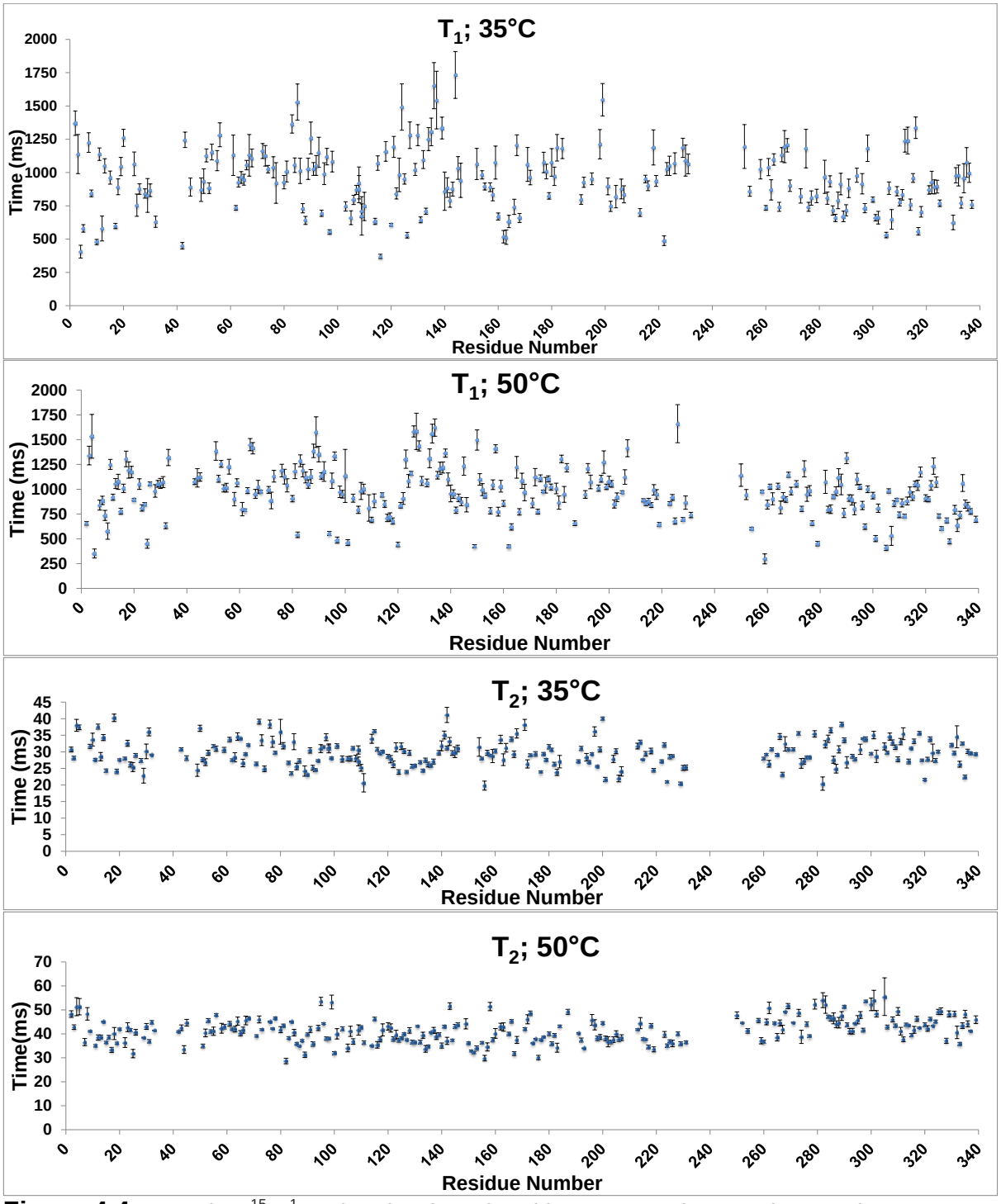



**Figure 4.4:** $T_1$ and $T_2$ $^{15}N$-$^{1}H$ relaxation times for Dbh at 35°C and 50°C. The error bars represent the error of the rate fitting for each residue. The palm domain consists of residues 1-19 and 78-171, the fingers domain consists of residues 20-77, the thumb domain consists of residues 172-231, and the LF domain consists of residues 246-344. The linker region (residues 232-245) and the C-terminal tail (residues 345-354) are unassigned and hence no relaxation data can be determined for these regions.

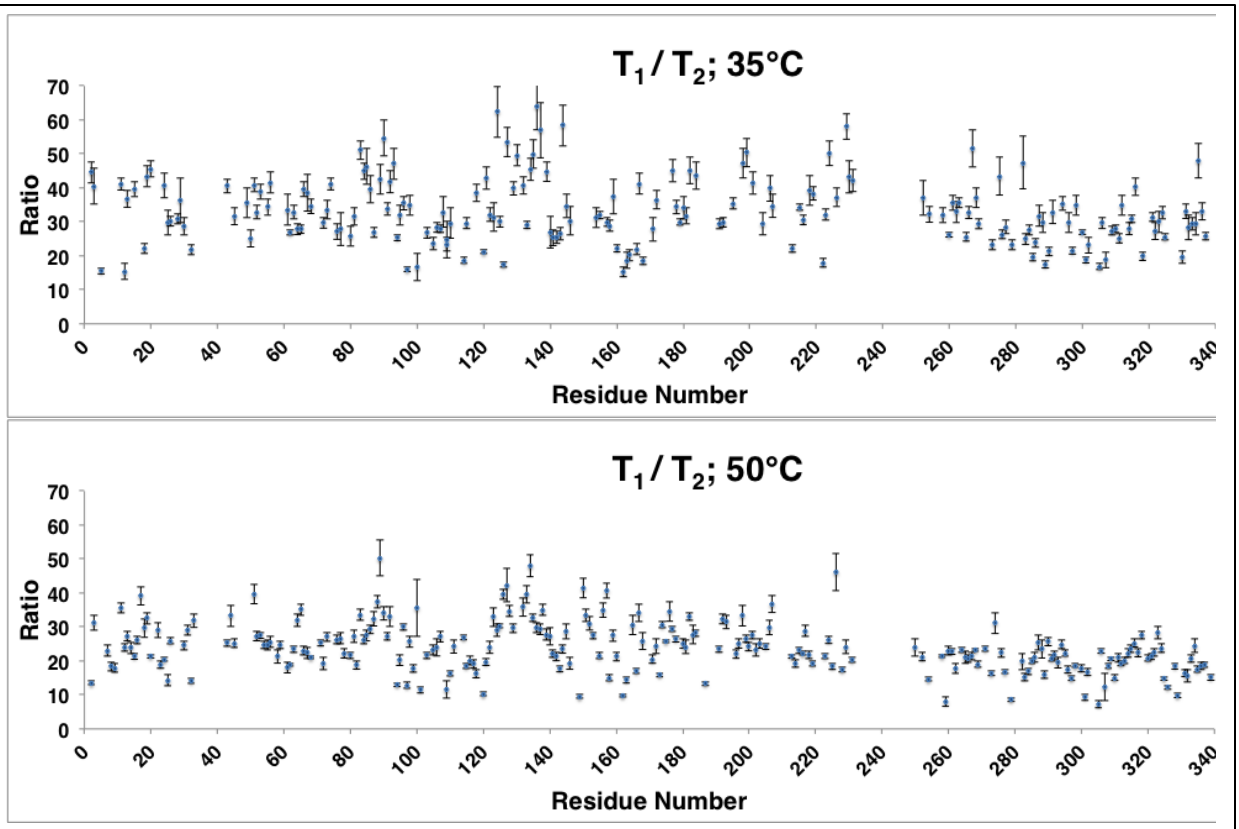



**Figure 4.5:** $T_1$ / $T_2$ $^{15}N$-$^{1}H$ ratios for Dbh at 35°C and 50°C. The average of the $T_1$ / $T_2$ ratios for rigid positions was used to arrive at an estimate of the overall rotational correlation time for Dbh. The error bars represent the propagation of the errors of the rate fits for the individual $T_1$ and $T_2$ experiments. The palm domain consists of residues 1-19 and 78-171, the fingers domain consists of residues 20-77, the thumb domain consists of residues 172-231, and the LF domain consists of residues 246-344. The linker region (residues 232-245) and the C-terminal tail (residues 345-354) are unassigned and hence no relaxation data can be determined for these regions.

**References:**


Abragam, A. (1961). The Principles of Nuclear Magnetism (Clarendon Press, Oxford).

Anthis, N. (2015). Protein Correlation Time Calculators.

Bloch, F. (1946). Nuclear Induction. Phys Rev 70, 460-474.

Bloch, F., Hansen, W.W., and Packard, M. (1946). The Nuclear Induction Experiment. Phys Rev 70, 474-485.

Bloembergen, N. (1948). Relaxation Effects in Nuclear Magnetic Resonance Absorption.

Bloembergen, N., Purcell, E.M., and Pound, R.V. (1948). Relaxation Effects in Nuclear Magnetic Resonance Absorption. Phys Rev 73, 679-712.

Chang, S.L., Hinck, A.P., and Ishima, R. (2007). Model-free analysis for large proteins at high magnetic field strengths. J Biomol NMR 38, 315-324.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C., and Gronenborn, A.G. (1990). Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. J Am Chem Soc 112, 4989.

d'Auvergne, E.J., and Gooley, P.R. (2008a). Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. J Biomol NMR 40, 107-119.

d'Auvergne, E.J., and Gooley, P.R. (2008b). Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. J Biomol NMR 40, 121-133.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry 33, 5984-6003.

Fisette, O., Lague, P., Gagne, S., and Morin, S. (2012). Synergistic applications of MD and NMR for the study of biological systems. J Biomed Biotechnol 2012, 254208.

Garcia de la Torre, J., Huertas, M.L., and Carrasco, B. (2000). HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. J Magn Reson 147, 138-146.

Goel, A., Tripet, B.P., Tyler, R.C., Nebert, L.D., and Copie, V. (2010). Backbone amide dynamics studies of Apo-L75F-TrpR, a temperature-sensitive mutant of the tryptophan repressor protein (TrpR): comparison with the (15)N NMR relaxation profiles of wild-type and A77V mutant Apo-TrpR repressors. Biochemistry 49, 8006-8019.

Inman, K.G., Baldisseri, D.M., Miller, K.E., and Weber, D.J. (2001). Backbone dynamics of the calcium-signaling protein apo-S100B as determined by 15N NMR relaxation. Biochemistry 40, 3439-3448.

Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 28, 8972-8979.

Lipari, G., and Szabo, A. (1982). Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. J Am Chem Soc 104, 4546-4559.

Mandel, A.M., Akke, M., and Palmer, A.G., 3rd (1995). Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. J Mol Biol 246, 144-163.

Metcalfe, E.E., Zamoon, J., Thomas, D.D., and Veglia, G. (2004). (1)H/(15)N heteronuclear NMR spectroscopy shows four dynamic domains for phospholamban reconstituted in dodecylphosphocholine micelles. Biophys J 87, 1205-1214.

Morin, S. (2011). A practical guide to protein dynamics from 15N spin relaxation in solution. Prog Nucl Magn Reson Spectrosc 59, 245-262.

Orekhov, V.Y., Nolde, D.E., Golovanov, A.P., Korzhnev, D.M., and Arseniev, A.S. (1995). Processing of heteronuclear NMR relaxation data with the new software DASHA. Appl Magn Reson 9, 581-588.

Palmer III, A.G., Rance, M., and Wright, P.E. (1991). Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy. J Am Chem Soc 113, 4371-4380.

Prompers, J.J., and Bruschweiler, R. (2002). General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. J Am Chem Soc 124, 4522-4534.

Reddy, T., and Rainey, J.K. (2010). Interpretation of biomolecular NMR spin relaxation parameters. Biochem Cell Biol 88, 131-142.

Ryabov, Y.E., Geraghty, C., Varshney, A., and Fushman, D. (2006). An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. J Am Chem Soc 128, 15432-15444.

Sahu, S.C., Bhuyan, A.K., Majumdar, A., and Udgaonkar, J.B. (2000). Backbone dynamics of barstar: a (15)N NMR relaxation study. Proteins 41, 460-474.

Theret, I., Baladi, S., Cox, J.A., Gallay, J., Sakamoto, H., and Craescu, C.T. (2001). Solution structure and backbone dynamics of the defunct domain of calcium vector protein. Biochemistry 40, 13888-13897.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59, 687-696.

Yao, S., Hinds, M.G., and Norton, R.S. (1998). Improved Estimation of Protein Rotational Correlation Times from 15N Relaxation Measurements. J Magn Reson 131, 347-350.

Zhu, G., Xia, Y., Nicholson, L.K., and Sze, K.H. (2000). Protein dynamics measurements by TROSY-based NMR experiments. J Magn Reson 143, 423-426.

## Chapter 5 – Molecular Dynamics Simulations of Dbh and $Dbh_{RKS(243-245)}$*

**These studies were performed in collaboration with the Luo laboratory at UC Irvine.*

*Special thanks to Prof. Ray Luo for the use of his laboratory's computational resources.*

### Rationale and Strategy

Hydrogen-deuterium exchange and nuclear spin relaxation spectroscopy provide abundant, rich, and detailed dynamic information on proteins, such as the degree and time scale of motions and relative order. Nevertheless, it is difficult to extend this data to a visual model of protein motion, limiting a hypothesis that explains the relationship of the dynamics to protein function. Fortunately, it is possible to obtain additional dynamic information as well as model the experimental data through computational methods, in particular molecular dynamics simulations. Combined with experimental methods, computational methods provide additional support for any hypothesis. In particular, nuclear spin relaxation and molecular dynamics simulation complement each other well. The strengths of one method cover the weaknesses of the other, and they can be used to refine each other to give the most detailed and accurate description of the data. For instance, the degree of motion of bond vectors can be determined experimentally by NMR, but the interpretation of the motion in the context of the structure is not straightforward. The MD simulation can provide a visual interpretation of structural motion, yet at the moment only ns-μs trajectory lengths are feasible in most settings. In addition, even with steady improvements in force fields MD simulations for large biomolecules remain classical approximations of motion and should be validated with experimental data. These concerns can be mitigated by the fact that parameters generated by NMR spin relaxation studies can be directly compared to parameters

calculated from MD trajectories, such as the $S^2$ order parameter (Fisette et al., 2012).In this case, the MD simulation can be validated by how well the $S^2$ calculated from the MD trajectories match the $S^2$ values determined from NMR relaxation experiments.

There are a number of methodologies to calculate $S^2$ from MD trajectories, the most common of which is the iRED (isotropic reorientational eigenmode dynamics) method (Prompers and Bruschweiler, 2002). The iRED method was chosen because it has been demonstrated to be comparable to experimental order parameters (Gu et al., 2014; Stafford et al., 2015), as well as the fact that the analysis is built into the CPPTRAJ module of AMBER14. In principle, iRED involves the calculation of a covariance matrix of internuclear bond vector reorientations, combined with integration of each frame over an isotropic distribution of orientations (Prompers and Bruschweiler, 2001, 2002). Crucially, this approach obviates the need to separate overall tumbling motion from internal motions. For simulations much longer than overall tumbling time (the simulations of Dbh are up to ~40x tumbling time), the separability of internal from overall motion is much more difficult; calculation of order parameters by iRED is appropriate for trajectories of hundreds of nanoseconds. To obtain the most accurate calculation of order parameters by iRED, it has been empirically demonstrated that the ideal window of simulation time to use is approximately 5 times the correlation time (Gu et al., 2014). Therefore, the simulations of Dbh were performed to 500ns, to provide sufficient time for four separate 125 iRED windows (approximate tumbling for Dbh is ~22.5 ns at 35°C, estimated from the $T_1/T_2$ ratio at 35°C (Yao et al., 1998)]).

Principal component analysis (PCA), also is another technique frequently used to analyze MD trajectories (Mu et al., 2005; Skjaerven et al., 2011; Franco-Gonzalez et al.,

2013; Jaeger and Pfaendtner, 2013; Stafford et al., 2013; Sittel et al., 2014). PCA is dimensional reduction procedure which can reveal the most dominant motions over the course of the trajectory (Hayward and de Groot, 2008). PCA involves diagonalizing the covariance matrix $C$ of the system to obtain eigenvalues of the diagonal matrix $\Lambda$ and associated eigenvectors contained in matrix $V$ (Eq. 1).

$$(1) C = V\Lambda V^T$$

The eigenvectors with the highest eigenvalues describe the first few principal components of the system (Hayward and de Groot, 2008). The eigenvectors can then be projected along the atoms of a system to isolate and visualize low frequency, coordinated motions of the system from higher frequency components (Wolf and Kirschner, 2013). The first few principal components are usually sufficient to capture the majority of the overall motion of the system (Amadei et al., 1993), and this motion often important in defining the function of the protein. In this way, PCA eliminates high-frequency vibrational motions and more clearly shows overall domain movements (Hayward and de Groot, 2008). Therefore, I used PCA to demonstrate the major movements of the domains of Dbh with respect to each other over the course of the trajectories. PCA analysis revealed that the major motions were rotation little finger domain propagated by rotation of the linker, and the flexing of the thumb and fingers domains with respect to the palm.

I performed simulations on Dbh at 35°C and 50°C for 500ns for direct comparison with the HDX and spin relaxation experiments. The AMBER14 MD software package (Case et al., 2015) was used to generate the MD trajectories, which utilizes the AMBER

12FFSB force field, along with the TIP4P-Ew explicit water model (Horn et al., 2004). AMBER is a very popular and well-established MD simulation software – multitudes of MD studies have been published using the software. TIP4P-Ew provides superior modeling of water over a wide range of temperatures in comparison to other explicit water models(Horn et al., 2004). Since my simulations were performed at 35°C and 50°C, TIP4P-Ew is the logical choice for simulation at two elevated temperatures.

The apo state of the protein from PDB entry 1K1S (Silvian et al., 2001) was used to generate the MD trajectories. The dynamics of the apo state can reveal intrinsic motions of the protein in the absence of DNA contacts, which provides a baseline for changes that occur upon substrate binding. Then, one can understand which motions present in the apo form could be important for function of the enzyme.

Different windows of the Dbh trajectory show larger movements in the LF domain; consequently, the averaging of multiple windows should provide a better estimation of the order parameter for the LF domain. The MD simulations of Dbh indicated a higher degree of order in the polymerase core of the protein, with lower order in the little finger domain. However, it can be deduced from the trajectories that the lower order parameters of the little finger result from the reorientations with respect to the core of Dbh over tens of nanoseconds.

In addition, I performed additional simulations at 35°C and 50°C for 250ns on a structural variant of Dbh, which has three residues in the linker region substituted with corresponding residues from the related protein Dpo4 (KIP 243-245 → RKS, henceforth referred to as $Dbh_{RKS(243-245)}$. The $Dbh_{RKS(243-245)}$ variant has been shown to resemble

Dpo4 in its lesion bypass properties and the orientation of the LF domain during DNA and nucleotide binding (Mukherjee et al., 2014). Given the substitution of a proline residue for polar serine and the substitution of a solvent-exposed isoleucine for a charged lysine, I would expect the mutation to increase the degrees of freedom of rotation for the linker, and should increase the movement of LF with respect to the polymerase core. This increased movement in fact was captured in the MD simulations, with the $Dbh_{RKS(243-245)}$ variant possessing similar order parameters in the core at 35°C, and increased order parameters in the LF at both 35°C and 50°C.

## Methods

*Simulation Preparation and Execution*

Amber14 software packages were used to run all simulations (Case et al., 2015). PDB structure 1K1S (apo form of Dbh) was used as the starting structure for the simulation (Silvian et al., 2001). 1K1S is missing coordinates for residues 36-38 in the beta-5-6 loop in the fingers domain and residues 345-354 in the C-terminal tail. Residues 36-38 were built in using the program MODELLER (Webb and Sali, 2014); the structure outputted from MODELLER (1K1S_loop) was superimposable on 1K1S with an RMSD of ~0.5Å (Silvian et al., 2001). Visual inspection of the structures also revealed no gross discrepancies. The C-terminal tail is not resolved in any structures of Dbh and is disordered, so the extended conformation peptide was generated in xleap, and appended to residue 344 of 1K1S_loop. The resulting structure (1K1S_looptail) was inspected for clashes, and no significant issues were found. Since there is no apo crystal structure for $Dbh_{RKS(243-245),}$ a PDB file for 1K1S_looptail was generated, manually

edited the residue codes of backbone coordinates for residues KIP 243-245 to RKS, and deleted the side chain atom coordinates. The missing side chains for RKS 243-245/1K1S_looptail were generated by *xleap*. Amber FF12SB was used as the force field. Both structure models of Dbh were placed in an octahedral box of size (250Å x 250Å x 250Å) with periodic boundaries, solvated with TIP4P-Ew water molecules (Horn et al., 2004), and neutralized with $Cl^-$ ions. The structures were energy minimized over a total of 2500 cycles with SANDER. Bond vibration of hydrogen atoms was constrained with the SHAKE algorithm (Ryckaert et al., 1977) allowing a time step of 2 fs to be used. For heating through the first portion of the equilibration of the models, a structural restraint in the form of $k(\Delta^2)$ ($k = 2$ kcal/mol−Å$^2$) using the minimized structure as a reference. Gradual heating of the restrained models to simulation temperature (35°C and 50°C) using a Langevin dynamics thermostat with a collision frequency set to 2 $ps^{-1}$ and random number seeding was run over 100 picoseconds. This was followed by density equilibration using a Berendsen barostat with a $\tau_p$ value (pressure relaxation time) of 2ps, and with the same structural restraint over 100 picoseconds with SANDER. Equilibration with gradually decreasing structural restraints was performed under constant pressure conditions over 1ns, followed by 5ns of free equilibration. Production runs were performed using PMEMD under constant volume conditions for 500ns for Dbh for at 308K and 323K, and for 250ns for $Dbh_{RKS(243-245)}$ at 308K and 323K, with coordinates written to trajectory every 10 picoseconds to ensure sufficient sampling for order parameter calculations.

*Trajectory data analysis*

Data analysis was performed using CPPTRAJ, a trajectory analysis program included in AMBER14 (Roe and Cheatham III, 2013). RMSD plots were for the trajectory were generated for Cα atoms, aligned and referenced to the first frame, using the rms command. Principal component analysis was performed according the CPPTRAJ diagmatrix command for Cα atoms and eigenvectors were for the first five principal components were projected along the Cα atom coordinates to visualize the motions of the principal components. A representative script of the analysis can be found in Appendix C.

Isotropically reorientational eigenmode dynamics (iRED) was used to generate $S^2_{MD}$ values for the trajectories, using the built-in **ired** matrix analysis in CPPTRAJ. The iRED methodology involves calculating the isotropically average covariance matrix of the spatial functions of the spin interactions in the N-H bond vector (Eq. 2) (Prompers and Bruschweiler, 2001, 2002).

$$(2)\ M_{ij}^{\mathrm{iRED}} = \frac{1}{2}\langle 3\left(\mathbf{u}_{\mathrm{LF},i}\mathbf{u}_{\mathrm{LF},j}\right)^2 - 1\rangle_{T_{\mathrm{iRED}}}$$

Then, an eigenvalue decomposition of the iRED matrix (Eq. 3) is performed to obtain the $S^2$ value (Prompers and Bruschweiler, 2002).

$$(3)\ S_k^2 = 1 - \sum_{m=6}^{N} \lambda_{m\,|\langle m|k\rangle|^2}$$

.

Chosen trajectory time windows for the iRED method have been found to best correspond with experimental data at ~5$\tau_c$ (Gu et al., 2014). Since the experimental correlation time has been estimated to be ~24ns for Dbh, a time window of approximately equal to 5$\tau_c$, in this case 125 ns, was chosen for each ired calculation. The iRED values for four non-overlapping 125ns windows for Dbh and two 125ns non-overlapping windows for $Dbh_{RKS(243-245)}$ were averaged to generate the final values. A representative script of the analysis can be found in Appendix C.

The hydrogen-bond occupancy analysis of the entire length of the trajectories was performed using the **hbond** command of CPPTRAJ (Roe and Cheatham III, 2013). The distance cutoff used for the analysis was 3.5Å, the atom mask selected for backbone hydrogen bonds only, and the angle cutoff for the bond was 135°. VMD (Humphrey et al., 1996) was used to visualize all trajectories and the projections of the principal components, and to generate the snapshots of the structures displayed in Figure 5.2.

**Results and Discussion**

Simulations of Dbh were run for 500ns at 35°C and 50°C, and for 250ns for $Dbh_{RKS(243-245)}$ at 35°C and 50°C. The RMSD values for the course of the simulation, from the first frame are displayed in Figure 5.1.The protein model remained stable over the course of each simulation. Much of the increase in the RMSD values is due to the change in the orientation of little finger domain with respect to the polymerase core. Order parameters were calculated according the iRED method as described in the methods section and are plotted by residue in Figures 5.2 and 5.3.

*Linker flexibility allows substantial reorientation of the LF*

The dominant motion in the simulations of Dbh is the overall movement of the little finger domain, propagated through the flexible linker. Principal component analysis analyzes the lowest frequencies motions which contribute the major fluctuations over the course of the trajectory. The projection of each principal component eigenmodes onto the structure can highlight any major conformational changes occurring during the simulation. Inspection of the principal component projections to the structure of Dbh reveal the dominant modes of motion for the LF domain, which rotates and flexes with respect to the polymerase core. The RMSF from an average structure of the Cα atoms of each residue calculated from the first five principal components of Dbh is plotted in Figure 5.4. Similarly, the displacement of the first two principal components from the average structure is projected onto the on the positions of Cα atoms of each residue is displayed in Figures 5.5 and 5.6, providing a visual model of the motion captured by the principal components. It is quite apparent that the linker domain does not move as significantly at 35°C, as the energy in the system is not enough to disrupt the H-bonding contacts between the palm domain β5 strand, the linker short β9 strand, and the β14 strand of the little finger domain. The main motion at 35°C is a hinging or flexing motion that changes the angle of the LF with respect to the core without rotation about the plane formed through the center of the LF and palm domains. This motion is very similar in amplitude and angle to the change in orientation from the apo crystal structure of Dbh to the DNA-bound structure (3BQ0) (Wilson and Pata, 2008). However, at 50°C, given the flexibility of the linker, there is enough thermal energy to separate these bridging H-bonds. The LF is then free to change conformation relative to the palm domain, and

rotations in the backbone in the linker region are responsible for the change in orientation.

Dbh is a very slow polymerase at ambient temperature, but is increasingly active at higher temperatures (Potapova et al., 2002). Wilson and Pata also noted in the ternary complexes of Dbh with DNA and correct incoming nucleotide, that the positioning of the 3-'OH primer terminus was further away from the ideal distance from the 5'-$\alpha$-phosphate for catalysis (Wilson and Pata, 2008). Therefore, it may be necessary for the LF to have more degrees of freedom during ligand binding to properly orient the substrates for efficient catalysis. Clearly, further simulations to observe the motion of the LF in the ligand-bound form of Dbh are needed to confirm whether the increased LF motion at higher temperature does improve conformational sampling of the active site.

The order parameter calculations also capture the motion of the LF; the $S^2_{MD}$ values are systematically lower for the LF domain over the course each simulation, and especially at the higher temperature. Comparing the Dbh $S^2_{MD}$ to the $Dbh_{RKS(243-245)}$ $S^2_{MD}$ values, the values are slightly lower for $Dbh_{RKS(243-245)}$ at 35°C, primarily in the little finger and thumb domains (Figure 5.3, top panel). A drastic reorientation of the linker in the $Dbh_{RKS(243-245)}$ simulation at 35°C was not observed, even with the greater conformation flexibility of the $Dbh_{RKS(243-245)}$ linker. In the palm and fingers domains, the difference between the order parameters is even smaller. Even with serine in place of proline at the 245 position in the linker of $Dbh_{RKS(243-245)}$, the LF is not able to undergo motion as dramatic as Dbh at 50°C. However, for $Dbh_{RKS(243-245)}$ at 50°C the $S^2_{MD}$ values are lower than Dbh at 50°C not only in the LF domain, but portions of the secondary structure in the palm as well (Figure 5.3, bottom panel). The linker of $Dbh_{RKS(243-245)}$ allows a more

drastic reorientation of the LF domain. Additionally, the lack of stable hydrogen bonding interaction from the linker to the β5 strand of the palm (residues 98 through 103) increases β5 strand away from hydrogen-bonding distance β6 strand and the rest of the palm. The principal component RMSF from average (Figure 5.4) and $S^2_{MD}$ clearly demonstrates the increased movement of β5 strand and other regions in the palm. The increase in motion of the palm occurs in Dbh at 50°C as well, but to a lesser degree. This movement clearly alters the conformation of the β5-β6 loop on which the metal-coordinating active site residue D105 is located and the adjacent residue E106 on the start of the β6 strand. Since $Dbh_{RKS(243-245)}$ has lesion bypass properties and activity closer to Dpo4, which is also far more active enzyme than Dbh at lower temperatures (Mukherjee et al., 2014), the greater motion of the LF could explain a different mechanism for lesion accommodation and bypass. The increased motion of residues near the active site in the palm could also indicate a mechanism for the greater activity of $Dbh_{RKS(243-245)}$. Nevertheless, it remains to be seen if the increased motion of the β5 strand $Dbh_{RKS(243-245)}$ compared to Dbh would also be observed in simulations and dynamics experiments of the DNA bound forms of the enzymes.

Further analysis of the MD trajectories, PC pseudo-trajectories, and order parameters reveals significant motion in other areas of the protein. The thumb domain (residues 171-231) is particularly mobile at 50°C for both Dbh and $Dbh_{RKS(243-245)}$. The thumb flexes away from and toward the palm domain. This motion could represent a “gripping” motion involved in DNA binding and translocation, as observed in an experimental FRET study on Dpo4 (Xu et al., 2009). Further dynamics experiments on DNA-bound forms of Dbh would be necessary to test this hypothesis. Unsurprisingly, the C-terminal

tail and the β3-4 loop of the protein are very dynamic. The C-terminal behaves as a random coil during the simulation; there is some partial folding . Since it is unresolved in DNA-bound forms of Dbh, it does not appear to have functional significance in binding u. In the apo crystal structures for Dbh, the residues near β3-4 loop had the highest B-factors in the protein, and the loop itself was not resolved in the apo [PDB entries 1K1S and 1IM4, (Silvian et al., 2001; Zhou et al., 2001)] and binary forms of Dbh [PDB entries 3BQ0 and 3BQ2, (Wilson and Pata, 2008)]. The loop is only resolved in the ternary structure of Dbh with nucleotide bound [PDB entry 3BQ1, (Wilson and Pata, 2008)], where it helps position the template strand.

*Comparison with HDX*

To roughly compare to the results of the MD simulation to the hydrogen-deuterium exchange NMR data, the percentage of hydrogen bond occupancy over the course of the trajectories was determined. It must be noted that the intrinsic exchange rate of amide hydrogens occurs on the order of seconds, while the MD simulations for WT-Dbh were performed for only 500ns. Therefore, the comparison cannot be quantitative and is only intended to illustrate residues that are likely important loci of stability in Dbh. Nevertheless, the amide hydrogens that are most protected in the HDX-NMR experiments should be within hydrogen bonding distance of an acceptor during the course of the simulation.

The most protected residues in the HDX-NMR experiments are populated at least 80% of the time in the simulation, providing evidence that the protein remains stable over the course of experiment. One of the more intriguing observations from the HDX experiment

was that atoms in the β13 strand facing the β10 contiguous with linker (338L, 336V, 334I, and 332R) exchange in the dead-time of the experiment, even though they are hydrogen bonded with the β10 strand and not solvent-exposed. By comparison, the residues of the β13 strand facing the interior of the β-sheet toward the β11 strand (331V, 333R, 335G, 337K) are protected during HDX and hydrogen bonded over the course of the simulation. The simulation of Dbh at 50°C reveals the basis for the HDX result: the β10 flexes away from the β13 strand, breaking the bonding as far the 246H NH to 338L CO H-bond, which is occupied in 98% of frames at 35°C but only 22% of frames at 50°C. This reflects a local opening event which can account for the fast exchange of amides that are within hydrogen bonding distance in the crystal structure. Functionally, the flexing and partial unzipping of the β13 strand actual restricts the reorientation of the LF for a portion of the simulation, as hydrogen bonds are then formed transiently in the linker (239E CO to 241K NH, 240N CO to 242S NH, 242S CO to 244I NH; 22.4%, 7.8%, and 7.1% occupancy, respectively). It remains to be seen how the motion of the LF at 50°C affects DNA positioning before and after binding of a nucleotide.

**Conclusions**

The greatest motion observed in all of the simulations was the flexing or rotation of the LF domain with respect to the polymerase core. The motion of the LF was greatly increased in the 50°C simulations due to a breaking of the interaction of the LF through the linker to the palm, allowing rotation of the LF about the linker. The $Dbh_{RKS(243\text{-}245)}$ simulation also revealed greater flexibility compared to Dbh, even in regions not directly connected to the more flexible linker, such as the fingers and palm domains. Future

simulations with DNA and nucleotide bound models of Dbh will be necessary to determine how the nanosecond time scale motions of differ from the apo Dbh upon ligand binding, and to extract further information about DNA/nucleotide binding and positioning.

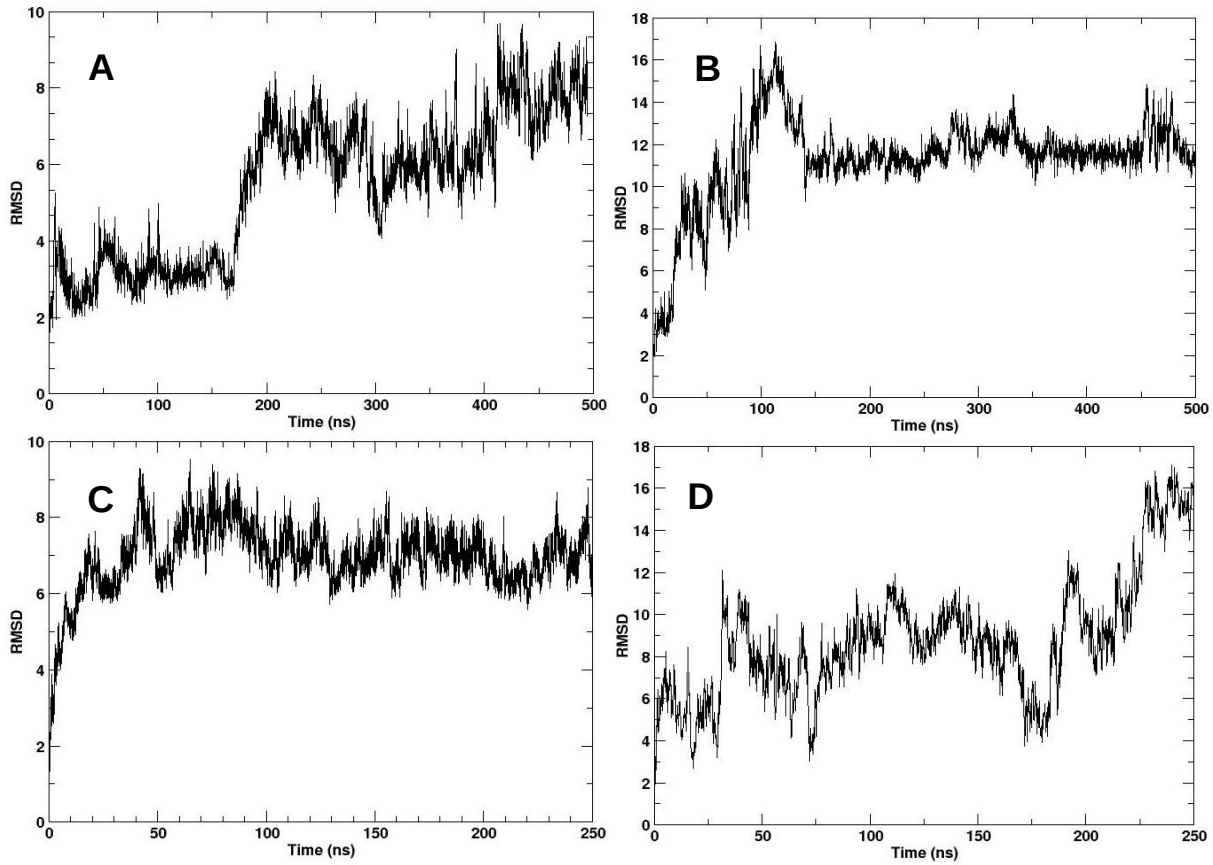



**Figure 5.1:** Root mean square deviation of Cα atoms (measured in Å) of Dbh at 35°C (A) and 50°C (B), and of $Dbh_{RKS(243-245)}$ at 35°C (C) and 50°C (D), referenced to the first frame of the production run. The increase in the RMSD can be primarily attributed to the changes in orientation of the little finger domain from its starting conformation.

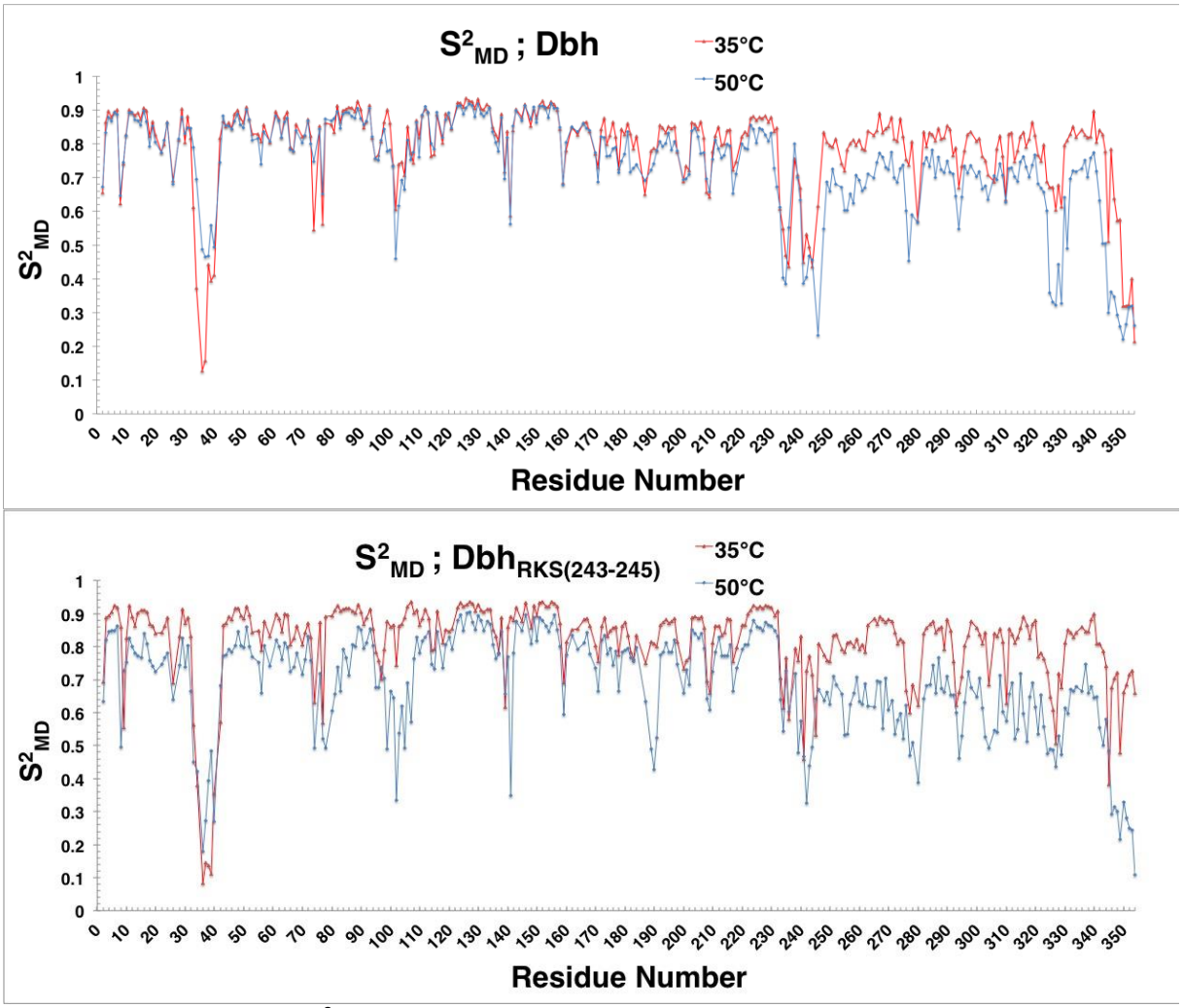



**Figure 5.2:** MD iRED ($S^2_{MD}$) backbone amide order parameters of Dbh at 35°C and 50°C (upper graph), and of $Dbh_{RKS(243-245)}$ at 35°C and 50°C (lower graph), plotted by residue number. The values for 35°C are red and the values for 50°C are blue in each graph. The palm domain consists of residues 1-19 and 78-171, the fingers domain consists of residues 20-77, the thumb domain consists of residues 172-231, and the LF domain consists of residues 246-344. The linker region (residues 232-245) and the C-terminal tail (residues 345-354) account for the rest of the enzyme. The increased mobility of the LF domain is captured in the simulations at 50°C, and the increased motion is more pronounced for the $Dbh_{RKS(243-245}$. The motion of the thumb domain is also increased at 50°C for both WT-Dbh and $Dbh_{RKS(243-245}$. For WT-Dbh, the palm and thumb domains show comparable rigidity at both temperatures. In contrast, $Dbh_{RKS(243-245}$ shows increased motion throughout the enzyme at 50°C.

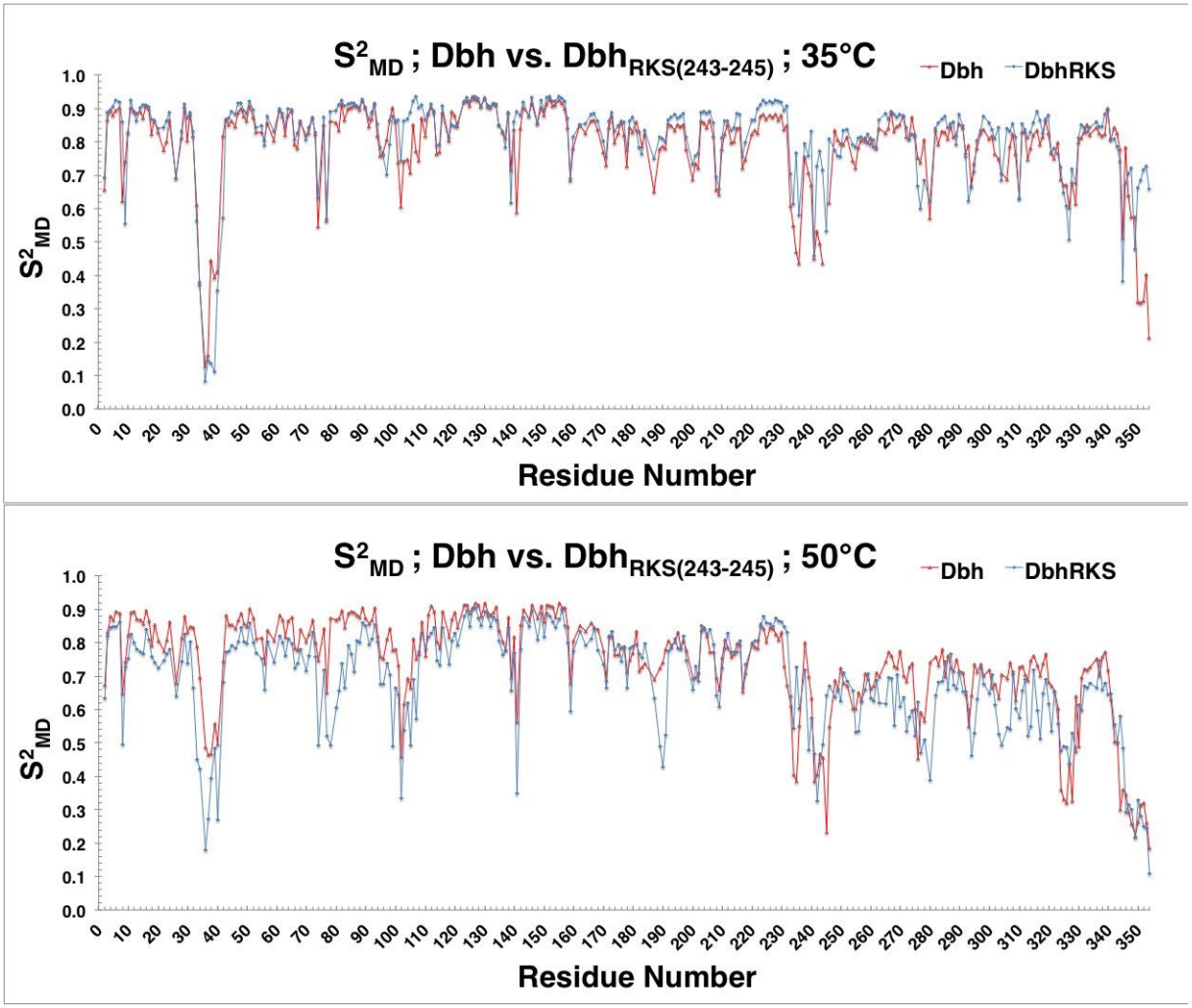



**Figure 5.3:** Difference in MD iRED ($S^2_{MD}$) backbone amide order parameters of Dbh vs. $Dbh_{RKS(243-245)}$ at 35°C (upper graph), and 50°C (lower graph), plotted by residue number. The order parameters are comparable at 35°C (slightly higher for $Dbh_{RKS(243-245)}$ in a few stretches of secondary structure), but significantly lower in many regions for $Dbh_{RKS(243-245)}$ at 50°C. The palm domain consists of residues 1-19 and 78-171, the fingers domain consists of residues 20-77, the thumb domain consists of residues 172-231, and the LF domain consists of residues 246-344. The linker region (residues 232-245) and the C-terminal tail (residues 345-354) account for the rest of the enzyme.

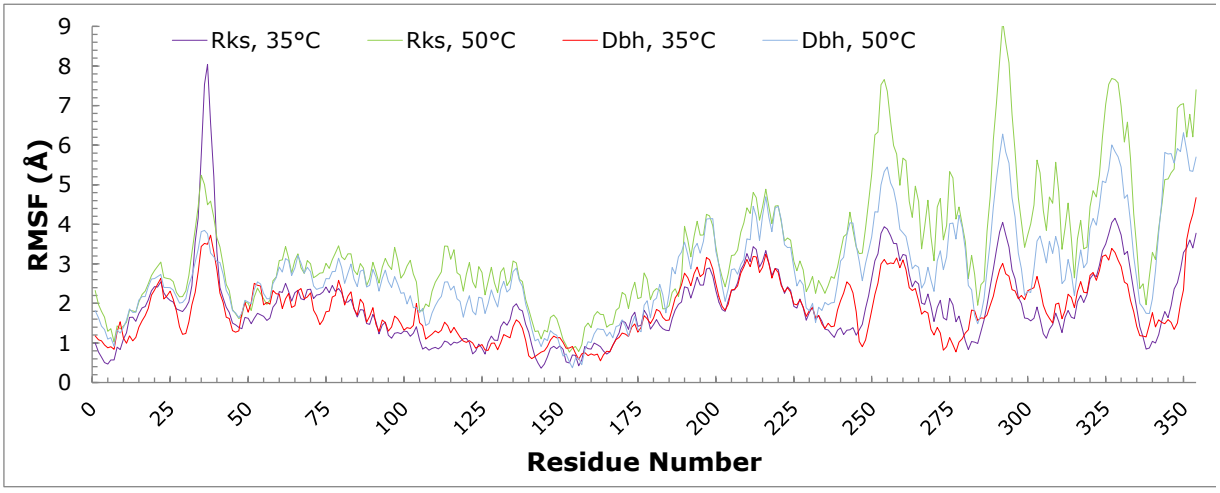



**Figure 5.4:** Root mean square fluctuation by residue of Cα atoms (measured in Å) from an averaged reference structure of the first five principal components Dbh at 35°C and 50°C, and of $Dbh_{RKS(243-245)}$ at 35°C and 50°C. The principal components were generated using data from the entire production trajectory. Increased motion is observed at 50°C, especially in the little finger domain (residues 246-354).

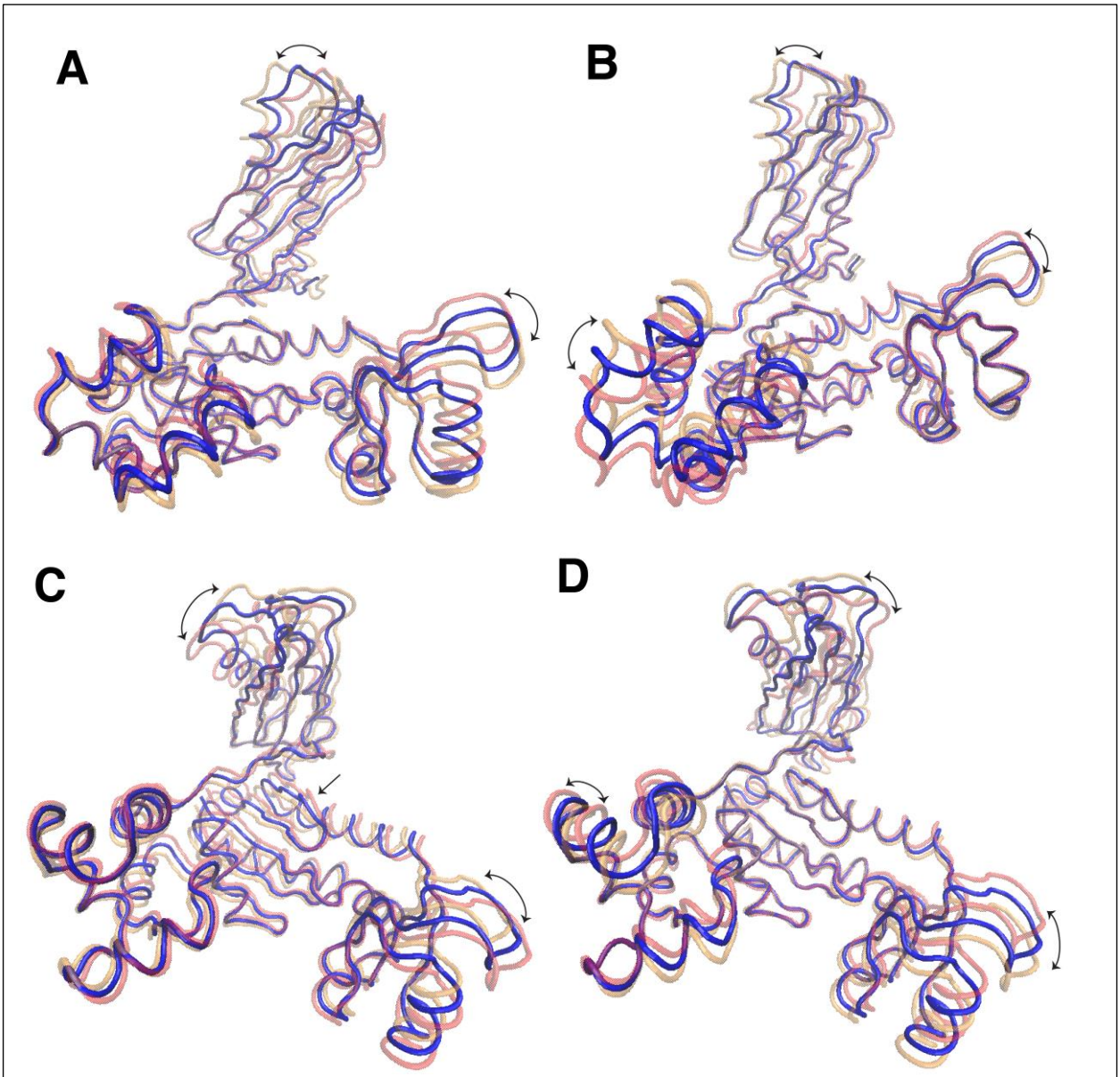



**Figure 5.5:** Projection of principal components 1 (left panels) and 2 (right panels) onto the coordinates of Cα atoms of Dbh at 35°C (A, PC1; B, PC2) and 50°C (C, PC1; D, PC2). The resulting pseudo-trajectories capture the major low-frequency motions of the main trajectory. The backbone is traced through the positions of the Cα atoms. The average structure or centroid about which the principal components oscillate is rendered in solid blue, while the maxima of the oscillation of each principal component are depicted in transparent red and orange. The arrows indicate direction of the oscillation about the centroid. The LF domain is located at the top of the structure, the palm in the center, the thumb on the lower left, and the fingers on the lower right. The arrow pointing to the β5 strand of the palm in panel C highlights increased motion of the strand at 50°C, which can alter the conformation of key active site residues.

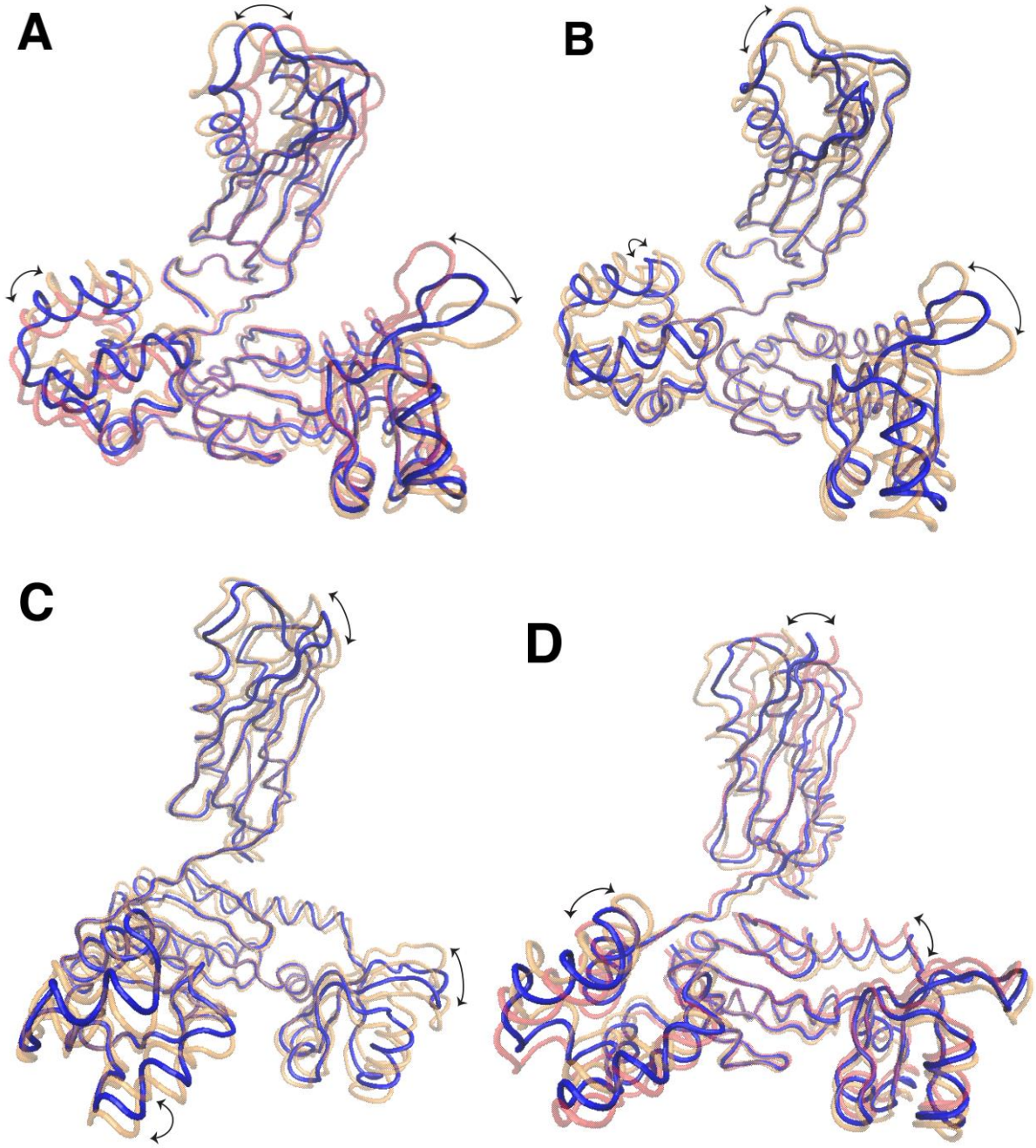



**Figure 5.6:** Projection of principal components 1 (left panels) and 2 (right panels) onto the coordinates of Cα atoms of $Dbh_{RKS(243-245)}$ at 35°C (A, PC1; B, PC2) and 50°C (C, PC1; D, PC2). Principal components capture the major low-frequency motions of the trajectory. The backbone is traced through the positions of the Cα atoms. The average structure or centroid about which the principal components oscillate is rendered in solid blue, while the maxima of the oscillation of each principal component are depicted in transparent red and orange. The arrows indicate direction of the oscillation about the centroid. The LF domain is located at the top of the structure, the palm in the center, the thumb on the lower left, and the fingers on the lower right.

## References:


Amadei, A., Linssen, A.B., and Berendsen, H.J. (1993). Essential dynamics of proteins. Proteins 17, 412-425.

Case, D.A., Berryman, J.T., Betz, R.M., Cerutti, D.S., Cheatham III, T.E., Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W.*, et al.* (2015). AMBER 2015 (University of California, San Francisco).

Fisette, O., Lague, P., Gagne, S., and Morin, S. (2012). Synergistic applications of MD and NMR for the study of biological systems. J Biomed Biotechnol 2012, 254208.

Franco-Gonzalez, J.F., Cruz, V.L., Ramos, J., and Martinez-Salazar, J. (2013). Conformational flexibility of the ErbB2 ectodomain and trastuzumab antibody complex as revealed by molecular dynamics and principal component analysis. J Mol Model 19, 1227-1236.

Gu, Y., Li, D.-W., and Brüschweiler, R. (2014). NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. J Chem Theory Comput 10, 2599-2607.

Hayward, S., and de Groot, B.L. (2008). Normal modes and essential dynamics. Methods Mol Biol 443, 89-106.

Horn, H.W., Swope, W.C., Pitera, J.W., Madura, J.D., Dick, T.J., Hura, G.L., and Head-Gordon, T. (2004). Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J Chem Phys 120, 9665-9678.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph 14, 33-38, 27-38.

Jaeger, V.W., and Pfaendtner, J. (2013). Structure, dynamics, and activity of xylanase solvated in binary mixtures of ionic liquid and water. ACS Chem Biol 8, 1179-1186.

Mu, Y., Nguyen, P.H., and Stock, G. (2005). Energy landscape of a small peptide revealed by dihedral angle principal component analysis. Proteins 58, 45-52.

Mukherjee, P., Wilson, R.C., Lahiri, I., and Pata, J.D. (2014). Three Residues of the Interdomain Linker Determine the Conformation and Single-Base Deletion Fidelity of Y-family Translesion Polymerases. J Biol Chem.

Potapova, O., Grindley, N.D., and Joyce, C.M. (2002). The mutational specificity of the Dbh lesion bypass polymerase and its implications. J Biol Chem 277, 28157-28166.

Prompers, J.J., and Bruschweiler, R. (2001). Reorientational eigenmode dynamics: a combined MD/NMR relaxation analysis method for flexible parts in globular proteins. J Am Chem Soc 123, 7305-7313.

Prompers, J.J., and Bruschweiler, R. (2002). General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. J Am Chem Soc 124, 4522-4534.

Roe, D.R., and Cheatham III, T.E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput 9, 3084-3095.

Ryckaert, J.P., Ciccotti, G., and Berendsen, A. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkenes. J Comp Phys 23.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8, 984-989.

Sittel, F., Jain, A., and Stock, G. (2014). Principal component analysis of molecular dynamics: on the use of Cartesian vs. internal coordinates. J Chem Phys 141, 014111.

Skjaerven, L., Martinez, A., and Reuter, N. (2011). Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit. Proteins 79, 232-243.

Stafford, K.A., Robustelli, P., and Palmer, A.G., 3rd (2013). Thermal adaptation of conformational dynamics in ribonuclease H. PLoS Comput Biol 9, e1003218.

Stafford, K.A., Trbovic, N., Butterwick, J.A., Abel, R., Friesner, R.A., and Palmer, A.G., 3rd (2015). Conformational preferences underlying reduced activity of a thermophilic ribonuclease H. J Mol Biol 427, 853-866.

Webb, B., and Sali, A. (2014). Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics 47, 5.6.1-5.6.32.

Wilson, R.C., and Pata, J.D. (2008). Structural insights into the generation of single-base deletions by the Y family DNA polymerase dbh. Mol Cell 29, 767-779.

Wolf, A., and Kirschner, K.N. (2013). Principal component and clustering analysis on molecular dynamics data of the ribosomal L11.23S subdomain. J Mol Model 19, 539-549.

Xu, C., Maxwell, B.A., Brown, J.A., Zhang, L., and Suo, Z. (2009). Global conformational dynamics of a Y-family DNA polymerase during catalysis. PLoS Biol 7, e1000225.

Yao, S., Hinds, M.G., and Norton, R.S. (1998). Improved Estimation of Protein Rotational Correlation Times from 15N Relaxation Measurements. J Magn Reson 131, 347-350.

Zhou, B.L., Pata, J.D., and Steitz, T.A. (2001). Crystal structure of a DinB lesion bypass DNA polymerase catalytic fragment reveals a classic polymerase catalytic domain. Mol Cell 8, 427-437.

## Chapter 6 – Response of Dbh to Temperature Change

### Rationale and Strategy

Thus far, I have presented the response of the dynamics of Dbh to temperature changes using three separate experimental methods: hydrogen-deuterium exchange NMR, nuclear spin relaxation, and molecular dynamics simulation. The data has revealed that the core folds of the palm and LF domains are very rigid and impart a high degree of stability to Dbh, while the fingers and thumb are more dynamic, but clearly are stably folded at 50°C. In addition, the MD simulation revealed greater flexibility in the LF domain at higher temperature, offering a possible explanation for the lower activity of Dbh at lower temperatures than a related polymerase, Dpo4. In this chapter, I present additional experiments for the response of Dbh to changes in temperature, the first of which is “temperature factor” or “temperature coefficient”. The temperature coefficient measures the change in chemical shift with increasing temperature, typically the $^{1}H$ chemical shift of the N-H backbone amide (Baxter and Williamson, 1997). The change in chemical shift with temperature is usually linear, and shifted downfield with increasing temperature; the slope of the line yields the temperature coefficient (Cierpicki and Otlewski, 2001). The magnitude of the change in proton chemical shift depends on the strength of the hydrogen bond; the change will be smaller for stronger hydrogen bonds.

The temperature coefficient can provide a prediction whether residues are participating in intramolecular hydrogen bonds, especially when this is combined with data from hydrogen-deuterium exchange experiments. Considered alone, it does not always correlate with the state of hydrogen-bonding (Tomlinson and Williamson, 2012), but can

be in combination with other experiments, such as hydrogen-deuterium exchange NMR (Hong et al., 2013), and the $^{3h}$J-NC' through-hydrogen-bond coupling. In most cases, the temperature coefficient is more positive than -4.6 ppb/K for hydrogen-bonded residues, predicting the presence of a hydrogen bond in more than 85% of the time (Cierpicki and Otlewski, 2001; Cordier and Grzesiek, 2002). The chemical shift of backbone amides can be subject to deshielding by ring currents in neighboring aromatic residues, which can cause the temperature coefficient to be more positive even in the absence of hydrogen bonding (Merutka et al., 1995; Cierpicki and Otlewski, 2001). In addition, it has been shown that amide proton temperature coefficients can be used to calculate the thermal expansion of the hydrogen bond, as the chemical shift is exquisitely sensitive to changes in the orientation and length of the hydrogen bond (Hong et al., 2013). The thermal expansion of hydrogen bonds is more pronounced in loop regions that are involved in weaker hydrogen bonds to solvent; therefore, the temperature coefficient should be more negative for these amides (Tilton et al., 1992; Cierpicki and Otlewski, 2001). It is possible to detect the presence of the hydrogen bond directly by NMR (Cordier et al., 2008), but this experiment was extremely time-consuming for even a small protein (ubiquitin, 8.6 kDa), and is likely infeasible for a protein as large at Dbh (40.8 kDa). Since the hydrogen-deuterium exchange data is a strong indication of the presence of a hydrogen bond, I have compared the values of the temperature coefficients with the protection factor data for residues for which both values could be calculated. The data also provides further insight in the utility and interpretation of temperature coefficients in the prediction of protein intramolecular hydrogen bonds.

*Investigating the possibility of cold denaturation above the freezing point of water in Dbh*

It is extremely well established that proteins unfold at high temperatures, but it is also possible for proteins to denature at cold temperatures. This implies that the stability curve for proteins us U-shaped, with its free energy minimum at some temperature wherein it is stably folded and functional. The cold denaturation temperature for most proteins is below the freezing point of water (or above freezing with added denaturants or in acidic or basic pH) (Privalov, 1990; Azuaga et al., 1992; Griko and Privalov, 1992; Babu et al., 2004; Whitten et al., 2006; Jaremko et al., 2013; Vajpai et al., 2013), but some proteins have been found to cold denature above zero °C (Pastore et al., 2007; Buchner et al., 2012). Since the U-shaped stability curve for thermostable proteins may either be shifted to the right instead of simply deeper than mesostable proteins, some thermostable proteins may cold denature above zero Celsius. After an HSQC of Dbh at 5°C showed a possible increase in random coil, we decided to investigate if Dbh cold denatures above zero °C. Additionally, it was observed that Dbh would not crystallize at 4°C, but would readily crystallize at room temperature, which could indicate partial aggregation at lower temperature (J. Pata, personal communication). To investigate possible denaturation of Dbh at low temperatures, we performed four separate experiments to look for evidence of unfolding in Dbh from 1-10°C: differential scanning calorimetry (DSC), dynamic light scattering (DLS), hydrogen-deuterium exchange NMR (HDX), and circular dichroism (CD). Differential scanning calorimetry can directly detect a change in heat capacity in a protein which results from a thermodynamic transition,

such as unfolding. Dynamic light scattering is sensitive to the hydrodynamic radius of a macromolecule; an unfolded protein should show an increase in radius versus the more compact folded form. Hydrogen-deuterium exchange NMR will detect unfolding when referenced to HDX experiments where the protein is stable; peaks that are visible at the stable temperature will disappear at lower temperature due to unfolding. HDX can even show local or transient unfolding, provided that the decrease in intrinsic exchange rate at lower temperature is corrected for. Finally, circular dichroism can measure loss of secondary structure and increase in random coil structure as a protein unfolds. If Dbh unfolds above zero °C, the combination of these experiments should provide abundant evidence of unfolding. Unfortunately, the experiments provided conflcting evidence of Dbh unfolding at temperatures just above freezing. Future experiments will be necessary to conclusively determine the degree and nature of structural changes of Dbh at lower temperature.

**Methods**

***Temperature Coefficient***

Conventional 2D-$^{15}$N-$^{1}$H HSQC spectra were taken at 35°C, 45°C, 50°C, 55°C, and 65°C, on samples containing at least 0.5mM $^{15}$N-labeled Dbh in 20mM HEPES, 50mMNaCl, 50μM EDTA, pH 7.5, with 10% (v/v) $D_2O$ added to the solution for spin locking. The data was processed with NMRpipe (Delaglio et al., 1995), and visualized using CcPNMR Analysis (Vranken et al., 2005). The chemical shift values for the resolvable amide protons were measured at each temperature, and then fit to a linear regression equation to obtain the temperature factor. For the hydrogen-bonding analysis, UCSF Chimera (Pettersen et al., 2004) was used to analyze residues within

hydrogen bonding distance from PDB structure 1K1S (Silvian et al., 2001), and CPPTRAJ (Roe and Cheatham III, 2013) was used to analyze the hydrogen bonds in the MD trajectories, as described in the methods section of Chapter 5.

### *Low-Temperature Biophysical Experiments to Investigate Possible Cold Denaturation*

*Differential Light Scattering (DLS)*

1 mL of Dbh buffer solution containing Dbh protein (50mM sodium phosphate, 100mM NaCl, 50μM EDTA, pH 7.5) was filtered through a 0.1μM polyvinylidene fluoride membrane to remove any large particles in the solution. The protein concentration was measured by light absorbance at 280nm and determined to be 1.27 mg/mL. The protein solution was pipetted into a disposable plastic cuvette with a path length of 1cm and placed in a Malvern Zetasizer Nano DLS instrument. The temperature was equilibrated at 35°C and measurements were taken for 30s and repeated ten times. The temperature was then lowered to 5°C, allowed to equilibrate for 10 minutes at 5°C, then measurements were taken as above.

*Differential Scanning Calorimetry (DSC)*

Dbh buffer solution containing Dbh protein was centrifuged at 10,000x *g* for 5 minutes to sediment any aggregates in solution, then extensively degassed under vacuum. The concentration of protein was then determined to be 1.7 mg/mL by light absorbance at 280nm. The sample was then transferred via syringe to the sample cell of a MicroCal VP-DSC microcalorimeter until the cell was filled; trapped air was expunged from the cell using extra solution in the syringe. The reference cell was filled with buffer solution

in an identical manner. After placing the cap over the calorimetry cells, the system pressure fluctuated between 5-6 psi. The equilibration time at the starting temperature before starting a scan was 15 minutes. During a continuous downscan from 70°C to 1°C and upscan from 1°C to 70°C, some protein precipitation was observed, negatively affecting the quality of the data. Therefore, alternating upscans and downscans were taken between 1°C and 35°C for one protein sample, and then from 30 to 70°C on a second protein sample. The first few degrees of each scan are part of instrument equilibration period, hence the $C_p$/dT values cannot be recorded. The upscan rate was 90°C/hr and the downscan rate was 60°C/hr. Four upscans and four downscans in each temperature range; the first has a different "thermal history" than the subsequent scans, and the following scans were performed to ensure repeatability. The scans were normalized for protein concentration to obtain the final curves. For the lysozyme sample, the same buffer was used as for Dbh, and the protein concentration was 2.1 mg/mL. One upscan and one downscan were taken, from 10°C to 80°C, at the same scan rates used for Dbh.

*Hydrogen Deuterium Exchange NMR (HDX-NMR)*

A sample of $^{15}$N-labeled Dbh solution (50mM sodium phosphate, 100mM NaCl, 50μM EDTA, pH 7.5) was lyophilized, then re-suspended in cold buffer at 4°C (same as above) containing 10% (v/v) $D_2O$, pipetted into a Shigemi NMR tube, and placed at 4°C for 18 hours. The sample was then placed in an 800MHz Oxford NMR spectrometer at 35°C and a 2D $^{15}$N-$^{1}$H TROSY-HSQC (Pervushin et al., 1997) spectrum was collected. The data was processed with NMRpipe (Delaglio et al., 1995) and visualized with CcPNMR analysis (Vranken et al., 2005).

*Circular Dichroism Spectroscopy (CD)*

200 µL of 0.28mg/mL Dbh protein in buffered (same as above) solution was transferred into a CD cuvette with a 1mm path length. The cuvette was placed in the sample cell Jasco J-810 spectropolarimeter equipped with a Peltier temperature controller, and the cell was calibrated at 35°C. A spectrum was collected at 35°C from 260nm to 185nm, with a scan rate of 50nm/min, data points collected every 0.5nm, and 10 total scans. The cell was then cooled at rate of 0.5°C/min to 7°C, with the ellipticity value measured at 208nm every 2°C, then after a full spectrum was taken at 7°C with same parameters as above. The sample was then removed from the CD cuvette, the cuvette was cleaned, and a second 200µL 0.28mg/mL Dbh protein sample was transferred into the cuvette. The cuvette was placed into the cell equilibrated at 35°C, another full CD spectrum was taken at this temperature, and the sample was heated to 65°C at a rate of 0.5°C/min, with the ellipticity value at 208nm taken every 2°C. One full spectrum was then take at 65°C. After adjustments to the chiller unit, a third 0.28mg/mL Dbh sample was used to collect full CD spectra again at 7°C and then at 2°C.

## Results and Discussion

### *Amide Temperature Coefficients*

The thermostability of Dbh allows NMR measurements to be performed at elevated temperatures; therefore, HSQCs were collected at 35°C to 65°C to measure the temperature factor over a wide temperature range. Figure 6.1 displays the HSQC spectra collected at each temperature; the spectra at 65°C shows Dbh remains well-folded at this temperature. The temperature coefficients were calculated for 255 of 277 (92%) of assigned backbone amide proton chemical shifts, which are displayed in

Figure 6.2. Of these residues, 213 are more positive than -4.6ppb/K, which was the limit to determine hydrogen-bonded residues suggested by Cordier and Grzesiek in their study of the temperature coefficients of ubiquitin (Cordier and Grzesiek, 2002). It must be noted that the limit was empirically determined and based only on ubiquitin; therefore, this limit may or may not have predictive value for the presence of hydrogen bonds in Dbh. Assuming that the temperature coefficient is influenced primarily by hydrogen-bonding in residues known to participate in secondary structure, Hong et al. (based on NMR and MD studies of ubiquitin) proposed that the thermal expansion coefficient of the hydrogen bond could also be estimated (Hong et al., 2013), which is shown in Eq. 1.

$$(1)\ \frac{d\mathrm{r}}{d\mathrm{T}} = \frac{1.16}{\delta - (4.06\mathrm{ppm})} \cdot \frac{d\delta}{d\mathrm{T}}\,\mathrm{Å}$$

However, small changes electronic environment around the amide hydrogen – for example, positioning of ring current or changes in magnetic anisotropy – could also have a significant effect on the temperature coefficient (Cordier and Grzesiek, 2002; Hong et al., 2013). Using Eq. 1 and the chemical shift value at 50°C, the average thermal expansion coefficient (*d*r/*d*T) of the hydrogen-bonded backbone amides of Dbh is 7.1 x $10^{-4}$ ± 3.9 x $10^{-4}$ Å/K, which is very similar to the values obtained by Hong et al. in their study of GB3 domain of protein G [7.9 x $10^{-4}$ ± 5.1 x $10^{-4}$ Å/K(Hong et al., 2013)]. Thus, it appears the expansion upon heating in hydrogen bonds from a thermostable protein compared a mesophilic protein domain is similar, at least for these two proteins. In combination with HDX and molecular dynamics data collected on Dbh, it is possible to make more definitive observations about the relationship between the temperature coefficient and hydrogen bonding strength.

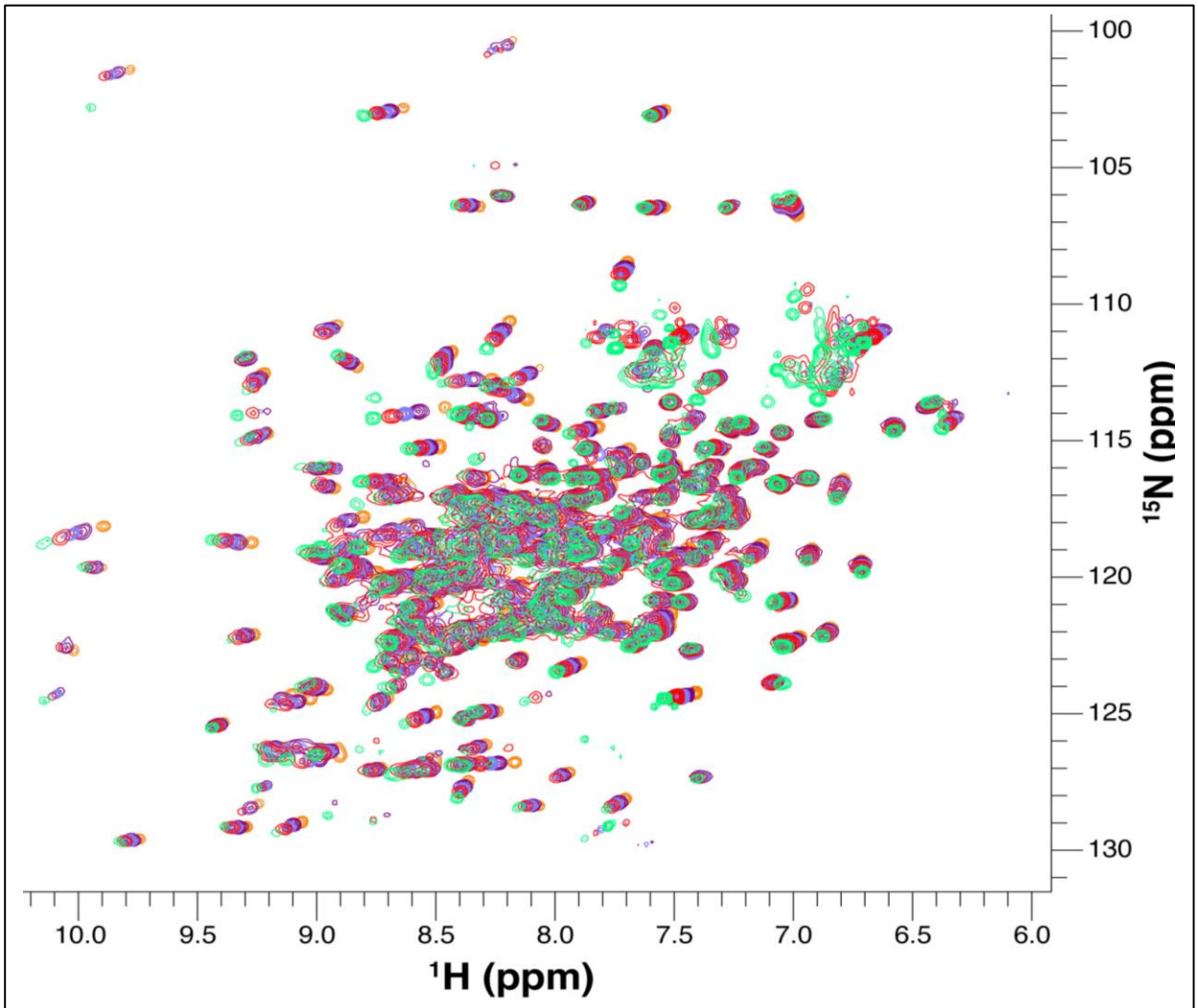



**Figure 6.1:** Overlay of $^{15}$N-$^{1}$H HSQCs of Dbh at 35ºC (green), 45ºC (red), 50ºC (blue), 55ºC (purple), and 65ºC (orange). The proton chemical shifts are shifted upfield with increasing temperature.

I compared the hydrogen-deuterium exchange rates for the 68 backbone amides for which a rate could be calculated to their respective temperature coefficients. The protection factors and temperature coefficients were uncorrelated, with an $R^2$ value of 0.004. This is not surprising since the protection factors are majorly dependent on tertiary structure blocking; intramolecular hydrogen bonds certainly contribute to the protection factor, but are not the sole determinant. Therefore, the protection factor is not

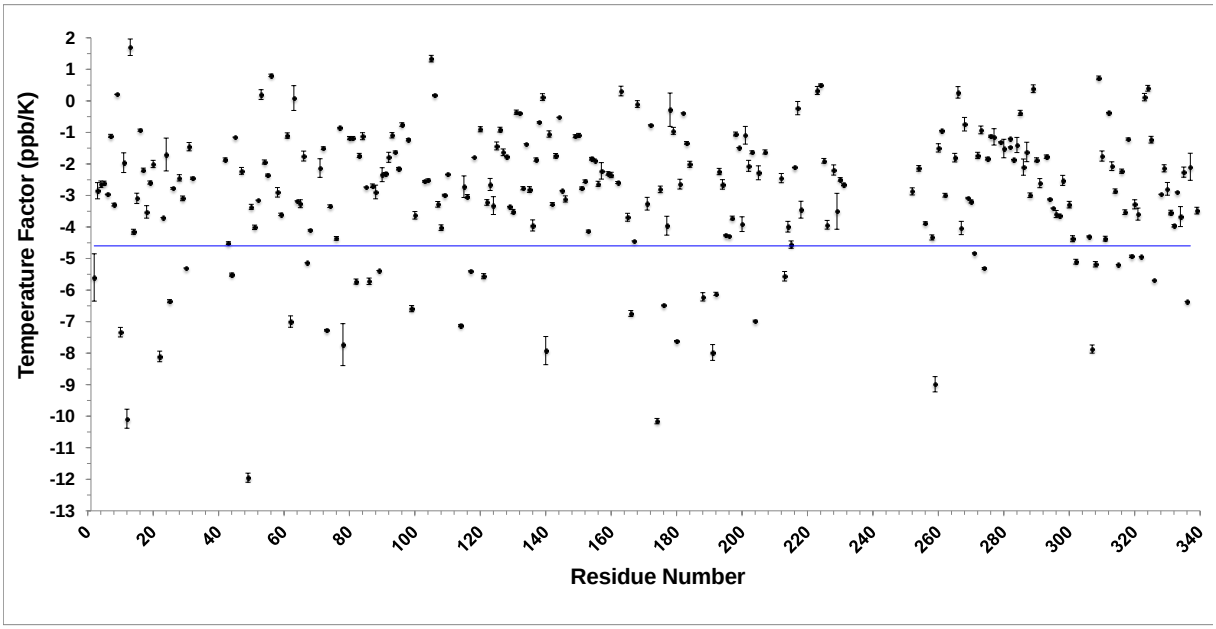



**Figure 6.2:** Temperature coefficients of the $^{1}$H(N) backbone amide chemical shifts. The blue line represents the cutoff point for hydrogen-bonded residues according to the criterion proposed by Cordier and Grzesiek. Residues above this line should be involved in the intramolecular hydrogen-bonded, below the line should be in hydrogen bonds with solvent. However, this is not a direct measurement of hydrogen bonding – it is only a prediction.

an outright measure of hydrogen bond strength. On the other hand, temperature coefficients are known to be quite sensitive to both the distance and geometry of the hydrogen bond, as well as small changes in the local electronic structure (Hong et al., 2013). Nevertheless, it is not necessary for the two factors to be correlated to reveal useful information about the strength of backbone hydrogen bonds. If an amide proton has a small temperature coefficient *and* is well-protected from exchange with solvent, it is more than reasonable to conclude that the hydrogen bond is present and strong. In the case of a thermostable protein, the presence of a strong, stable hydrogen-bond is indicative of the rigidity and stability of the local structure.

Temperature coefficients were determined for 255 backbone amides of Dbh, with 200 of those amides involved in main-chain hydrogen bonds. The average value of the

temperature coefficient of the residues of Dbh with detectable protection in the HDX experiment (-2.55 ppb/K) is lower than the average value of the temperature coefficient of the residues exhibiting fast exchange in the HDX experiment (-3.05 ppb/K). For backbone amides that are within hydrogen-bonding distance for main-chain hydrogen-bonds (in either the PDB or occupied >20% in the MD simulation), the average temperature coefficient is -2.55 ppb/K, a very similar result to the temperature coefficient for the exchange-protected set of amides. None of the main-chain bonded amides have a temperature coefficient more negative than -6.5 ppb/K. In contrast, the average of the temperature coefficient for amides which are not involved in backbone hydrogen bonds is 4.06 ppb/K. It is clear that many of the residues on flexible loops (for which NMR signals can be detected – many unassigned residues are located on the loop regions) have particularly negative temperature coefficients (e.g. 12F, -10.08 ppb/K; 49E, -11.95 ppb/K; 174D, -10.16 ppb/K; 259V -8.99 ppb/K).

However; not all of the backbone amides which are apparently not involved in main-chain hydrogen bonds have large negative temperature coefficients (e.g. 28L, -2.44 ppb/K; 139I, 0.23 ppb/K; 168R, -0.10 ppb/K; 309D 0.73 ppb/K). Many of these amides are located on short turns or at the beginning of $\alpha$-helices. Therefore, the value of the temperature coefficient is correlated with, yet is not necessarily predictive of, the intramolecular hydrogen bonding state of the amide proton. Nevertheless, for amides which have very negative temperature coefficients (in other words, a high sensitivity of the amide), it is reasonable to conclude that the amide is not involved in an intramolecular hydrogen bond. In conclusion, amide temperature coefficients alone are

not sufficient to determine intramolecular hydrogen bonding, but are informative in combination with other experiments that can also probe the state of intramolecular hydrogen bonding.

***Low-temperature biophysical experiments***

*Hydrogen-deuterium exchange NMR*

If Dbh is undergoing partial unfolding at lower temperatures, HDX-NMR is capable of detecting changes in the structure of Dbh. Peaks that are visible in the 15 minute HDX spectrum at 35°C represent structured regions of Dbh protected from exchange. If these peaks are missing in 18 hour cold-incubated spectrum, then it is fair to conclude that there is structural opening of unfolding occurring in that region of Dbh. The intrinsic rate of amide hydrogen exchange is ~21 times lower at 4°C, compared to 35°C (see Ch. 3, Eq. 7 for rate correction, (Bai et al., 1993). The first time point collected at 35°C was after ~15 minutes, and the spectrum for the cold-incubated sample was collected after ~18 hours. However, no significant difference was observed between the peaks in the two spectra (Figure 6.3), as almost all of the peaks visible in the 35°C spectrum are present in the cold-incubated 35°C spectrum. Therefore, the HDX data indicates that there is no unfolding event at 4°C which exposes amides that are protected at 35°C.

*Circular Dichroism*

CD spectroscopy is sensitive to changes in the secondary structure of proteins, and can discriminate between α-helical, β-sheet, and random coil structures. Therefore, if Dbh partially or completely unfolds at lower temperatures, the CD spectrum should shift from a mixed β-sheet/α-helical CD signature (since the atomic structure of Dbh is known) to a more random coil signature. Random coil CD signatures should show a more positive

CD signal in the 205nm-230nm range, and a more negative signal in the 190nm-205nm range. Examining the CD spectra of Dbh taken at 2°C, 7°C, 35°C, and 65°C in Figure 6.3 shows no drastic changes within the secondary structure of Dbh within this large temperature range. However, there is some increase in signal in the 190nm-205nm, indicating an overall increase in helical structure. In the case of β-lactoglobulin, the cold denatured state showed an expanded, helical conformation when examined by CD spectroscopy (Yamada et al., 2005). The evidence does not suggest that Dbh denatures into a random coil at this temperature by CD spectroscopy. Therefore, it may be that Dbh partially transitions to alternative conformation that is more helical and expanded. Any expansion occurring could be detected by dynamic light scattering. Nevertheless, a change in the hydrogen bonding pattern of Dbh at lower temperature should have been detectable by HDX-NMR, and no significant change was observed.

*Dynamic Light Scattering*

Dynamic light scattering measures macromolecular particle size (hydrodynamic radius) by tracking the autocorrelation function of the light scattered by the solution over time. The rate of Brownian diffusion (slower for larger particles) affects how quickly the autocorrelation function decays and determines the particle size and relative number of particles in solution (Lorber et al., 2012) . Following a hypothetical unfolding event, denatured Dbh monomers should display an increase in size compared to compact, stably-folded monomers. Figure 6.5 displays the size distribution for Dbh at 5°C and 35°C; note that the large particles scatter far more light than the smaller particles (the intensity scales with a $r^6$ dependence). The hydrodynamic diameter of the Dbh monomer falls around 6-7nm, which is greater than what is predicted for hydrodynamic

theory perfectly spherical protein of 41kDa (~4.5nm) (Erickson, 2009). Given that Dbh has a rather extended Y-shape, a size greater than that predicted for a sphere is expected. However; the presence of larger particles in the solution (70-100nm) suggests impurities or oligomers that were able to pass through the 0.1µm PVDF filter. The high polydispersity index of the solution (averaging around 35% for most measurements) indicates a collection of states (broad monomer peak) and also the presence of oligomers or impurities. A desirable polydispersity index would be below <20%. There does not appear to be a significant difference in the two monomers at the two temperatures; however, a measurement with a small PDI (and sharper monomer distribution) would be needed to confirm this result. Therefore, a future DLS experiment should use a lower protein concentration of Dbh, ideally as low as the instrument can get an accurate reading, to prevent oligomer formation and bring the PDI into an acceptable range.

*Differential Scanning Calorimetry*

DSC is capable of directly detecting unfolding or folding transitions in by precisely monitoring the change in heat capacity with temperature. The technique is typically used to monitor the thermal unfolding of proteins with temperature and to measure the change in heat capacity. In the case of Dbh, DSC was used to investigate an unfolding transition upon a *decrease* in temperature. There is little published data on DSC downscans to monitor transitions upon temperature decreases, usually using denaturants to achieve cold denaturation at higher temperature (Azuaga et al., 1992; Griko and Kutyshenko, 1994; Romero-Romero et al., 2011).This complicates the interpretation of downscan curves, whereas interpretation of upscans is more clear. For

instance, the shape of the curve in the region from 70°C to 30°C in Figure 6.6 (bottom) and Figure 6.7 (bottom) appears to be normal for the instrument, since water/water downscans in the instrument manual show this shape in this region as well. For this reason, the downscan of Dbh was compared to a downscan hen egg white lysozyme. There are published curves of upscans of lysozyme; therefore, I took an upscan to verify the instrument settings. The upscan of lysozyme (Figure 6.7, bottom) shows a thermal transition at approximately 74°C, which matches the published denaturation temperature for lysozyme (Cooper et al., 2000). The downscan of Dbh (Figure 6.6) shows a peak from approximately 25°C to 5°C (indicating the calorimeter needs to remove less heat from the sample cell compared to reference, which could due to the protein absorbing extra heat due to unfolding), which is not present in the downscan of lysozyme. However, the peak is very broad, and overall the downscan of Dbh is U-shaped. In contrast, the Dbh and lysozyme upscans show a linear increase with increasing temperature, and the lysozyme transition peak is quite sharp, typical of cooperative thermal unfolding events. On the other hand, cold denaturation can result in transition to a compact, but structurally heterogeneous ensemble of states where the efficiency of hydrophobic packing is disrupted (Sadqi et al., 2009; Jaremko et al., 2013). Broader transition peaks for cold denaturation have been observed for other proteins, and may be a result of the unfolding intermediates becoming kinetically trapped at lower temperature (Romero-Romero et al., 2011). A scan rate adjustment can reveal if there is a kinetic barrier to the transition – lower scans rates will cause the peak to be more sharp (Romero-Romero et al., 2011). A future DSC experiment on Dbh could use a lower scan rate to reveal any changes the transition peak. Also, the use of

cryoprotectants and an adjustment on the instrument minimum temperature would allow access to temperatures below 1°C. These experiments could allow further interpretation of the nature of the transition.

*Conclusions*

Based on the results of the above experiments, it is difficult to reach a definitive conclusion about the presence of cold denaturation of Dbh. The observation of that Dbh does not crystallize well 4°C, and the altered HSQC at 5°C indicate significant conformational heterogeneity. However, this apparent conformational heterogeneity did not alter the HDX HSQC spectrum at 35°C after incubation for 18 hours at 4°C. This indicates that the secondary structure and hydrophobic core of Dbh is not significant disrupted at low temperature. It was noted in a cold denaturation study of ubiquitin that the residual hydrophobic core in cold-denatured ubiquitin resembled the native core as determined by HDX (Sivaraman et al., 2001; Babu et al., 2004). This is confirmed by the observation by CD that Dbh actually appears to have increasing helical content at lower temperatures. Increased helical content has also been observed by CD in cold denaturation (Yamada et al., 2005). Nevertheless, the DSC measurement appears to indicate some form of structural transition, albeit a gradual, second-order process, as the cold transition peak is broad and flattened (Romero-Romero et al., 2011). In a recent study on CylR2, it was observed that there were only small changes in CD signal to -8°C; while the NMR spectra indicated that the structure remained compact, non-native tertiary contacts and increased backbone dynamics were observed (Jaremko et al., 2013). Similar observations have been made by NMR for ubiquitin, where a non-cooperative unfolding process yields an ensemble of states, with regions of the protein

possessing compact, native-like structure, and other regions with increased water penetration into the structure (Babu et al., 2004; Whitten et al., 2006). In contrast, thermal denaturation typically follows a cooperative, two-state process, and usually yields expanded and disordered structures, without retaining a hydrophobic core (Dobson et al., 1998).

In summary, the DSC experiment indicated a slow transition below 20°C, with the CD spectra showing increased $\alpha$-helical structure. In contrast, the HDX-NMR experiment incubated at 4°C was essentially identical to the HDX-NMR at 35°C, and the DLS experiment was inconclusive. The data suggests that Dbh may be undergoing a cold denaturation process at temperatures below 20°C, but retains compact with native-like levels of secondary structure. Further experiments will be necessary to observe disruption of native structure and formation of non-native tertiary contacts and water penetration, possibly using the approach of Jaremko et. al. for CylR2; they explored well-resolved changes in $^{1}$H-$^{1}$H NOE contacts (Jaremko et al., 2013). Ultimately, the characterization of the possible cold denatured state of Dbh merits further study, as it would be a rare case of a protein that undergoes this process above zero °C.

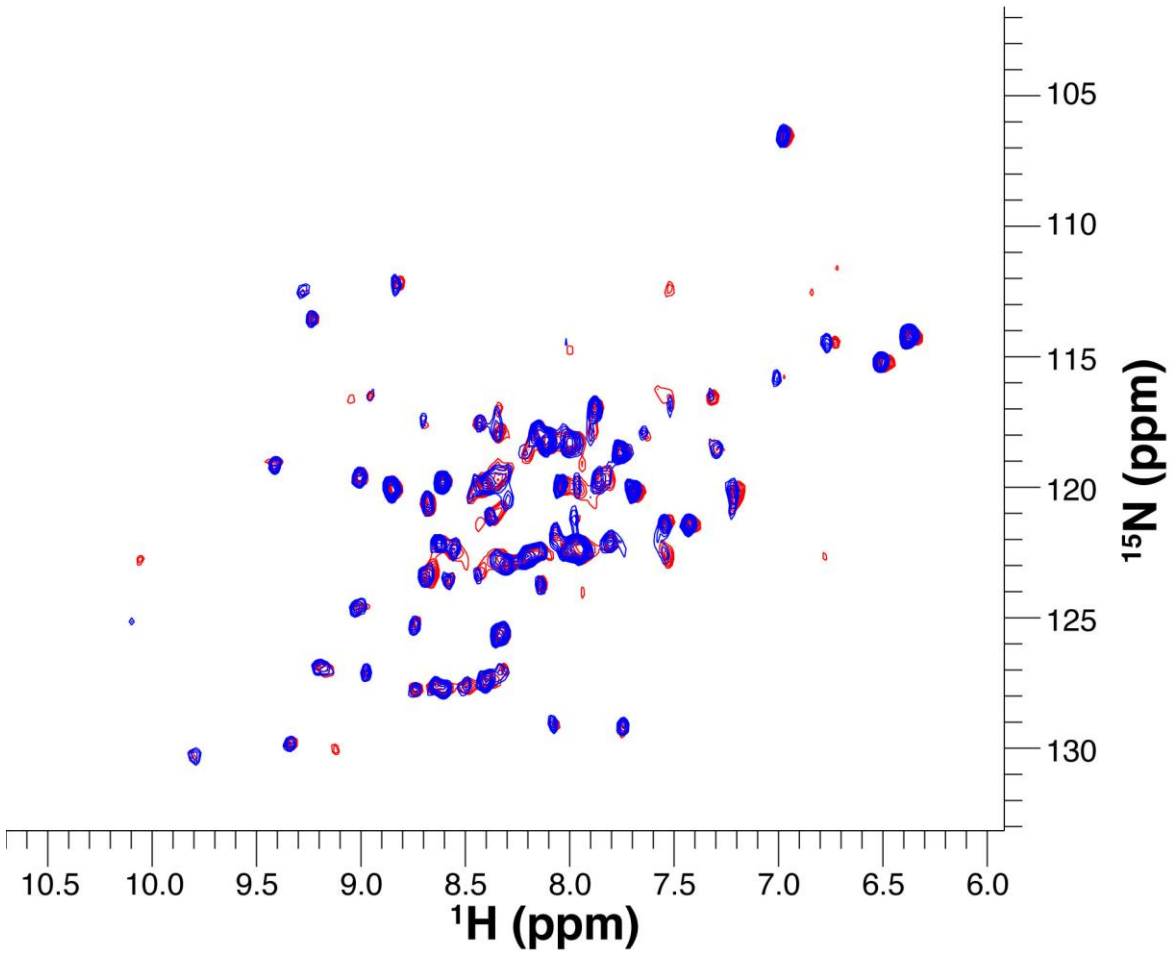



**Figure 6.3:** Overlay of hydrogen deuterium exchange $^{15}$N-$^{1}$H HSQC at 35°C (red) after 18hr incubation at 4°C (red) and 35°C ~15min after reconstitution from lyophilized protein (blue). No significant differences were observed between the two spectra, indicating that incubation at 4°C does not expose protected residues to solvent exchange via a denaturation process. The intrinsic exchange amide exchange rate is ~21 times lower at 4°C than 35°C, but the incubation time at 4°C was 72 times as long as the dead time in the 35°C experiment.

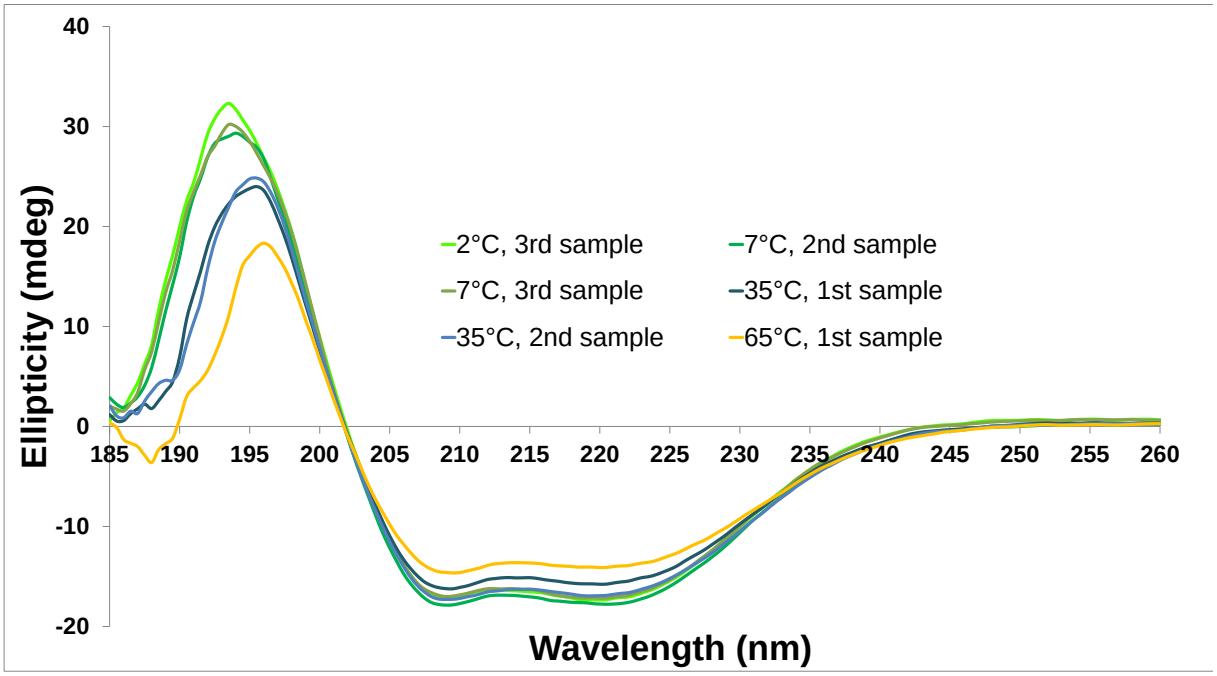



**Figure 6.4:** CD spectra of Dbh from 185 to 260nm at varying temperatures: 2°C (lime), 7°C (red, green), 35°C (blue, purple), and 65°C (orange). The increase in signal in the region from 190 to 200nm indicates an increase in α-helical structure. The double minimum at 208nm and 222nm also indicates a significant amount of helical structure.

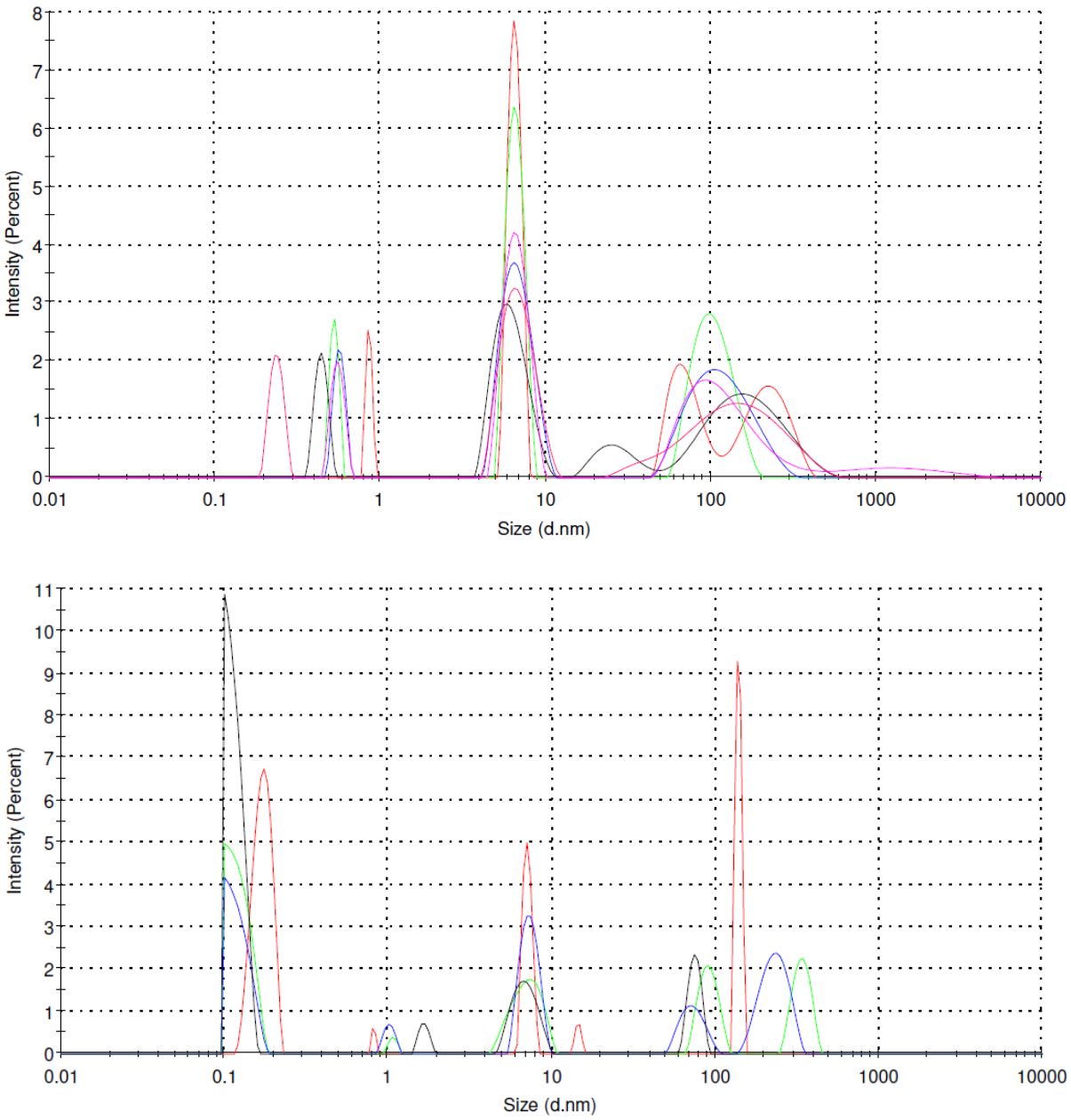



**Figure 6.5:** DLS measurements of Dbh (1.27mg/mL) at 5°C (top panel) and 35°C (bottom panel), by intensity. The size of Dbh monomer is between 6-7nm diameter. There is slight decrease in diameter at 5°C for the monomer. However, the presence of large particles around 100nm in the solution indicates a lower protein concentration should be used.

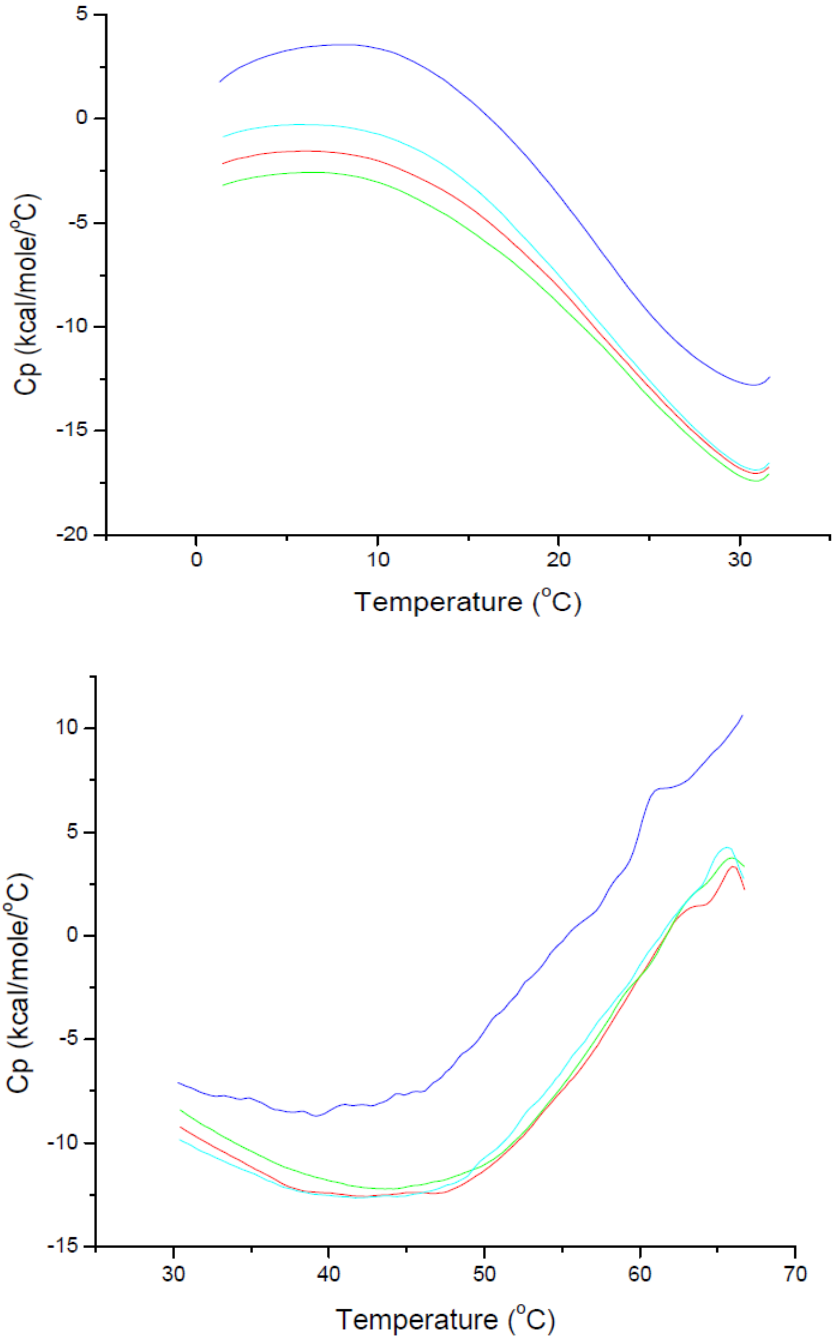



**Figure 6.6:** DSC downscan curves of Dbh from 35°C to 1°C (top panel) and from 70°C to 30°C. The first downscan in a series has a separate "thermal history" (that is, it is not preceded by a previous scan) which results in the curve shifting vertically. Three subsequent scans have identical thermal history, and were taken to test reversibility of any possible transition. The range from 70°C to 30°C resembles the curve for hen egg white lysozyme (Figure 6.7, top), but the rise from 25°C to the peak around 5°C differs from lysozyme, indicating a possible structural transition. It should be noted that DSC does not provide any information about what specific structural changes could be occurring.

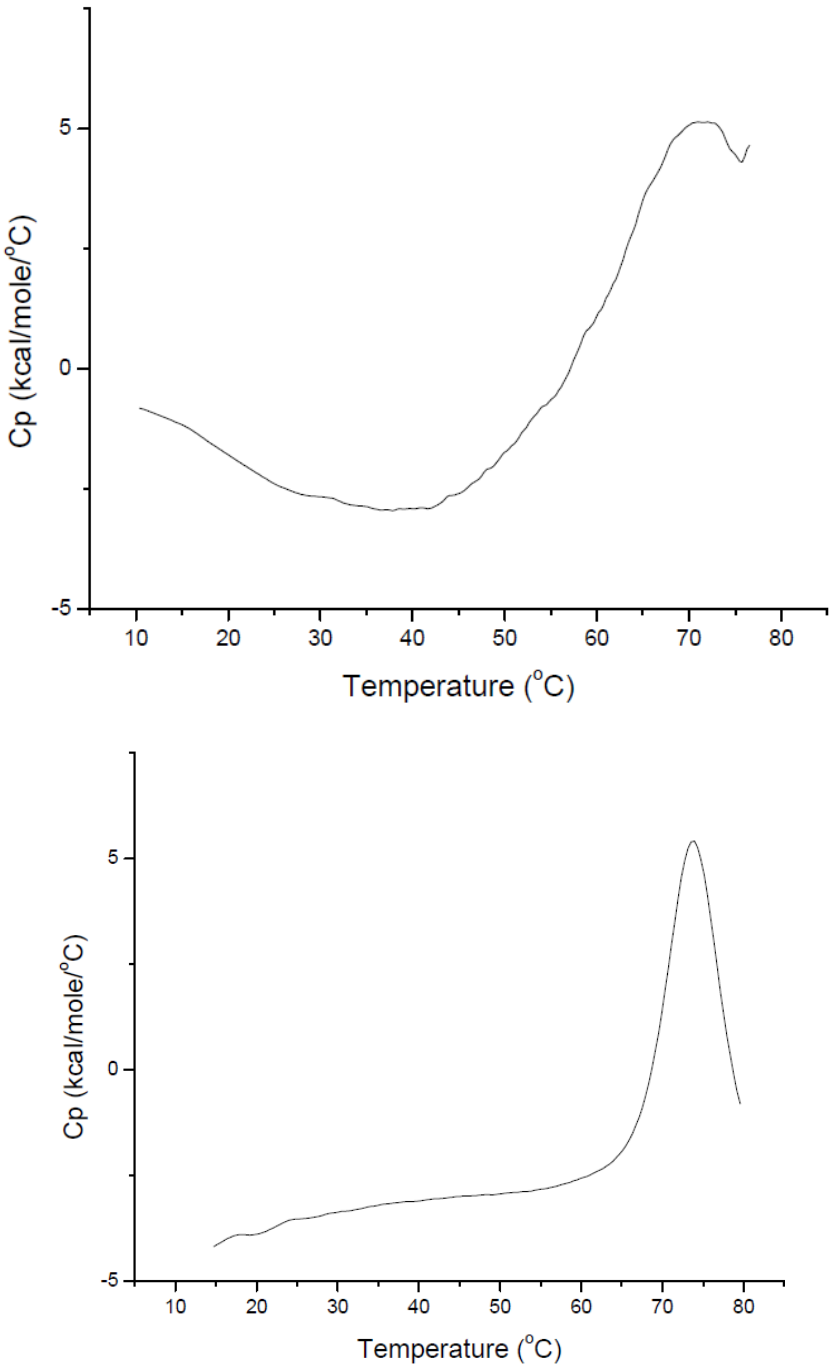



**Figure 6.7:** DSC downscan curves of hen egg white lysozyme from 75°C to 10°C (top panel) and an upscan (bottom panel) of HEWL from 10°C to 80°C, the maximum temperature achievable for this particular instrument. For the downscan, the curve does not rise as sharply between 25°C and 10°C as Dbh. The upscan shows the typical peak around the denaturation temperature (~74°C), which matches published values for HEWL. The $\Delta C_p$ is measured as the difference between the baseline value before and after the transition peak; unfortunately, measurements above 80°C were not possible for this instrument due to a malfunction in the pressure cap.

## References:


Azuaga, A.I., Galisteo, M.L., Mayorga, O.L., Cortijo, M., and Mateo, P.L. (1992). Heat and cold denaturation of beta-lactoglobulin B. FEBS Lett 309, 258-260.

Babu, C.R., Hilser, V.J., and Wand, A.J. (2004). Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. Nat Struct Mol Biol 11, 352-357.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. Proteins 17, 75-86.

Baxter, N.J., and Williamson, M.P. (1997). Temperature dependence of 1H chemical shifts in proteins. J Biomol NMR 9, 359-369.

Buchner, G.S., Shih, N., Reece, A.E., Niebling, S., and Kubelka, J. (2012). Unusual cold denaturation of a small protein domain. Biochemistry 51, 6496-6498.

Cierpicki, T., and Otlewski, J. (2001). Amide proton temperature coefficients as hydrogen bond indicators in proteins. J Biomol NMR 21, 249-261.

Cooper, A., Nutley, M.A., and Wadood, A. (2000). Differential scanning microcalorimetry in Protein-Ligand Interactions: hydrodynamics and calorimetry. (Oxford University Press).

Cordier, F., and Grzesiek, S. (2002). Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. J Mol Biol 317, 739-752.

Cordier, F., Nisius, L., Dingley, A.J., and Grzesiek, S. (2008). Direct detection of N-H[...]O=C hydrogen bonds in biomolecules by NMR spectroscopy. Nat Protoc 3, 235-241.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Dobson, C.M., Sali, A., and Karplus, M. (1998). Protein folding: A perspective from theory and experiment. Angew Chem Int Ed 37, 868-893.

Erickson, H.P. (2009). Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. In Biol Proced Online, pp. 32-51.

Griko , Y.V., and Kutyshenko, V.P. (1994). Differences in the processes of beta-lactoglobulin cold and heat denaturations. Biophys J 67, 356-363.

Griko, Y.V., and Privalov, P.L. (1992). Calorimetric study of the heat and cold denaturation of beta-lactoglobulin. Biochemistry 31, 8810-8815.

Hong, J., Jing, Q., and Yao, L. (2013). The protein amide (1)H(N) chemical shift temperature coefficient reflects thermal expansion of the N-H...O=C hydrogen bond. J Biomol NMR 55, 71-78.

Jaremko, M., Jaremko, L., Kim, H.Y., Cho, M.K., Schwieters, C.D., Giller, K., Becker, S., and Zweckstetter, M. (2013). Cold denaturation of a protein dimer monitored at atomic resolution. Nat Chem Biol 9, 264-270.

Lorber, B., Fischer, F., Bailly, M., Roy, H., and Kern, D. (2012). Protein analysis by dynamic light scattering: methods and techniques for students. Biochem Mol Biol Educ 40, 372-382.

Merutka, G., Dyson, H.J., and Wright, P.E. (1995). 'Random coil' 1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. J Biomol NMR 5, 14-24.

Pastore, A., Martin, S.R., Politou, A., Kondapalli, K.C., Stemmler, T., and Temussi, P.A. (2007). Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. J Am Chem Soc 129, 5374-5375.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94, 12366-12371.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605-1612.

Privalov, P.L. (1990). Cold denaturation of proteins. Crit Rev Biochem Mol Biol 25, 281-305.

Roe, D.R., and Cheatham III, T.E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput 9, 3084-3095.

Romero-Romero, M.L., Ingles-Prieto, A., Ibarra-Molero, B., and Sanchez-Ruiz, J.M. (2011). Highly anomalous energetics of protein cold denaturation linked to folding-unfolding kinetics. PLoS One 6, e23050.

Sadqi, M., de Alba, E., Perez-Jimenez, R., Sanchez-Ruiz, J.M., and Munoz, V. (2009). A designed protein as experimental model of primordial folding. Proc Natl Acad Sci U S A 106, 4127-4132.

Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nat Struct Biol 8, 984-989.

Sivaraman, T., Arrington, C.B., and Robertson, A.D. (2001). Kinetics of unfolding and folding from amide hydrogen exchange in native ubiquitin. Nat Struct Biol 8, 331-333.

Tilton, R.F., Jr., Dewan, J.C., and Petsko, G.A. (1992). Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K. Biochemistry 31, 2469-2481.

Tomlinson, J.H., and Williamson, M.P. (2012). Amide temperature coefficients in the protein G B1 domain. J Biomol NMR 52, 57-64.

Vajpai, N., Nisius, L., Wiktor, M., and Grzesiek, S. (2013). High-pressure NMR reveals close similarity between cold and alcohol protein denaturation in ubiquitin. Proc Natl Acad Sci U S A 110, E368-376.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59, 687-696.

Whitten, S.T., Kurtz, A.J., Pometun, M.S., Wand, A.J., and Hilser, V.J. (2006). Revealing the nature of the native state ensemble through cold denaturation. Biochemistry 45, 10163-10174.

Yamada, Y., Yajima, T., Fujiwara, K., Arai, M., Ito, K., Shimizu, A., Kihara, H., Kuwajima, K., Amemiya, Y., and Ikeguchi, M. (2005). Helical and expanded conformation of equine beta-lactoglobulin in the cold-denatured state. J Mol Biol 350, 338-348.

# Appendix A: HDX rates and protection factors for residues where signal could be detected for at least one temperature

Legend:

Stable = No decay detected over the course of the experiment
$k_{ex}$ = measured HDX rate
PF = Protection factor, for stable residues minimum possible protection factor is given
E = Exponent, base 10

**Observed Signals for HDX, 35°C**

| Res # | $k_{ex}$ | $k_{ex}$ fit error | PF | PF Error | If Stable, Min PF |
|---|---|---|---|---|---|
| **2Ile** | 1.43E-03 | 2.84E-04 | 6.63E+05 | 1.31E+05 | |
| **3Val** | Stable | | Stable | | 1.818E+07 |
| **4Ile** | Stable | | Stable | | 2.087E+07 |
| **5Phe** | Stable | | Stable | | 5.242E+07 |
| **6Val** | Stable | | Stable | | 3.544E+07 |
| **7Asp** | Stable | | Stable | | |
| **8Phe** | 1.27E-04 | 2.77E-05 | 8.03E+06 | 1.76E+06 | |
| **13Ala** | Stable | | Stable | | 1.776E+08 |
| **15Val** | 1.51E-04 | 1.53E-05 | 5.62E+06 | 5.69E+05 | |
| **18Val** | 9.07E-04 | 6.76E-05 | 4.17E+05 | 3.11E+04 | |
| **19Leu** | 4.31E-04 | 1.02E-04 | 1.18E+06 | 2.80E+05 | |
| **20Asn** | Stable | | Stable | | 2.948E+08 |
| **28Leu** | 2.87E-04 | 9.94E-05 | 1.41E+06 | 4.90E+05 | |
| **30Val** | 3.19E-04 | 4.95E-05 | 1.21E+06 | 1.88E+05 | |
| **50Ala** | 1.30E-03 | 7.45E-05 | 1.46E+06 | 8.35E+04 | |
| **51Arg** | Stable | | Stable | | 1.860E+08 |
| **55Val** | 9.38E-03 | 1.70E-03 | 8.42E+04 | 1.53E+04 | |
| **64Ala** | 1.08E-02 | 6.36E-03 | 3.28E+05 | 1.94E+05 | |
| **68Ala** | Stable | | Stable | | 9.110E+07 |
| **74Val** | 1.10E-04 | 1.40E-05 | 5.47E+06 | 6.98E+05 | |
| **84Phe** | 2.60E-03 | 5.49E-04 | 5.93E+05 | 1.26E+05 | |
| **85Ser** | 1.86E-03 | 4.14E-04 | 3.88E+06 | 8.67E+05 | |
| **88Ile** | 4.66E-05 | 2.31E-05 | 1.77E+07 | 8.80E+06 | |
| **89Met** | 1.49E-04 | 3.79E-05 | 1.03E+07 | 2.62E+06 | |
| **91Leu** | Stable | | Stable | | 8.502E+07 |
| **92Leu** | 2.57E-04 | 1.97E-05 | 1.69E+06 | 1.29E+05 | |
| **99Ile** | 5.67E-03 | 7.27E-04 | 1.16E+05 | 1.49E+04 | |
| **106Glu** | 1.23E-04 | 1.23E-05 | 4.45E+06 | 4.46E+05 | |
| **107Ala** | 6.17E-05 | 2.47E-05 | 3.07E+07 | 1.23E+07 | |
| **108Tyr** | Stable | | Stable | | 8.308E+07 |
| **109Leu** | Stable | | Stable | | 4.566E+07 |
| **122Gly** | 4.13E-03 | 1.83E-03 | 2.52E+06 | 1.12E+06 | |
| **124Glu** | 1.92E-04 | 4.71E-05 | 2.54E+06 | 6.21E+05 | |
| **127Arg** | 1.52E-03 | 1.37E-04 | 2.11E+06 | 1.90E+05 | |
| **128Lys** | Stable | | Stable | | 2.815E+07 |
| **130Lys** | Stable | | Stable | | 2.342E+08 |
| **132Glu** | 4.29E-03 | 7.77E-04 | 3.05E+05 | 5.53E+04 | |
| **133Ile** | 4.01E-05 | 3.01E-05 | 8.80E+06 | 6.61E+06 | |
| **134Leu** | Stable | | Stable | | 7.577E+07 |
| **135Glu** | 4.20E-04 | 3.17E-05 | 1.21E+06 | 9.16E+04 | |
| **137Glu** | 2.87E-03 | 3.89E-04 | 3.81E+05 | 5.17E+04 | |
| **138Lys** | 7.52E-03 | 2.98E-03 | 2.30E+05 | 9.09E+04 | |
| **139Ile** | 7.52E-03 | 1.18E-05 | 8.73E+04 | 1.37E+02 | |
| **142Thr** | Stable | | Stable | | 9.989E+07 |
| **143Val** | 6.58E-05 | 3.52E-05 | 1.29E+07 | 6.88E+06 | |
| **144Gly** | Stable | | Stable | | 9.539E+07 |
| **145Val** | Stable | | Stable | | 4.892E+07 |
| **146Ala** | Stable | | Stable | | 2.087E+08 |
| **152Ala** | Stable | | Stable | | 4.566E+07 |

| | | | | | |
|---|---|---|---|---|---|
| **155Ile** | Stable | | Stable | | 1.121E+08 |
| **156Ala** | Stable | | Stable | | 9.539E+07 |
| **167Ile** | 2.59E-04 | 3.85E-05 | 1.39E+06 | 2.06E+05 | |
| **175Phe** | 3.53E-03 | 9.99E-04 | 2.88E+05 | 8.14E+04 | |
| **178Glu** | 3.33E-03 | 8.60E-04 | 5.20E+05 | 1.34E+05 | |
| **184Ile** | 9.77E-03 | 1.63E-03 | 3.61E+04 | 6.01E+03 | |
| **194Arg** | 1.41E-03 | 7.52E-05 | 3.79E+06 | 2.02E+05 | |
| **195Leu** | Stable | | Stable | | 2.039E+07 |
| **201Gln** | 3.80E-04 | 6.90E-05 | 4.77E+06 | 8.66E+05 | |
| **206Ile** | 5.78E-03 | 1.67E-03 | 5.70E+04 | 1.64E+04 | |
| **223Ala** | 9.00E-03 | 2.54E-03 | 3.92E+05 | 1.11E+05 | |
| **226Leu** | 2.66E-04 | 9.67E-05 | 2.96E+06 | 1.08E+06 | |
| **262Ile** | 2.76E-04 | 1.51E-04 | 1.31E+06 | 7.17E+05 | |
| **265Tyr** | 3.75E-03 | 6.13E-04 | 2.20E+05 | 3.60E+04 | |
| **266Leu** | Stable | | Stable | | 2.087E+07 |
| **267Lys** | 2.07E-03 | 1.86E-03 | 7.27E+05 | 6.55E+05 | |
| **269Ala** | 3.89E-04 | 2.34E-05 | 9.08E+06 | 5.46E+05 | |
| **270Ile** | Stable | | Stable | | 8.700E+07 |
| **271Asn** | 1.47E-03 | 2.82E-04 | 3.32E+06 | 6.39E+05 | |
| **273Ala** | 4.41E-03 | 2.62E-03 | 4.29E+05 | 2.54E+05 | |
| **280Ile** | 1.29E-04 | 6.77E-05 | 5.72E+06 | 3.00E+06 | |
| **283Arg** | 5.15E-03 | 7.80E-04 | 8.05E+05 | 1.22E+05 | |
| **284Ile** | Stable | | Stable | | 4.261E+07 |
| **285Thr** | 1.16E-04 | 3.17E-05 | 1.16E+07 | 3.18E+06 | |
| **286Val** | Stable | | Stable | | 4.781E+07 |
| **287Ile** | Stable | | Stable | | 7.754E+07 |
| **288Ala** | Stable | | Stable | | 4.892E+07 |
| **289Ile** | Stable | | Stable | | 2.087E+07 |
| **290Met** | Stable | | Stable | | 9.110E+07 |
| **291Glu** | 4.20E-04 | 4.04E-05 | 2.54E+06 | 2.44E+05 | |
| **296Leu** | 1.01E-03 | 3.34E-04 | 4.10E+05 | 1.35E+05 | |
| **298Lys** | Stable | | Stable | | 6.159E+07 |
| **315Ala** | 1.06E-04 | 1.66E-05 | 1.82E+07 | 2.85E+06 | |
| **317Asp** | Stable | | Stable | | 2.815E+08 |
| **318Leu** | Stable | | Stable | | 1.121E+08 |
| **319Leu** | Stable | | Stable | | 5.489E+07 |
| **320Arg** | 3.96E-04 | 4.82E-05 | 5.01E+06 | 6.11E+05 | |
| **321Glu** | 6.33E-03 | 4.26E-04 | 2.17E+05 | 1.46E+04 | |
| **322Leu** | 3.01E-04 | 2.95E-05 | 1.66E+06 | 1.63E+05 | |
| **323Leu** | 1.04E-02 | 3.69E-03 | 4.17E+04 | 1.47E+04 | |
| **331Val** | 1.11E-03 | 8.74E-04 | 1.01E+06 | 7.97E+05 | |
| **333Arg** | 1.34E-04 | 7.13E-05 | 3.99E+07 | 2.13E+07 | |
| **335Gly** | Stable | | Stable | | 6.449E+07 |
| **337Lys** | Stable | | Stable | | 3.087E+08 |
| **339Asp** | 5.71E-03 | 1.55E-03 | 1.45E+05 | 3.94E+04 | |

## Observed Signals for HDX, 50°C

| Res # | $k_{ex}$ | $k_{ex}$ fit error | PF | PF Error | If Stable, Min PF |
|---|---|---|---|---|---|
| **3Val** | Stable | | Stable | | 9.879E+07 |
| **4Ile** | Stable | | Stable | | 1.134E+08 |
| **5Phe** | Stable | | Stable | | 2.849E+08 |
| **6Val** | Stable | | Stable | | 1.926E+08 |
| **7Asp** | Stable | | Stable | | 3.053E+08 |
| **8Phe** | 1.32E-03 | 1.10E-04 | 3.01E+06 | 2.51E+05 | |
| **19Leu** | 1.60E-04 | 8.74E-05 | 1.25E+07 | 6.83E+06 | |
| **89Met** | 5.62E-03 | 7.97E-04 | 1.07E+06 | 1.51E+05 | |
| **92Leu** | 1.31E-03 | 1.52E-04 | 1.29E+06 | 1.50E+05 | |
| **106Glu** | 1.32E-03 | 7.78E-05 | 1.62E+06 | 9.56E+04 | |
| **107Ala** | 9.45E-05 | 4.28E-05 | 7.81E+07 | 3.54E+07 | |
| **108Tyr** | 1.79E-03 | 2.54E-04 | 3.13E+06 | 4.45E+05 | |
| **109Leu** | 1.23E-04 | 9.35E-05 | 2.51E+07 | 1.91E+07 | |
| **124Glu** | 1.10E-03 | 2.80E-04 | 1.73E+06 | 4.40E+05 | |

| | | | | | |
|---|---|---|---|---|---|
| **127Arg** | 1.12E-02 | 6.92E-03 | 1.12E+06 | 6.95E+05 | |
| **128Lys** | Stable | | Stable | | 1.273E+09 |
| **129Ile** | Stable | | Stable | | 2.064E+08 |
| **130Lys** | 3.68E-04 | 2.43E-05 | 1.52E+07 | 1.00E+06 | |
| **133Ile** | Stable | | Stable | | 1.108E+08 |
| **134Leu** | 2.34E-04 | 3.80E-05 | 6.91E+06 | 1.12E+06 | |
| **135Glu** | 4.78E-03 | 4.78E-04 | 4.16E+05 | 4.16E+04 | |
| **139Ile** | 1.66E-04 | 1.98E-05 | 1.54E+07 | 1.84E+06 | |
| **142Thr** | 2.80E-04 | 4.46E-05 | 2.30E+07 | 3.67E+06 | |
| **143Val** | 6.56E-04 | 9.48E-05 | 5.02E+06 | 7.25E+05 | |
| **144Gly** | Stable | | Stable | | 1.134E+09 |
| **145Val** | Stable | | Stable | | 2.482E+08 |
| **146Ala** | Stable | | Stable | | 6.092E+08 |
| **152Ala** | Stable | | Stable | | 5.185E+08 |
| **155Ile** | Stable | | Stable | | 9.220E+07 |
| **156Ala** | Stable | | Stable | | 4.951E+08 |
| **167Ile** | 2.39E-04 | 5.33E-05 | 5.88E+06 | 1.31E+06 | |
| **195Leu** | Stable | | Stable | | 2.599E+08 |
| **266Leu** | 1.77E-04 | 4.82E-05 | 1.74E+07 | 4.76E+06 | |
| **269Ala** | 8.67E-04 | 9.80E-05 | 1.59E+07 | 1.79E+06 | |
| **270Ile** | Stable | | Stable | | 1.566E+08 |
| **284Ile** | 7.48E-04 | 1.16E-04 | 4.31E+06 | 6.67E+05 | |
| **285Thr** | 3.23E-03 | 2.63E-03 | 1.62E+06 | 1.31E+06 | |
| **286Val** | 8.90E-04 | 1.91E-04 | 3.70E+06 | 7.93E+05 | |
| **287Ile** | Stable | | Stable | | 1.134E+08 |
| **288Ala** | Stable | | Stable | | 4.951E+08 |
| **289Ile** | Stable | | Stable | | 1.566E+08 |
| **290Met** | Stable | | Stable | | 4.839E+08 |
| **291Glu** | 3.85E-03 | 1.84E-03 | 1.08E+06 | 5.16E+05 | |
| **315Ala** | 1.82E-03 | 1.68E-04 | 4.16E+06 | 3.86E+05 | |
| **318Leu** | 9.08E-04 | 1.09E-04 | 1.99E+06 | 2.40E+05 | |
| **319Leu** | Stable | | Stable | | 1.364E+08 |
| **320Arg** | 8.30E-03 | 4.78E-03 | 9.31E+05 | 5.36E+05 | |
| **322Leu** | 9.38E-03 | 3.14E-03 | 2.07E+05 | 6.93E+04 | |
| **335Gly** | Stable | | Stable | | 9.220E+08 |
| **337Lys** | Stable | | Stable | | 5.556E+08 |

# Appendix B: $T_1$ and $T_2$ rates, and $T_1/T_2$ ratios for Dbh at 35°C and 50°C

### Observed Signals for Spin Relaxation Parameters, 35°C

| Res # | $T_1$ (ms) | $T_1$ fit error (ms) | $T_2$(ms) | $T_2$ fit error (ms) | $T_1/T_2$ | $T_1/T_2$ error |
|---|---|---|---|---|---|---|
| **2IleN** | 1371 | 183 | 30.7 | 1.4 | 44.6 | 6.3 |
| **3ValN** | 1139 | 294 | 28.1 | 1.1 | 40.5 | 10.6 |
| **11PheN** | 1138 | 94 | 27.6 | 0.8 | 41.2 | 3.6 |
| **12PheN** | 580 | 187 | 37.6 | 1.5 | 15.4 | 5.0 |
| **13AlaN** | 1050 | 107 | 28.5 | 2.5 | 36.8 | 4.9 |
| **15ValN** | 962 | 98 | 24.3 | 1.0 | 39.5 | 4.3 |
| **18ValN** | 893 | 119 | 40.3 | 2.2 | 22.2 | 3.2 |
| **19LeuN** | 1045 | 138 | 24.1 | 1.4 | 43.4 | 6.2 |
| **20AsnN** | 1261 | 128 | 27.6 | 0.5 | 45.6 | 4.7 |
| **24LysN** | 1065 | 178 | 26.2 | 1.5 | 40.7 | 7.2 |
| **25GlyN** | 754 | 167 | 25.4 | 2.7 | 29.7 | 7.3 |
| **26LysN** | 876 | 79 | 28.9 | 1.1 | 30.3 | 3.0 |
| **28LeuN** | 842 | 69 | 27.4 | 1.2 | 30.8 | 2.9 |
| **29ValN** | 828 | 253 | 22.8 | 4.4 | 36.3 | 13.2 |
| **30ValN** | 867 | 95 | 30.1 | 4.4 | 28.8 | 5.3 |
| **32ValN** | 631 | 83 | 29.1 | 0.4 | 21.7 | 2.9 |
| **43ValN** | 1246 | 115 | 30.7 | 1.0 | 40.6 | 4.0 |
| **45ThrN** | 891 | 134 | 28.1 | 1.4 | 31.7 | 5.0 |
| **49GluN** | 871 | 176 | 24.5 | 3.7 | 35.6 | 9.0 |
| **50AlaN** | 935 | 184 | 37.1 | 1.8 | 25.2 | 5.1 |
| **51ArgN** | 1128 | 97 | 27.7 | 1.3 | 40.8 | 4.0 |
| **52LysN** | 883 | 80 | 26.9 | 2.0 | 32.9 | 3.9 |
| **53LeuN** | 1154 | 118 | 29.7 | 1.5 | 38.8 | 4.4 |
| **55ValN** | 1089 | 151 | 31.7 | 0.9 | 34.4 | 4.9 |
| **56LysN** | 1283 | 179 | 30.9 | 1.6 | 41.5 | 6.2 |
| **61IleN** | 1129 | 303 | 33.7 | 1.3 | 33.5 | 9.1 |
| **62IleN** | 735 | 35 | 27.5 | 0.8 | 26.7 | 1.5 |
| **63LysN** | 929 | 75 | 28.3 | 2.8 | 32.8 | 4.1 |
| **64AlaN** | 963 | 88 | 34.6 | 2.3 | 27.8 | 3.1 |
| **65MetN** | 946 | 82 | 34.0 | 0.4 | 27.8 | 2.4 |
| **66GlnN** | 1056 | 71 | 26.6 | 2.3 | 39.6 | 4.4 |
| **67IleN** | 1130 | 313 | 29.3 | 0.7 | 38.5 | 10.7 |
| **68AlaN** | 1109 | 133 | 32.1 | 0.6 | 34.6 | 4.2 |
| **72IleN** | 1161 | 113 | 39.1 | 1.5 | 29.7 | 3.1 |
| **73TyrN** | 1125 | 154 | 33.5 | 3.3 | 33.6 | 5.7 |
| **74ValN** | 1025 | 56 | 24.9 | 1.6 | 41.2 | 3.5 |
| **76MetN** | 1040 | 157 | 38.3 | 2.5 | 27.1 | 4.5 |
| **77ArgN** | 922 | 305 | 33.0 | 3.4 | 27.9 | 9.7 |
| **80IleN** | 928 | 101 | 36.1 | 7.5 | 25.7 | 6.0 |
| **81TyrN** | 1008 | 157 | 31.8 | 1.9 | 31.7 | 5.3 |
| **83AlaN** | 1365 | 137 | 26.7 | 1.0 | 51.2 | 5.5 |
| **84PheN** | 1055 | 108 | 23.5 | 1.0 | 44.9 | 5.0 |
| **85SerN** | 1529 | 271 | 33.1 | 4.8 | 46.2 | 10.6 |
| **86AsnN** | 1013 | 191 | 25.6 | 1.9 | 39.6 | 8.0 |
| **87ArgN** | 732 | 70 | 27.3 | 1.2 | 26.8 | 2.8 |
| **89MetN** | 1028 | 160 | 24.2 | 3.1 | 42.5 | 8.6 |
| **90AsnN** | 1260 | 239 | 23.0 | 0.6 | 54.7 | 10.5 |
| **91LeuN** | 1027 | 96 | 30.4 | 1.6 | 33.8 | 3.6 |
| **92LeuN** | 1053 | 149 | 25.2 | 1.5 | 41.8 | 6.4 |
| **93AsnN** | 1152 | 226 | 24.5 | 0.9 | 47.1 | 9.4 |
| **94LysN** | 694 | 47 | 27.3 | 0.5 | 25.4 | 1.8 |
| **95HisN** | 991 | 164 | 31.0 | 2.4 | 32.0 | 5.8 |
| **96AlaN** | 1119 | 123 | 31.5 | 0.6 | 35.5 | 4.0 |
| **97AspN** | 554 | 34 | 34.4 | 2.1 | 16.1 | 1.4 |
| **98LysN** | 1082 | 155 | 31.0 | 2.7 | 34.9 | 5.8 |
| **100 GluN** | 986 | 459 | 59.4 | 7.7 | 16.6 | 8.0 |
| **103SerN** | 745 | 71 | 27.8 | 1.9 | 26.8 | 3.1 |
| **105AspN** | 660 | 104 | 27.9 | 1.3 | 23.6 | 3.9 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **106GluN** | 795 | 69 | 28.0 | 1.6 | 28.4 | 2.9 |
| **107AlaN** | 869 | 69 | 31.1 | 1.1 | 28.0 | 2.4 |
| **108TyrN** | 922 | 240 | 28.0 | 1.9 | 32.9 | 8.8 |
| **109LeuN** | 675 | 288 | 27.1 | 2.6 | 24.9 | 10.9 |
| **110AspN** | 745 | 213 | 25.2 | 2.0 | 29.6 | 8.8 |
| **114LysN** | 633 | 46 | 33.9 | 2.7 | 18.7 | 2.0 |
| **115ValN** | 1071 | 108 | 36.2 | 1.0 | 29.6 | 3.1 |
| **118AsnN** | 1159 | 148 | 30.1 | 0.8 | 38.5 | 5.0 |
| **120GluN** | 608 | 21 | 28.7 | 1.1 | 21.2 | 1.1 |
| **121AsnN** | 1192 | 158 | 27.8 | 2.2 | 42.9 | 6.6 |
| **122GlyN** | 843 | 83 | 26.2 | 2.0 | 32.2 | 4.0 |
| **123IleN** | 982 | 251 | 31.4 | 2.7 | 31.3 | 8.4 |
| **124GluN** | 1492 | 348 | 23.9 | 1.4 | 62.4 | 15.0 |
| **125LeuN** | 951 | 78 | 31.7 | 2.0 | 30.0 | 3.1 |
| **126AlaN** | 529 | 42 | 30.2 | 1.4 | 17.5 | 1.6 |
| **127ArgN** | 1280 | 199 | 23.9 | 1.0 | 53.5 | 8.6 |
| **129IleN** | 1023 | 90 | 25.7 | 1.1 | 39.9 | 3.9 |
| **130LysN** | 1281 | 157 | 25.8 | 0.8 | 49.6 | 6.3 |
| **132GluN** | 1097 | 135 | 26.9 | 0.9 | 40.8 | 5.2 |
| **133IleN** | 709 | 47 | 24.3 | 0.9 | 29.2 | 2.2 |
| **134LeuN** | 1251 | 173 | 27.5 | 1.2 | 45.5 | 6.6 |
| **135GluN** | 1305 | 207 | 26.2 | 1.4 | 49.9 | 8.4 |
| **136LysN** | 1652 | 346 | 25.8 | 1.0 | 63.9 | 13.6 |
| **137GluN** | 1544 | 433 | 27.1 | 1.5 | 57.0 | 16.3 |
| **139IleN** | 1334 | 164 | 29.7 | 1.3 | 44.9 | 5.9 |
| **140ThrN** | 860 | 288 | 31.8 | 3.0 | 27.0 | 9.4 |
| **141ValN** | 887 | 147 | 35.0 | 2.4 | 25.3 | 4.5 |
| **142ThrN** | 794 | 108 | 31.2 | 1.1 | 25.5 | 3.6 |
| **143ValN** | 876 | 104 | 33.1 | 2.4 | 26.4 | 3.7 |
| **144GlyN** | 1733 | 352 | 29.7 | 1.3 | 58.4 | 12.1 |
| **145ValN** | 1034 | 170 | 29.9 | 3.2 | 34.6 | 6.8 |
| **146AlaN** | 941 | 257 | 31.0 | 2.0 | 30.3 | 8.5 |
| **154IleN** | 983 | 62 | 31.4 | 5.6 | 31.2 | 5.9 |
| **155IleN** | 895 | 53 | 27.9 | 0.9 | 32.0 | 2.1 |
| **157AspN** | 889 | 66 | 29.7 | 1.5 | 29.9 | 2.7 |
| **158LysN** | 834 | 93 | 29.0 | 0.5 | 28.8 | 3.2 |
| **159SerN** | 1074 | 248 | 28.7 | 4.0 | 37.4 | 10.1 |
| **160LysN** | 670 | 53 | 30.3 | 1.2 | 22.1 | 2.0 |
| **162AsnN** | 518 | 93 | 33.7 | 2.6 | 15.4 | 3.0 |
| **163GlyN** | 514 | 107 | 27.6 | 3.5 | 18.6 | 4.6 |
| **164LeuN** | 633 | 93 | 31.2 | 2.6 | 20.3 | 3.4 |
| **166ValN** | 742 | 114 | 33.8 | 1.4 | 21.9 | 3.5 |
| **167IleN** | 1205 | 155 | 29.3 | 2.0 | 41.2 | 6.0 |
| **168ArgN** | 659 | 69 | 35.6 | 2.8 | 18.5 | 2.4 |
| **171GluN** | 1061 | 255 | 38.1 | 3.4 | 27.8 | 7.1 |
| **172ValN** | 962 | 109 | 26.4 | 2.5 | 36.5 | 5.4 |
| **177AsnN** | 1076 | 154 | 23.9 | 0.6 | 44.9 | 6.5 |
| **178GluN** | 1007 | 110 | 29.3 | 1.0 | 34.4 | 3.9 |
| **179LeuN** | 825 | 49 | 27.5 | 1.0 | 30.0 | 2.1 |
| **180AspN** | 1076 | 207 | 31.5 | 1.3 | 34.1 | 6.7 |
| **181IleN** | 972 | 139 | 30.6 | 0.9 | 31.7 | 4.6 |
| **182AspN** | 1186 | 199 | 26.3 | 1.2 | 45.2 | 7.9 |
| **183GluN** | 2072 | 323 | 23.8 | 1.9 | 87.1 | 15.2 |
| **184IleN** | 1180 | 151 | 27.0 | 3.7 | 43.7 | 8.2 |
| **191LeuN** | 798 | 72 | 27.1 | 0.9 | 29.4 | 2.8 |
| **192AlaN** | 927 | 77 | 31.0 | 1.3 | 29.9 | 2.8 |
| **195LeuN** | 953 | 89 | 26.9 | 0.9 | 35.4 | 3.5 |
| **198LeuN** | 1213 | 221 | 25.7 | 0.5 | 47.2 | 8.6 |
| **199GlyN** | 1547 | 242 | 30.7 | 1.5 | 50.4 | 8.3 |
| **201GlnN** | 896 | 127 | 21.7 | 1.3 | 41.4 | 6.4 |
| **204ArgN** | 819 | 168 | 27.8 | 2.1 | 29.4 | 6.4 |
| **206IleN** | 876 | 149 | 21.9 | 2.0 | 39.9 | 7.7 |
| **207LeuN** | 836 | 128 | 24.1 | 2.8 | 34.7 | 6.7 |
| **213GluN** | 700 | 59 | 31.7 | 1.7 | 22.1 | 2.2 |
| **215GluN** | 951 | 57 | 27.7 | 0.3 | 34.3 | 2.1 |
| **216LysN** | 901 | 73 | 29.4 | 1.4 | 30.6 | 2.9 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **218ThrN** | 1188 | 264 | 30.3 | 1.7 | 39.3 | 9.0 |
| **219GlyN** | 937 | 82 | 24.5 | 1.1 | 38.3 | 3.7 |
| **222LysN** | 488 | 73 | 27.1 | 0.9 | 18.0 | 2.8 |
| **223AlaN** | 1027 | 96 | 32.1 | 1.4 | 32.0 | 3.3 |
| **224LeuN** | 1050 | 145 | 20.9 | 0.7 | 50.2 | 7.2 |
| **226LeuN** | 1070 | 156 | 28.7 | 0.8 | 37.2 | 5.5 |
| **229LeuN** | 1186 | 143 | 20.4 | 0.9 | 58.1 | 7.4 |
| **230AlaN** | 1090 | 233 | 25.2 | 1.5 | 43.3 | 9.6 |
| **231GlnN** | 1065 | 148 | 25.3 | 1.6 | 42.1 | 6.4 |
| **258AspN** | 1025 | 146 | 32.1 | 0.6 | 31.9 | 4.6 |
| **260LysN** | 735 | 35 | 27.9 | 0.9 | 26.3 | 1.5 |
| **261ValN** | 1041 | 129 | 29.2 | 0.6 | 35.7 | 4.5 |
| **262IleN** | 870 | 152 | 26.4 | 2.0 | 33.0 | 6.3 |
| **263LeuN** | 1098 | 86 | 30.7 | 1.0 | 35.7 | 3.0 |
| **265TyrN** | 742 | 67 | 29.1 | 0.8 | 25.5 | 2.4 |
| **266LeuN** | 1134 | 106 | 34.6 | 1.7 | 32.7 | 3.5 |
| **267LysN** | 1194 | 241 | 23.1 | 1.4 | 51.6 | 10.9 |
| **268LysN** | 1204 | 101 | 32.4 | 4.3 | 37.2 | 5.9 |
| **269AlaN** | 901 | 87 | 30.7 | 1.1 | 29.4 | 3.0 |
| **273AlaN** | 824 | 107 | 35.6 | 1.0 | 23.2 | 3.1 |
| **274TyrN** | 2156 | 396 | 26.5 | 2.9 | 81.3 | 17.4 |
| **275AsnN** | 1179 | 289 | 27.2 | 1.7 | 43.4 | 11.0 |
| **276LysN** | 742 | 67 | 28.3 | 0.8 | 26.3 | 2.5 |
| **277ValN** | 808 | 107 | 28.4 | 0.8 | 28.5 | 3.9 |
| **279GlyN** | 824 | 100 | 35.4 | 1.9 | 23.3 | 3.1 |
| **282MetN** | 968 | 249 | 20.4 | 4.0 | 47.4 | 15.4 |
| **283ArgN** | 807 | 92 | 32.3 | 2.2 | 25.0 | 3.3 |
| **284IleN** | 934 | 83 | 33.9 | 2.2 | 27.5 | 3.0 |
| **285ThrN** | 715 | 71 | 36.5 | 1.5 | 19.6 | 2.1 |
| **286ValN** | 659 | 54 | 27.6 | 2.1 | 23.9 | 2.7 |
| **287IleN** | 791 | 125 | 24.9 | 2.6 | 31.8 | 6.0 |
| **288AlaN** | 913 | 162 | 30.7 | 2.2 | 29.7 | 5.7 |
| **289IleN** | 671 | 79 | 38.3 | 1.3 | 17.5 | 2.2 |
| **290MetN** | 717 | 83 | 33.5 | 1.5 | 21.4 | 2.7 |
| **291GluN** | 882 | 139 | 26.8 | 3.4 | 32.9 | 6.7 |
| **294AspN** | 979 | 106 | 27.7 | 1.0 | 35.3 | 4.0 |
| **296LeuN** | 916 | 153 | 30.8 | 3.1 | 29.8 | 5.8 |
| **297SerN** | 732 | 69 | 34.0 | 0.4 | 21.6 | 2.1 |
| **298LysN** | 1182 | 198 | 33.8 | 1.4 | 35.0 | 6.0 |
| **300LysN** | 796 | 39 | 29.4 | 0.7 | 27.1 | 1.5 |
| **301LysN** | 661 | 48 | 35.1 | 2.3 | 18.8 | 1.8 |
| **302PheN** | 661 | 95 | 28.6 | 3.8 | 23.1 | 4.5 |
| **305GlyN** | 531 | 41 | 31.6 | 2.1 | 16.8 | 1.7 |
| **306IleN** | 882 | 94 | 29.7 | 1.0 | 29.7 | 3.3 |
| **307SerN** | 648 | 147 | 34.6 | 2.1 | 18.7 | 4.4 |
| **309AspN** | 860 | 69 | 31.4 | 1.3 | 27.4 | 2.5 |
| **310AsnN** | 776 | 53 | 27.8 | 1.4 | 27.9 | 2.4 |
| **311AlaN** | 833 | 79 | 33.1 | 1.9 | 25.2 | 2.8 |
| **312TyrN** | 1239 | 171 | 35.5 | 3.6 | 34.9 | 6.0 |
| **314ValN** | 758 | 85 | 27.1 | 1.5 | 28.0 | 3.5 |
| **315AlaN** | 959 | 66 | 31.1 | 0.6 | 30.8 | 2.2 |
| **316GluN** | 1336 | 161 | 33.2 | 1.3 | 40.3 | 5.1 |
| **318LeuN** | 705 | 80 | 35.6 | 1.1 | 19.8 | 2.3 |
| **321GluN** | 869 | 63 | 27.7 | 1.2 | 31.3 | 2.7 |
| **322LeuN** | 925 | 169 | 33.8 | 1.4 | 27.4 | 5.1 |
| **323LeuN** | 892 | 106 | 29.7 | 5.2 | 30.0 | 6.3 |
| **324ValN** | 894 | 97 | 27.4 | 1.2 | 32.7 | 3.8 |
| **325ArgN** | 769 | 49 | 30.1 | 0.8 | 25.6 | 1.8 |
| **330AsnN** | 625 | 109 | 32.0 | 1.1 | 19.5 | 3.5 |
| **331ValN** | 979 | 115 | 29.6 | 1.1 | 33.1 | 4.1 |
| **332ArgN** | 977 | 156 | 34.5 | 6.6 | 28.3 | 7.1 |
| **333ArgN** | 774 | 82 | 26.3 | 1.8 | 29.5 | 3.7 |
| **334IleN** | 957 | 219 | 32.5 | 1.0 | 29.5 | 6.8 |
| **335GlyN** | 1076 | 223 | 22.4 | 1.2 | 48.1 | 10.3 |
| **336ValN** | 998 | 145 | 30.2 | 1.3 | 33.1 | 5.0 |
| **337LysN** | 761 | 60 | 29.6 | 0.9 | 25.7 | 2.2 |

## Observed Signals for Spin Relaxation Parameters, 50°C

| Res # | $T_1$ (ms) | $T_1$ fit error (ms) | $T_2$(ms) | $T_2$ fit error (ms) | $T_1/T_2$ | $T_1/T_2$ error |
|---|---|---|---|---|---|---|
| **2IleN** | 656 | 36 | 48.3 | 2.7 | 13.6 | 1.1 |
| **3ValN** | 1339 | 187 | 42.8 | 1.9 | 31.3 | 4.6 |
| **7AspN** | 840 | 100 | 36.5 | 3.0 | 23.0 | 3.3 |
| **8PheN** | 887 | 86 | 48.3 | 5.2 | 18.4 | 2.7 |
| **9AspN** | 736 | 103 | 41.1 | 0.6 | 17.9 | 2.5 |
| **11PheN** | 1250 | 100 | 35.0 | 1.0 | 35.7 | 3.0 |
| **12PheN** | 920 | 61 | 38.5 | 2.2 | 23.9 | 2.1 |
| **13AlaN** | 1057 | 102 | 38.6 | 1.7 | 27.4 | 2.9 |
| **14GlnN** | 1075 | 152 | 45.1 | 0.5 | 23.8 | 3.4 |
| **15ValN** | 778 | 59 | 36.4 | 1.0 | 21.4 | 1.7 |
| **16GluN** | 1009 | 85 | 38.6 | 1.6 | 26.1 | 2.5 |
| **17GluN** | 1306 | 157 | 33.3 | 2.1 | 39.2 | 5.3 |
| **18ValN** | 1194 | 164 | 40.1 | 5.1 | 29.7 | 5.6 |
| **19LeuN** | 1174 | 114 | 36.1 | 1.0 | 32.5 | 3.3 |
| **20AsnN** | 894 | 29 | 42.0 | 0.8 | 21.3 | 0.8 |
| **22GlnN** | 1052 | 106 | 36.4 | 4.1 | 28.9 | 4.4 |
| **23TyrN** | 812 | 63 | 42.8 | 3.7 | 19.0 | 2.2 |
| **24LysN** | 849 | 34 | 41.6 | 1.3 | 20.4 | 1.0 |
| **25GlyN** | 450 | 91 | 31.8 | 3.5 | 14.2 | 3.3 |
| **26LysN** | 1055 | 38 | 40.7 | 2.4 | 25.9 | 1.8 |
| **30ValN** | 1059 | 101 | 43.1 | 2.5 | 24.6 | 2.7 |
| **31SerN** | 1072 | 110 | 36.9 | 1.4 | 29.1 | 3.2 |
| **32ValN** | 632 | 59 | 44.8 | 1.5 | 14.1 | 1.4 |
| **43ValN** | 1077 | 58 | 42.5 | 2.4 | 25.3 | 2.0 |
| **44AlaN** | 1114 | 185 | 33.5 | 3.2 | 33.3 | 6.4 |
| **45ThrN** | 1123 | 85 | 44.7 | 2.8 | 25.2 | 2.5 |
| **51ArgN** | 1386 | 186 | 34.9 | 1.6 | 39.7 | 5.6 |
| **52LysN** | 1105 | 69 | 40.4 | 3.2 | 27.3 | 2.8 |
| **53LeuN** | 1256 | 68 | 45.5 | 1.8 | 27.6 | 1.8 |
| **54GlyN** | 1009 | 72 | 40.9 | 1.5 | 24.6 | 2.0 |
| **55ValN** | 1016 | 88 | 41.1 | 3.4 | 24.7 | 3.0 |
| **56LysN** | 1227 | 149 | 47.9 | 1.1 | 25.6 | 3.2 |
| **58GlyN** | 900 | 135 | 42.3 | 4.1 | 21.3 | 3.8 |
| **59MetN** | 1065 | 81 | 43.0 | 1.9 | 24.7 | 2.2 |
| **61IleN** | 799 | 136 | 44.2 | 2.0 | 18.1 | 3.2 |
| **62IleN** | 789 | 25 | 42.0 | 0.9 | 18.8 | 0.7 |
| **63LysN** | 986 | 58 | 42.0 | 2.4 | 23.5 | 1.9 |
| **64AlaN** | 1449 | 123 | 45.3 | 3.4 | 32.0 | 3.6 |
| **65MetN** | 1413 | 115 | 40.3 | 2.2 | 35.1 | 3.4 |
| **66GlnN** | 953 | 91 | 41.6 | 2.6 | 22.9 | 2.6 |
| **67IleN** | 1022 | 133 | 45.6 | 3.6 | 22.4 | 3.4 |
| **68AlaN** | 976 | 27 | 46.6 | 1.1 | 21.0 | 0.8 |
| **71AlaN** | 993 | 69 | 39.1 | 0.9 | 25.4 | 1.9 |
| **72IleN** | 887 | 167 | 46.0 | 2.9 | 19.3 | 3.8 |
| **73TyrN** | 1131 | 118 | 41.8 | 0.8 | 27.1 | 2.9 |
| **76MetN** | 1188 | 126 | 45.1 | 0.5 | 26.3 | 2.8 |
| **77ArgN** | 1127 | 149 | 42.2 | 0.5 | 26.7 | 3.6 |
| **78LysN** | 1039 | 132 | 46.5 | 1.0 | 22.3 | 2.9 |
| **80IleN** | 905 | 65 | 41.7 | 2.5 | 21.7 | 2.0 |
| **81TyrN** | 1182 | 150 | 43.4 | 1.4 | 27.2 | 3.6 |
| **82GluN** | 541 | 56 | 28.7 | 2.3 | 18.8 | 2.4 |
| **83AlaN** | 1287 | 120 | 38.3 | 1.0 | 33.6 | 3.3 |
| **84PheN** | 1188 | 126 | 45.1 | 0.5 | 26.3 | 2.8 |
| **85SerN** | 1118 | 110 | 40.6 | 2.7 | 27.6 | 3.3 |
| **86AsnN** | 1052 | 86 | 35.8 | 1.2 | 29.4 | 2.6 |
| **87ArgN** | 1129 | 140 | 34.9 | 0.5 | 32.3 | 4.0 |
| **88IleN** | 1387 | 134 | 37.2 | 0.8 | 37.3 | 3.7 |
| **89MetN** | 1575 | 308 | 31.3 | 2.2 | 50.3 | 10.5 |
| **90AsnN** | 1354 | 155 | 39.7 | 1.4 | 34.1 | 4.1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **91LeuN** | 1134 | 64 | 41.8 | 2.5 | 27.1 | 2.2 |
| **92LeuN** | 1178 | 198 | 35.8 | 1.9 | 32.9 | 5.8 |
| **94LysN** | 554 | 39 | 42.5 | 2.2 | 13.0 | 1.1 |
| **95HisN** | 1086 | 156 | 53.5 | 3.4 | 20.3 | 3.2 |
| **96AlaN** | 1333 | 85 | 44.3 | 0.7 | 30.1 | 2.0 |
| **97AspN** | 487 | 61 | 38.0 | 1.4 | 12.8 | 1.7 |
| **98LysN** | 974 | 116 | 38.0 | 1.1 | 25.6 | 3.1 |
| **99IleN** | 949 | 77 | 53.3 | 5.8 | 17.8 | 2.4 |
| **100GluN** | 1135 | 535 | 31.9 | 1.4 | 35.6 | 16.8 |
| **101ValN** | 463 | 59 | 40.0 | 4.3 | 11.6 | 1.9 |
| **103SerN** | 910 | 81 | 42.1 | 1.8 | 21.6 | 2.1 |
| **105AspN** | 792 | 73 | 34.1 | 3.3 | 23.2 | 3.1 |
| **106GluN** | 982 | 172 | 41.1 | 4.3 | 23.9 | 4.9 |
| **107AlaN** | 1002 | 99 | 36.8 | 2.3 | 27.2 | 3.2 |
| **109LeuN** | 811 | 267 | 70.2 | 21.8 | 11.5 | 5.2 |
| **110AspN** | 691 | 54 | 42.6 | 1.5 | 16.2 | 1.4 |
| **111ValN** | 884 | 140 | 36.4 | 1.5 | 24.3 | 4.0 |
| **114LysN** | 943 | 43 | 35.0 | 0.6 | 26.9 | 1.3 |
| **115ValN** | 851 | 68 | 46.2 | 1.5 | 18.4 | 1.6 |
| **116GluN** | 711 | 73 | 35.4 | 2.7 | 20.0 | 2.6 |
| **117GlyN** | 720 | 84 | 37.7 | 2.0 | 19.1 | 2.5 |
| **118AsnN** | 679 | 62 | 41.7 | 5.0 | 16.3 | 2.5 |
| **120GluN** | 443 | 44 | 43.2 | 3.1 | 10.3 | 1.2 |
| **121AsnN** | 838 | 56 | 42.5 | 3.5 | 19.7 | 2.1 |
| **122GlyN** | 907 | 119 | 37.9 | 1.9 | 23.9 | 3.4 |
| **123IleN** | 1303 | 184 | 39.6 | 3.9 | 32.9 | 5.7 |
| **124GluN** | 1084 | 128 | 37.4 | 1.6 | 29.0 | 3.7 |
| **125LeuN** | 1160 | 40 | 38.3 | 1.1 | 30.3 | 1.4 |
| **126AlaN** | 1583 | 112 | 39.9 | 1.0 | 39.6 | 3.0 |
| **127ArgN** | 1584 | 367 | 37.5 | 0.8 | 42.2 | 9.8 |
| **128LysN** | 1437 | 102 | 41.5 | 2.3 | 34.7 | 3.1 |
| **129IleN** | 1083 | 90 | 36.6 | 1.6 | 29.6 | 2.8 |
| **132GluN** | 1312 | 189 | 36.6 | 1.6 | 35.9 | 5.4 |
| **133IleN** | 1561 | 189 | 39.4 | 2.5 | 39.6 | 5.4 |
| **134LeuN** | 1622 | 175 | 33.7 | 2.8 | 48.1 | 6.5 |
| **135GluN** | 1141 | 76 | 34.9 | 1.0 | 32.7 | 2.4 |
| **136LysN** | 1210 | 124 | 40.6 | 0.5 | 29.8 | 3.1 |
| **137GluN** | 1219 | 111 | 41.5 | 2.5 | 29.4 | 3.2 |
| **138LysN** | 1364 | 83 | 39.0 | 2.7 | 35.0 | 3.2 |
| **139IleN** | 1100 | 127 | 39.7 | 1.5 | 27.7 | 3.4 |
| **140ThrN** | 958 | 169 | 35.1 | 2.1 | 27.3 | 5.1 |
| **141ValN** | 951 | 87 | 43.0 | 1.2 | 22.1 | 2.1 |
| **142ThrN** | 789 | 62 | 37.1 | 3.0 | 21.3 | 2.4 |
| **143ValN** | 913 | 77 | 51.6 | 2.7 | 17.7 | 1.8 |
| **144GlyN** | 875 | 83 | 37.3 | 0.9 | 23.4 | 2.3 |
| **145ValN** | 1234 | 184 | 43.1 | 1.7 | 28.6 | 4.4 |
| **146AlaN** | 844 | 145 | 43.8 | 1.8 | 19.3 | 3.4 |
| **150IleN** | 1496 | 204 | 36.1 | 1.2 | 41.5 | 5.8 |
| **151LeuN** | 1099 | 116 | 32.9 | 0.8 | 33.4 | 3.6 |
| **152AlaN** | 996 | 120 | 32.1 | 0.8 | 31.0 | 3.8 |
| **153LysN** | 937 | 51 | 34.1 | 1.8 | 27.5 | 2.1 |
| **155IleN** | 783 | 62 | 36.3 | 1.0 | 21.6 | 1.8 |
| **156AlaN** | 1046 | 97 | 29.9 | 2.4 | 35.0 | 4.3 |
| **157AspN** | 1408 | 82 | 34.6 | 3.4 | 40.7 | 4.6 |
| **158LysN** | 774 | 91 | 51.4 | 3.4 | 15.1 | 2.0 |
| **159SerN** | 1030 | 118 | 37.5 | 1.7 | 27.5 | 3.4 |
| **160LysN** | 858 | 59 | 40.2 | 4.1 | 21.3 | 2.6 |
| **162AsnN** | 423 | 32 | 43.1 | 2.8 | 9.8 | 1.0 |
| **163GlyN** | 620 | 65 | 42.8 | 3.5 | 14.5 | 1.9 |
| **165GlyN** | 1221 | 219 | 39.9 | 1.5 | 30.6 | 5.6 |
| **166ValN** | 772 | 64 | 45.2 | 1.4 | 17.1 | 1.5 |
| **167IleN** | 1085 | 154 | 31.8 | 1.6 | 34.2 | 5.1 |
| **168ArgN** | 967 | 165 | 37.5 | 2.9 | 25.8 | 4.8 |
| **171GluN** | 862 | 98 | 42.0 | 1.1 | 20.5 | 2.4 |
| **172ValN** | 1122 | 174 | 46.1 | 3.8 | 24.4 | 4.3 |
| **173GlnN** | 774 | 48 | 48.6 | 1.9 | 15.9 | 1.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **174AspN** | 1111 | 68 | 36.2 | 0.3 | 30.7 | 1.9 |
| **175PheN** | 978 | 26 | 38.0 | 0.3 | 25.8 | 0.7 |
| **176LeuN** | 1040 | 157 | 30.1 | 1.8 | 34.5 | 5.6 |
| **177AsnN** | 1104 | 61 | 37.5 | 0.8 | 29.4 | 1.8 |
| **178GluN** | 1022 | 61 | 38.7 | 0.7 | 26.4 | 1.6 |
| **180AspN** | 1008 | 107 | 40.0 | 1.5 | 25.2 | 2.8 |
| **181IleN** | 865 | 134 | 35.8 | 1.7 | 24.1 | 3.9 |
| **182AspN** | 1306 | 77 | 39.4 | 1.5 | 33.1 | 2.3 |
| **183GluN** | 948 | 159 | 34.2 | 3.7 | 27.7 | 5.5 |
| **184IleN** | 1217 | 84 | 43.2 | 1.0 | 28.2 | 2.1 |
| **187IleN** | 657 | 49 | 49.2 | 2.1 | 13.4 | 1.1 |
| **191LeuN** | 947 | 76 | 40.3 | 0.6 | 23.5 | 1.9 |
| **192AlaN** | 1212 | 94 | 37.4 | 1.7 | 32.4 | 2.9 |
| **193ArgN** | 1073 | 138 | 34.0 | 0.2 | 31.6 | 4.1 |
| **196AsnN** | 1011 | 63 | 45.8 | 4.5 | 22.1 | 2.6 |
| **197GluN** | 1105 | 69 | 43.8 | 4.2 | 25.2 | 2.9 |
| **198LeuN** | 1274 | 227 | 38.1 | 1.9 | 33.4 | 6.2 |
| **199GlyN** | 1024 | 54 | 38.7 | 2.1 | 26.5 | 2.0 |
| **200IleN** | 1078 | 106 | 44.5 | 1.1 | 24.2 | 2.5 |
| **201GlnN** | 1054 | 68 | 38.1 | 2.0 | 27.7 | 2.3 |
| **202LysN** | 852 | 83 | 36.7 | 4.3 | 23.2 | 3.5 |
| **203LeuN** | 919 | 89 | 36.6 | 1.5 | 25.1 | 2.6 |
| **205AspN** | 971 | 28 | 40.0 | 2.5 | 24.3 | 1.7 |
| **206IleN** | 1123 | 146 | 37.7 | 1.7 | 29.8 | 4.1 |
| **207LeuN** | 1413 | 171 | 38.4 | 2.4 | 36.8 | 5.0 |
| **213GluN** | 889 | 33 | 41.8 | 0.8 | 21.3 | 0.9 |
| **214LeuN** | 860 | 51 | 44.5 | 4.7 | 19.3 | 2.3 |
| **215GluN** | 876 | 66 | 37.8 | 1.1 | 23.2 | 1.9 |
| **216LysN** | 841 | 48 | 37.6 | 0.6 | 22.3 | 1.3 |
| **217IleN** | 994 | 106 | 34.5 | 1.6 | 28.8 | 3.3 |
| **218ThrN** | 947 | 99 | 43.4 | 1.9 | 21.8 | 2.5 |
| **219GlyN** | 645 | 41 | 33.5 | 2.1 | 19.2 | 1.7 |
| **223AlaN** | 856 | 37 | 39.8 | 2.4 | 21.5 | 1.6 |
| **224LeuN** | 917 | 59 | 35.1 | 1.5 | 26.2 | 2.0 |
| **225TyrN** | 679 | 62 | 36.9 | 1.2 | 18.4 | 1.8 |
| **226LeuN** | 1661 | 383 | 36.0 | 2.7 | 46.1 | 11.2 |
| **228LysN** | 696 | 39 | 40.0 | 1.5 | 17.4 | 1.2 |
| **229LeuN** | 865 | 134 | 35.8 | 1.7 | 24.1 | 3.9 |
| **231GlnN** | 740 | 54 | 36.5 | 1.2 | 20.3 | 1.6 |
| **250LeuN** | 1142 | 227 | 47.8 | 2.8 | 23.9 | 5.0 |
| **252LeuN** | 946 | 102 | 44.6 | 0.5 | 21.2 | 2.3 |
| **254TyrN** | 603 | 29 | 41.3 | 1.9 | 14.6 | 1.0 |
| **258AspN** | 976 | 27 | 45.6 | 1.7 | 21.4 | 1.0 |
| **259ValN** | 298 | 101 | 37.1 | 2.6 | 8.0 | 2.8 |
| **260LysN** | 845 | 92 | 36.7 | 1.4 | 23.0 | 2.6 |
| **261ValN** | 1026 | 52 | 45.0 | 2.4 | 22.8 | 1.7 |
| **262IleN** | 904 | 121 | 50.9 | 4.7 | 17.8 | 2.9 |
| **263LeuN** | 1031 | 62 | 44.5 | 0.9 | 23.2 | 1.5 |
| **265TyrN** | 816 | 127 | 38.6 | 3.1 | 21.2 | 3.7 |
| **266LeuN** | 923 | 87 | 44.6 | 2.5 | 20.7 | 2.3 |
| **267LysN** | 897 | 58 | 41.8 | 3.5 | 21.4 | 2.2 |
| **268LysN** | 1146 | 51 | 49.2 | 0.2 | 23.3 | 1.0 |
| **269AlaN** | 982 | 78 | 51.6 | 2.2 | 19.0 | 1.7 |
| **271AsnN** | 1054 | 66 | 44.7 | 0.4 | 23.6 | 1.5 |
| **273AlaN** | 803 | 53 | 48.8 | 2.9 | 16.4 | 1.5 |
| **274TyrN** | 1206 | 162 | 38.7 | 5.5 | 31.1 | 6.1 |
| **276LysN** | 986 | 101 | 44.1 | 1.9 | 22.4 | 2.5 |
| **277ValN** | 659 | 46 | 39.0 | 1.1 | 16.9 | 1.3 |
| **279GlyN** | 452 | 44 | 52.4 | 4.1 | 8.6 | 1.1 |
| **282MetN** | 1074 | 233 | 54.1 | 6.2 | 19.9 | 4.9 |
| **283ArgN** | 796 | 65 | 52.2 | 7.3 | 15.2 | 2.5 |
| **284IleN** | 799 | 88 | 47.2 | 1.3 | 16.9 | 1.9 |
| **285ThrN** | 926 | 41 | 46.5 | 2.2 | 19.9 | 1.3 |
| **286ValN** | 980 | 90 | 46.5 | 4.6 | 21.1 | 2.8 |
| **287IleN** | 1117 | 181 | 44.0 | 2.9 | 25.4 | 4.4 |
| **288AlaN** | 1049 | 207 | 44.7 | 4.4 | 23.5 | 5.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **289IleN** | 761 | 94 | 47.4 | 3.7 | 16.0 | 2.3 |
| **290MetN** | 1313 | 111 | 51.2 | 1.4 | 25.6 | 2.3 |
| **291GluN** | 911 | 66 | 43.7 | 1.7 | 20.8 | 1.7 |
| **292AspN** | 892 | 104 | 41.1 | 2.2 | 21.7 | 2.8 |
| **293LeuN** | 803 | 121 | 41.1 | 2.4 | 19.5 | 3.1 |
| **294AspN** | 1103 | 117 | 44.4 | 0.3 | 24.8 | 2.6 |
| **295IleN** | 1025 | 51 | 45.9 | 3.4 | 22.3 | 2.0 |
| **296LeuN** | 836 | 83 | 47.9 | 2.8 | 17.5 | 2.0 |
| **297SerN** | 622 | 54 | 41.5 | 1.6 | 15.0 | 1.4 |
| **298LysN** | 1000 | 68 | 53.5 | 1.2 | 18.7 | 1.3 |
| **300LysN** | 936 | 66 | 52.3 | 5.2 | 17.9 | 2.2 |
| **301LysN** | 503 | 61 | 54.0 | 8.4 | 9.3 | 1.8 |
| **302PheN** | 809 | 85 | 48.4 | 2.6 | 16.7 | 2.0 |
| **305GlyN** | 410 | 53 | 55.5 | 15.8 | 7.4 | 2.3 |
| **306IleN** | 983 | 42 | 42.9 | 1.6 | 22.9 | 1.3 |
| **307SerN** | 530 | 191 | 43.4 | 23.5 | 12.2 | 7.9 |
| **308IleN** | 856 | 58 | 46.0 | 2.0 | 18.6 | 1.5 |
| **309AspN** | 894 | 29 | 43.4 | 1.4 | 20.6 | 1.0 |
| **310AsnN** | 743 | 58 | 49.4 | 3.2 | 15.0 | 1.5 |
| **311AlaN** | 862 | 82 | 41.1 | 3.2 | 21.0 | 2.6 |
| **312TyrN** | 730 | 39 | 37.7 | 2.0 | 19.3 | 1.4 |
| **313LysN** | 882 | 78 | 43.8 | 1.3 | 20.1 | 1.9 |
| **314ValN** | 972 | 102 | 43.6 | 0.4 | 22.3 | 2.3 |
| **315AlaN** | 930 | 86 | 39.4 | 1.3 | 23.6 | 2.3 |
| **316GluN** | 1053 | 56 | 41.7 | 2.9 | 25.3 | 2.2 |
| **317AspN** | 1040 | 103 | 46.3 | 2.0 | 22.5 | 2.4 |
| **318LeuN** | 1173 | 101 | 42.6 | 0.6 | 27.6 | 2.4 |
| **320ArgN** | 911 | 66 | 43.7 | 1.7 | 20.8 | 1.7 |
| **321GluN** | 899 | 52 | 42.3 | 2.3 | 21.3 | 1.7 |
| **322LeuN** | 1038 | 98 | 46.2 | 1.4 | 22.4 | 2.2 |
| **323LeuN** | 1233 | 165 | 43.4 | 0.2 | 28.4 | 3.8 |
| **324ValN** | 1072 | 105 | 45.1 | 1.2 | 23.7 | 2.4 |
| **325ArgN** | 731 | 40 | 49.2 | 2.1 | 14.8 | 1.0 |
| **326AspN** | 604 | 40 | 49.5 | 2.7 | 12.2 | 1.1 |
| **328ArgN** | 685 | 50 | 37.1 | 2.1 | 18.5 | 1.7 |
| **329ArgN** | 474 | 49 | 48.3 | 2.5 | 9.8 | 1.1 |
| **331ValN** | 793 | 88 | 48.4 | 2.2 | 16.4 | 2.0 |
| **332ArgN** | 635 | 124 | 40.6 | 5.2 | 15.6 | 3.6 |
| **333ArgN** | 741 | 72 | 35.7 | 1.4 | 20.7 | 2.2 |
| **334IleN** | 1062 | 169 | 43.4 | 2.3 | 24.5 | 4.1 |
| **335GlyN** | 850 | 83 | 48.3 | 3.2 | 17.6 | 2.1 |
| **336ValN** | 824 | 89 | 44.2 | 2.5 | 18.6 | 2.3 |
| **337LysN** | 777 | 57 | 41.2 | 0.6 | 18.8 | 1.4 |
| **339AspN** | 697 | 61 | 45.9 | 3.2 | 15.2 | 1.7 |

# Appendix C: $S^2_{iRED}$ order parameters from MD simulations of Dbh and $Dbh_{RKS(243-245)}$

| Residue Number | WT Dbh 35°C $S^2_{iRED}$ | $Dbh_{RKS(243-245)}$ 35°C $S^2_{iRED}$ | WT Dbh 50°C $S^2_{iRED}$ | $Dbh_{RKS(243-245)}$ 50°C $S^2_{iRED}$ |
|---|---|---|---|---|
| 2 | 0.66 | 0.69 | 0.67 | 0.63 |
| 3 | 0.86 | 0.89 | 0.83 | 0.82 |
| 4 | 0.90 | 0.89 | 0.88 | 0.85 |
| 5 | 0.88 | 0.90 | 0.87 | 0.85 |
| 6 | 0.89 | 0.92 | 0.89 | 0.85 |
| 7 | 0.90 | 0.92 | 0.89 | 0.86 |
| 8 | 0.62 | 0.86 | 0.65 | 0.50 |
| 9 | 0.74 | 0.56 | 0.74 | 0.73 |
| 10 | 0.83 | 0.83 | 0.82 | 0.75 |
| 11 | 0.90 | 0.92 | 0.89 | 0.83 |
| 12 | 0.89 | 0.89 | 0.89 | 0.80 |
| 13 | 0.88 | 0.86 | 0.87 | 0.78 |
| 14 | 0.89 | 0.90 | 0.87 | 0.77 |
| 15 | 0.87 | 0.91 | 0.86 | 0.77 |
| 16 | 0.91 | 0.91 | 0.90 | 0.84 |
| 17 | 0.90 | 0.90 | 0.87 | 0.81 |
| 18 | 0.82 | 0.87 | 0.79 | 0.76 |
| 19 | 0.87 | 0.86 | 0.85 | 0.74 |
| 20 | 0.83 | 0.84 | 0.81 | 0.72 |
| 22 | 0.78 | 0.84 | 0.77 | 0.75 |
| 23 | 0.80 | 0.86 | 0.81 | 0.77 |
| 24 | 0.86 | 0.89 | 0.86 | 0.78 |
| 26 | 0.69 | 0.69 | 0.68 | 0.64 |
| 28 | 0.81 | 0.83 | 0.81 | 0.74 |
| 29 | 0.90 | 0.91 | 0.88 | 0.82 |
| 30 | 0.80 | 0.87 | 0.83 | 0.74 |
| 31 | 0.88 | 0.89 | 0.85 | 0.80 |
| 32 | 0.82 | 0.83 | 0.85 | 0.66 |
| 33 | 0.61 | 0.56 | 0.79 | 0.45 |
| 34 | 0.37 | 0.38 | 0.70 | 0.42 |
| 36 | 0.13 | 0.08 | 0.49 | 0.18 |
| 37 | 0.16 | 0.14 | 0.46 | 0.27 |
| 38 | 0.44 | 0.14 | 0.47 | 0.39 |
| 39 | 0.39 | 0.11 | 0.56 | 0.48 |
| 40 | 0.41 | 0.35 | 0.49 | 0.27 |
| 42 | 0.82 | 0.57 | 0.74 | 0.68 |
| 43 | 0.87 | 0.87 | 0.88 | 0.77 |
| 44 | 0.85 | 0.87 | 0.85 | 0.78 |
| 45 | 0.86 | 0.89 | 0.85 | 0.79 |
| 46 | 0.85 | 0.88 | 0.84 | 0.78 |
| 47 | 0.89 | 0.92 | 0.87 | 0.80 |
| 48 | 0.90 | 0.92 | 0.89 | 0.85 |
| 49 | 0.88 | 0.89 | 0.86 | 0.80 |
| 50 | 0.86 | 0.89 | 0.85 | 0.80 |
| 51 | 0.91 | 0.92 | 0.90 | 0.86 |
| 52 | 0.87 | 0.90 | 0.87 | 0.80 |
| 53 | 0.83 | 0.84 | 0.81 | 0.77 |
| 55 | 0.83 | 0.85 | 0.82 | 0.75 |
| 56 | 0.81 | 0.79 | 0.74 | 0.66 |
| 57 | 0.86 | 0.88 | 0.84 | 0.80 |
| 59 | 0.80 | 0.83 | 0.81 | 0.74 |
| 61 | 0.90 | 0.90 | 0.88 | 0.82 |
| 62 | 0.88 | 0.89 | 0.87 | 0.80 |
| 63 | 0.82 | 0.84 | 0.82 | 0.76 |
| 64 | 0.88 | 0.90 | 0.86 | 0.81 |
| 65 | 0.89 | 0.89 | 0.88 | 0.80 |
| 66 | 0.79 | 0.81 | 0.78 | 0.72 |
| 67 | 0.78 | 0.82 | 0.78 | 0.74 |
| 68 | 0.86 | 0.86 | 0.84 | 0.78 |

| | | | | |
|---|---|---|---|---|
| 70 | 0.82 | 0.81 | 0.80 | 0.72 |
| 71 | 0.83 | 0.84 | 0.82 | 0.76 |
| 72 | 0.87 | 0.87 | 0.87 | 0.83 |
| 73 | 0.82 | 0.83 | 0.80 | 0.76 |
| 74 | 0.55 | 0.63 | 0.75 | 0.49 |
| 76 | 0.85 | 0.87 | 0.84 | 0.72 |
| 77 | 0.56 | 0.57 | 0.65 | 0.52 |
| 78 | 0.86 | 0.89 | 0.87 | 0.49 |
| 80 | 0.86 | 0.89 | 0.87 | 0.61 |
| 81 | 0.83 | 0.91 | 0.87 | 0.65 |
| 82 | 0.91 | 0.92 | 0.90 | 0.74 |
| 83 | 0.86 | 0.91 | 0.85 | 0.66 |
| 84 | 0.90 | 0.91 | 0.89 | 0.79 |
| 85 | 0.90 | 0.92 | 0.89 | 0.77 |
| 86 | 0.91 | 0.92 | 0.89 | 0.71 |
| 87 | 0.91 | 0.91 | 0.88 | 0.81 |
| 88 | 0.90 | 0.90 | 0.88 | 0.80 |
| 89 | 0.93 | 0.93 | 0.91 | 0.86 |
| 90 | 0.90 | 0.91 | 0.87 | 0.85 |
| 91 | 0.85 | 0.87 | 0.86 | 0.79 |
| 92 | 0.87 | 0.89 | 0.87 | 0.81 |
| 93 | 0.91 | 0.91 | 0.91 | 0.85 |
| 94 | 0.82 | 0.85 | 0.82 | 0.80 |
| 95 | 0.76 | 0.78 | 0.76 | 0.68 |
| 96 | 0.77 | 0.76 | 0.75 | 0.68 |
| 97 | 0.81 | 0.70 | 0.81 | 0.74 |
| 98 | 0.86 | 0.79 | 0.84 | 0.70 |
| 99 | 0.90 | 0.88 | 0.78 | 0.49 |
| 100 | 0.86 | 0.86 | 0.78 | 0.66 |
| 101 | 0.74 | 0.87 | 0.73 | 0.65 |
| 102 | 0.61 | 0.75 | 0.46 | 0.33 |
| 103 | 0.74 | 0.86 | 0.62 | 0.54 |
| 104 | 0.75 | 0.87 | 0.69 | 0.62 |
| 105 | 0.71 | 0.89 | 0.67 | 0.49 |
| 106 | 0.85 | 0.92 | 0.81 | 0.69 |
| 107 | 0.77 | 0.94 | 0.75 | 0.57 |
| 108 | 0.74 | 0.90 | 0.78 | 0.76 |
| 109 | 0.87 | 0.91 | 0.86 | 0.83 |
| 110 | 0.82 | 0.87 | 0.76 | 0.78 |
| 111 | 0.89 | 0.88 | 0.88 | 0.82 |
| 112 | 0.90 | 0.91 | 0.91 | 0.83 |
| 113 | 0.89 | 0.89 | 0.89 | 0.84 |
| 114 | 0.76 | 0.79 | 0.80 | 0.75 |
| 115 | 0.77 | 0.79 | 0.78 | 0.73 |
| 116 | 0.89 | 0.91 | 0.89 | 0.85 |
| 118 | 0.80 | 0.81 | 0.82 | 0.74 |
| 119 | 0.89 | 0.85 | 0.87 | 0.81 |
| 120 | 0.88 | 0.85 | 0.89 | 0.83 |
| 121 | 0.85 | 0.85 | 0.85 | 0.79 |
| 123 | 0.92 | 0.92 | 0.91 | 0.88 |
| 124 | 0.92 | 0.93 | 0.91 | 0.90 |
| 125 | 0.91 | 0.92 | 0.89 | 0.85 |
| 126 | 0.94 | 0.93 | 0.91 | 0.90 |
| 127 | 0.93 | 0.93 | 0.92 | 0.90 |
| 128 | 0.92 | 0.93 | 0.91 | 0.87 |
| 129 | 0.90 | 0.91 | 0.88 | 0.85 |
| 130 | 0.93 | 0.93 | 0.92 | 0.89 |
| 131 | 0.90 | 0.91 | 0.89 | 0.88 |
| 132 | 0.90 | 0.91 | 0.88 | 0.85 |
| 133 | 0.92 | 0.91 | 0.89 | 0.88 |
| 134 | 0.91 | 0.91 | 0.91 | 0.87 |
| 135 | 0.85 | 0.85 | 0.84 | 0.81 |
| 136 | 0.83 | 0.83 | 0.80 | 0.76 |
| 137 | 0.81 | 0.78 | 0.78 | 0.78 |
| 138 | 0.89 | 0.89 | 0.88 | 0.84 |
| 139 | 0.72 | 0.62 | 0.70 | 0.66 |
| 140 | 0.84 | 0.86 | 0.82 | 0.77 |
| 141 | 0.59 | 0.89 | 0.56 | 0.35 |
| 142 | 0.84 | 0.88 | 0.85 | 0.78 |
| 143 | 0.90 | 0.92 | 0.90 | 0.88 |

| | | | | |
|---|---|---|---|---|
| 145 | 0.88 | 0.88 | 0.87 | 0.85 |
| 146 | 0.92 | 0.93 | 0.91 | 0.90 |
| 148 | 0.85 | 0.86 | 0.87 | 0.81 |
| 149 | 0.90 | 0.92 | 0.91 | 0.89 |
| 150 | 0.88 | 0.89 | 0.86 | 0.82 |
| 151 | 0.91 | 0.93 | 0.91 | 0.89 |
| 152 | 0.93 | 0.94 | 0.91 | 0.88 |
| 153 | 0.91 | 0.92 | 0.91 | 0.86 |
| 154 | 0.91 | 0.92 | 0.88 | 0.85 |
| 155 | 0.92 | 0.94 | 0.92 | 0.87 |
| 156 | 0.92 | 0.93 | 0.90 | 0.90 |
| 157 | 0.90 | 0.92 | 0.91 | 0.85 |
| 158 | 0.84 | 0.87 | 0.85 | 0.80 |
| 159 | 0.68 | 0.69 | 0.68 | 0.60 |
| 160 | 0.78 | 0.82 | 0.80 | 0.78 |
| 162 | 0.85 | 0.85 | 0.85 | 0.84 |
| 164 | 0.83 | 0.85 | 0.84 | 0.79 |
| 166 | 0.86 | 0.88 | 0.86 | 0.81 |
| 167 | 0.86 | 0.89 | 0.85 | 0.84 |
| 168 | 0.84 | 0.86 | 0.84 | 0.78 |
| 170 | 0.77 | 0.80 | 0.77 | 0.73 |
| 171 | 0.73 | 0.75 | 0.69 | 0.67 |
| 172 | 0.84 | 0.86 | 0.82 | 0.82 |
| 173 | 0.88 | 0.89 | 0.82 | 0.83 |
| 174 | 0.80 | 0.83 | 0.76 | 0.78 |
| 175 | 0.83 | 0.85 | 0.76 | 0.79 |
| 176 | 0.86 | 0.86 | 0.79 | 0.74 |
| 177 | 0.82 | 0.86 | 0.79 | 0.79 |
| 178 | 0.73 | 0.78 | 0.71 | 0.66 |
| 179 | 0.84 | 0.86 | 0.75 | 0.78 |
| 180 | 0.83 | 0.87 | 0.77 | 0.79 |
| 181 | 0.86 | 0.83 | 0.84 | 0.79 |
| 182 | 0.83 | 0.78 | 0.72 | 0.77 |
| 183 | 0.79 | 0.76 | 0.73 | 0.76 |
| 184 | 0.82 | 0.83 | 0.74 | 0.80 |
| 187 | 0.65 | 0.75 | 0.69 | 0.63 |
| 189 | 0.78 | 0.81 | 0.72 | 0.49 |
| 190 | 0.79 | 0.81 | 0.74 | 0.43 |
| 191 | 0.78 | 0.80 | 0.78 | 0.52 |
| 192 | 0.85 | 0.87 | 0.81 | 0.77 |
| 193 | 0.85 | 0.87 | 0.79 | 0.78 |
| 194 | 0.83 | 0.88 | 0.80 | 0.81 |
| 195 | 0.85 | 0.87 | 0.83 | 0.78 |
| 196 | 0.84 | 0.88 | 0.78 | 0.78 |
| 197 | 0.85 | 0.88 | 0.81 | 0.82 |
| 198 | 0.78 | 0.82 | 0.78 | 0.75 |
| 200 | 0.69 | 0.73 | 0.69 | 0.66 |
| 201 | 0.73 | 0.76 | 0.70 | 0.73 |
| 202 | 0.72 | 0.77 | 0.71 | 0.68 |
| 203 | 0.86 | 0.89 | 0.84 | 0.85 |
| 204 | 0.86 | 0.89 | 0.85 | 0.84 |
| 205 | 0.84 | 0.89 | 0.82 | 0.83 |
| 206 | 0.87 | 0.89 | 0.77 | 0.84 |
| 207 | 0.82 | 0.86 | 0.77 | 0.80 |
| 208 | 0.66 | 0.70 | 0.70 | 0.64 |
| 209 | 0.64 | 0.66 | 0.66 | 0.61 |
| 210 | 0.78 | 0.81 | 0.75 | 0.72 |
| 211 | 0.82 | 0.86 | 0.81 | 0.78 |
| 212 | 0.85 | 0.86 | 0.79 | 0.83 |
| 213 | 0.80 | 0.83 | 0.76 | 0.77 |
| 214 | 0.80 | 0.84 | 0.77 | 0.77 |
| 215 | 0.84 | 0.89 | 0.80 | 0.77 |
| 216 | 0.84 | 0.88 | 0.79 | 0.81 |
| 217 | 0.72 | 0.76 | 0.65 | 0.67 |
| 218 | 0.75 | 0.80 | 0.71 | 0.73 |
| 220 | 0.82 | 0.87 | 0.80 | 0.79 |
| 221 | 0.83 | 0.86 | 0.79 | 0.81 |
| 222 | 0.83 | 0.90 | 0.78 | 0.80 |
| 223 | 0.88 | 0.91 | 0.86 | 0.85 |
| 224 | 0.88 | 0.93 | 0.84 | 0.88 |

| | | | | |
|---|---|---|---|---|
| 225 | 0.87 | 0.92 | 0.80 | 0.86 |
| 226 | 0.88 | 0.92 | 0.85 | 0.86 |
| 227 | 0.87 | 0.92 | 0.84 | 0.85 |
| 228 | 0.88 | 0.92 | 0.83 | 0.87 |
| 229 | 0.86 | 0.92 | 0.81 | 0.87 |
| 230 | 0.88 | 0.92 | 0.83 | 0.86 |
| 231 | 0.84 | 0.89 | 0.73 | 0.85 |
| 232 | 0.85 | 0.91 | 0.67 | 0.83 |
| 233 | 0.61 | 0.71 | 0.61 | 0.64 |
| 234 | 0.55 | 0.61 | 0.40 | 0.54 |
| 235 | 0.47 | 0.77 | 0.38 | 0.73 |
| 236 | 0.44 | 0.58 | 0.55 | 0.60 |
| 238 | 0.76 | 0.79 | 0.80 | 0.72 |
| 239 | 0.71 | 0.76 | 0.70 | 0.48 |
| 240 | 0.67 | 0.83 | 0.63 | 0.57 |
| 241 | 0.45 | 0.46 | 0.39 | 0.47 |
| 242 | 0.53 | 0.73 | 0.41 | 0.33 |
| 243 | 0.50 | 0.77 | 0.47 | 0.44 |
| 244 | 0.44 | 0.71 | 0.46 | 0.50 |
| 245 | | 0.53 | 0.23 | 0.64 |
| 246 | 0.62 | 0.81 | 0.55 | 0.67 |
| 248 | 0.83 | 0.78 | 0.69 | 0.64 |
| 249 | 0.81 | 0.76 | 0.66 | 0.66 |
| 250 | 0.79 | 0.75 | 0.72 | 0.63 |
| 251 | 0.79 | 0.83 | 0.68 | 0.71 |
| 252 | 0.81 | 0.84 | 0.67 | 0.68 |
| 254 | 0.74 | 0.79 | 0.60 | 0.66 |
| 255 | 0.72 | 0.78 | 0.60 | 0.53 |
| 256 | 0.78 | 0.81 | 0.65 | 0.53 |
| 257 | 0.80 | 0.82 | 0.62 | 0.62 |
| 258 | 0.81 | 0.80 | 0.71 | 0.66 |
| 259 | 0.80 | 0.82 | 0.69 | 0.71 |
| 260 | 0.81 | 0.79 | 0.66 | 0.63 |
| 261 | 0.79 | 0.80 | 0.67 | 0.62 |
| 262 | 0.78 | 0.78 | 0.71 | 0.69 |
| 263 | 0.84 | 0.86 | 0.70 | 0.62 |
| 265 | 0.82 | 0.89 | 0.74 | 0.62 |
| 266 | 0.84 | 0.87 | 0.77 | 0.70 |
| 267 | 0.89 | 0.89 | 0.76 | 0.69 |
| 268 | 0.83 | 0.88 | 0.73 | 0.55 |
| 269 | 0.84 | 0.87 | 0.72 | 0.70 |
| 270 | 0.85 | 0.88 | 0.78 | 0.61 |
| 271 | 0.88 | 0.88 | 0.70 | 0.64 |
| 272 | 0.81 | 0.84 | 0.69 | 0.53 |
| 273 | 0.81 | 0.81 | 0.73 | 0.58 |
| 274 | 0.87 | 0.82 | 0.74 | 0.60 |
| 275 | 0.82 | 0.81 | 0.60 | 0.52 |
| 276 | 0.75 | 0.67 | 0.45 | 0.62 |
| 277 | 0.74 | 0.60 | 0.59 | 0.47 |
| 278 | 0.81 | 0.68 | 0.57 | 0.51 |
| 280 | 0.57 | 0.62 | 0.74 | 0.39 |
| 282 | 0.84 | 0.84 | 0.76 | 0.64 |
| 283 | 0.79 | 0.86 | 0.73 | 0.68 |
| 284 | 0.83 | 0.87 | 0.78 | 0.68 |
| 285 | 0.83 | 0.88 | 0.70 | 0.74 |
| 286 | 0.81 | 0.84 | 0.76 | 0.66 |
| 287 | 0.85 | 0.85 | 0.72 | 0.77 |
| 288 | 0.81 | 0.86 | 0.72 | 0.67 |
| 289 | 0.82 | 0.79 | 0.75 | 0.66 |
| 290 | 0.85 | 0.88 | 0.72 | 0.71 |
| 291 | 0.84 | 0.85 | 0.71 | 0.65 |
| 292 | 0.77 | 0.75 | 0.65 | 0.65 |
| 293 | 0.79 | 0.62 | 0.55 | 0.60 |
| 294 | 0.67 | 0.66 | 0.64 | 0.46 |
| 295 | 0.73 | 0.71 | 0.73 | 0.53 |
| 296 | 0.78 | 0.80 | 0.71 | 0.63 |
| 297 | 0.83 | 0.83 | 0.74 | 0.73 |
| 298 | 0.84 | 0.88 | 0.70 | 0.68 |
| 300 | 0.81 | 0.86 | 0.72 | 0.65 |
| 301 | 0.82 | 0.84 | 0.67 | 0.70 |

| | | | | |
|---|---|---|---|---|
| 302 | 0.76 | 0.81 | 0.68 | 0.61 |
| 303 | 0.75 | 0.84 | 0.63 | 0.53 |
| 304 | 0.71 | 0.68 | 0.71 | 0.49 |
| 306 | 0.69 | 0.84 | 0.69 | 0.55 |
| 307 | 0.79 | 0.83 | 0.74 | 0.54 |
| 308 | 0.82 | 0.85 | 0.71 | 0.71 |
| 309 | 0.76 | 0.81 | 0.63 | 0.60 |
| 310 | 0.63 | 0.63 | 0.73 | 0.57 |
| 311 | 0.83 | 0.85 | 0.73 | 0.66 |
| 312 | 0.83 | 0.84 | 0.70 | 0.69 |
| 313 | 0.75 | 0.81 | 0.69 | 0.52 |
| 314 | 0.78 | 0.83 | 0.75 | 0.55 |
| 315 | 0.82 | 0.86 | 0.76 | 0.72 |
| 316 | 0.83 | 0.89 | 0.73 | 0.60 |
| 317 | 0.79 | 0.86 | 0.70 | 0.51 |
| 318 | 0.81 | 0.83 | 0.74 | 0.65 |
| 319 | 0.86 | 0.87 | 0.77 | 0.69 |
| 320 | 0.83 | 0.88 | 0.68 | 0.62 |
| 321 | 0.77 | 0.77 | 0.67 | 0.54 |
| 322 | 0.75 | 0.78 | 0.66 | 0.65 |
| 323 | 0.80 | 0.76 | 0.60 | 0.56 |
| 324 | 0.69 | 0.72 | 0.36 | 0.48 |
| 325 | 0.67 | 0.65 | 0.33 | 0.49 |
| 326 | 0.67 | 0.61 | 0.32 | 0.49 |
| 327 | 0.60 | 0.51 | 0.44 | 0.44 |
| 328 | 0.68 | 0.72 | 0.33 | 0.53 |
| 329 | 0.61 | 0.68 | 0.64 | 0.47 |
| 330 | 0.80 | 0.81 | 0.49 | 0.61 |
| 331 | 0.81 | 0.85 | 0.70 | 0.60 |
| 332 | 0.83 | 0.84 | 0.72 | 0.67 |
| 333 | 0.85 | 0.83 | 0.72 | 0.66 |
| 334 | 0.82 | 0.84 | 0.73 | 0.68 |
| 336 | 0.84 | 0.86 | 0.75 | 0.67 |
| 337 | 0.83 | 0.85 | 0.70 | 0.75 |
| 338 | 0.82 | 0.85 | 0.76 | 0.66 |
| 339 | 0.82 | 0.88 | 0.77 | 0.68 |
| 340 | 0.90 | 0.90 | 0.72 | 0.65 |
| 341 | 0.81 | 0.81 | 0.63 | 0.65 |
| 342 | 0.84 | 0.81 | 0.51 | 0.55 |
| 343 | 0.83 | 0.79 | 0.50 | 0.50 |
| 344 | 0.78 | 0.74 | 0.30 | 0.58 |
| 345 | 0.51 | 0.38 | 0.36 | 0.48 |
| 346 | 0.78 | 0.68 | 0.35 | 0.29 |
| 347 | 0.64 | 0.70 | 0.29 | 0.31 |
| 348 | 0.57 | 0.72 | 0.26 | 0.30 |
| 349 | 0.58 | 0.48 | 0.22 | 0.22 |
| 350 | 0.32 | 0.66 | 0.27 | 0.33 |
| 351 | 0.32 | 0.68 | 0.32 | 0.28 |
| 352 | 0.32 | 0.72 | 0.32 | 0.25 |
| 353 | 0.40 | 0.73 | 0.26 | 0.24 |
| 354 | 0.21 | 0.66 | 0.18 | 0.11 |
| 312 | 0.83 | 0.84 | 0.70 | 0.69 |
| 313 | 0.75 | 0.81 | 0.69 | 0.52 |
| 314 | 0.78 | 0.83 | 0.75 | 0.55 |
| 315 | 0.82 | 0.86 | 0.76 | 0.72 |
| 316 | 0.83 | 0.89 | 0.73 | 0.60 |
| 317 | 0.79 | 0.86 | 0.70 | 0.51 |
| 318 | 0.81 | 0.83 | 0.74 | 0.65 |
| 319 | 0.86 | 0.87 | 0.77 | 0.69 |
| 320 | 0.83 | 0.88 | 0.68 | 0.62 |
| 321 | 0.77 | 0.77 | 0.67 | 0.54 |
| 322 | 0.75 | 0.78 | 0.66 | 0.65 |
| 323 | 0.80 | 0.76 | 0.60 | 0.56 |
| 324 | 0.69 | 0.72 | 0.36 | 0.48 |
| 325 | 0.67 | 0.65 | 0.33 | 0.49 |
| 326 | 0.67 | 0.61 | 0.32 | 0.49 |
| 327 | 0.60 | 0.51 | 0.44 | 0.44 |
| 328 | 0.68 | 0.72 | 0.33 | 0.53 |
| 329 | 0.61 | 0.68 | 0.64 | 0.47 |
| 330 | 0.80 | 0.81 | 0.49 | 0.61 |

| | | | | |
|---|---|---|---|---|
| 331 | 0.81 | 0.85 | 0.70 | 0.60 |
| 332 | 0.83 | 0.84 | 0.72 | 0.67 |
| 333 | 0.85 | 0.83 | 0.72 | 0.66 |
| 334 | 0.82 | 0.84 | 0.73 | 0.68 |
| 336 | 0.84 | 0.86 | 0.75 | 0.67 |
| 337 | 0.83 | 0.85 | 0.70 | 0.75 |
| 338 | 0.82 | 0.85 | 0.76 | 0.66 |
| 339 | 0.82 | 0.88 | 0.77 | 0.68 |
| 340 | 0.90 | 0.90 | 0.72 | 0.65 |
| 341 | 0.81 | 0.81 | 0.63 | 0.65 |
| 342 | 0.84 | 0.81 | 0.51 | 0.55 |
| 343 | 0.83 | 0.79 | 0.50 | 0.50 |
| 344 | 0.78 | 0.74 | 0.30 | 0.58 |
| 345 | 0.51 | 0.38 | 0.36 | 0.48 |
| 346 | 0.78 | 0.68 | 0.35 | 0.29 |
| 347 | 0.64 | 0.70 | 0.29 | 0.31 |
| 348 | 0.57 | 0.72 | 0.26 | 0.30 |
| 349 | 0.58 | 0.48 | 0.22 | 0.22 |
| 350 | 0.32 | 0.66 | 0.27 | 0.33 |
| 351 | 0.32 | 0.68 | 0.32 | 0.28 |
| 352 | 0.32 | 0.72 | 0.32 | 0.25 |
| 353 | 0.40 | 0.73 | 0.26 | 0.24 |
| 354 | 0.21 | 0.66 | 0.18 | 0.11 |

# Appendix D: Hydrogen-bond analysis from MD simulations of Dbh

*Note: This includes only the donor-acceptors pairs that meet the hydrogen bond criterion (within 3.5Å and 135° angle) for at least 10% of the frames in each simulation

**35°C, 500ns simulation of Dbh**

| # Lifetimes | Max Lifetime | Avg. Lifetime | # Frames | % Occupancy | Donor-Acceptor Pair |
|---|---|---|---|---|---|
| 1188 | 505 | 40.521 | 48139 | 97.29% | ALA_146@O-VAL_3@N-H |
| 407 | 1215 | 120.4202 | 49011 | 99.05% | LEU_109@O-ILE_4@N-H |
| 1004 | 442 | 48.2361 | 48429 | 97.88% | GLY_144@O-PHE_5@N-H |
| 2812 | 326 | 12.734 | 35808 | 72.37% | ALA_107@O-VAL_6@N-H |
| 1162 | 915 | 41.4957 | 48218 | 97.45% | THR_142@O-ASP_7@N-H |
| 3125 | 23 | 1.6458 | 5143 | 10.39% | VAL_6@O-PHE_8@N-H |
| 3484 | 157 | 5.244 | 18270 | 36.92% | ASP_105@O-PHE_8@N-H |
| 3231 | 62 | 3.4751 | 11228 | 22.69% | ASP_7@O-ASP_9@N-H |
| 8062 | 54 | 3.1569 | 25451 | 51.44% | PHE_8@O-PHE_11@N-H |
| 10934 | 55 | 2.3303 | 25480 | 51.50% | TYR_10@O-ALA_13@N-H |
| 10052 | 38 | 3.1339 | 31502 | 63.67% | TYR_10@O-GLN_14@N-H |
| 7166 | 9 | 1.3302 | 9532 | 19.26% | PHE_11@O-GLN_14@N-H |
| 4383 | 348 | 10.0319 | 43970 | 88.87% | PHE_11@O-VAL_15@N-H |
| 5425 | 90 | 7.5228 | 40811 | 82.48% | PHE_12@O-GLU_16@N-H |
| 9584 | 42 | 3.0654 | 29379 | 59.38% | ALA_13@O-GLU_17@N-H |
| 4096 | 12 | 1.2793 | 5240 | 10.59% | GLN_14@O-GLU_17@N-H |
| 8862 | 51 | 2.8187 | 24979 | 50.48% | GLN_14@O-VAL_18@N-H |
| 7403 | 42 | 1.864 | 13799 | 27.89% | VAL_15@O-VAL_18@N-H |
| 5437 | 142 | 7.8389 | 42620 | 86.14% | VAL_15@O-LEU_19@N-H |
| 10780 | 22 | 1.8455 | 19894 | 40.21% | GLU_16@O-ASN_20@N-H |
| 6200 | 18 | 1.4002 | 8681 | 17.54% | GLU_17@O-ASN_20@N-H |
| 3453 | 156 | 12.1445 | 41935 | 84.75% | ASN_20@O-TYR_23@N-H |
| 7867 | 40 | 4.157 | 32703 | 66.09% | PRO_21@O-LYS_24@N-H |
| 7914 | 40 | 3.8884 | 30773 | 62.19% | TYR_23@O-LYS_26@N-H |
| 1339 | 438 | 35.9022 | 48073 | 97.16% | ILE_72@O-VAL_29@N-H |
| 3282 | 602 | 13.8041 | 45305 | 91.56% | THR_45@O-VAL_30@N-H |
| 2786 | 256 | 16.6378 | 46353 | 93.68% | VAL_74@O-SER_31@N-H |
| 1826 | 225 | 25.9014 | 47296 | 95.59% | ALA_42@O-VAL_32@N-H |
| 3348 | 116 | 8.0502 | 26952 | 54.47% | SER_40@O-SER_34@N-H |
| 2738 | 55 | 3.6308 | 9941 | 20.09% | GLY_35@O-THR_37@N-H |
| 7458 | 39 | 2.269 | 16922 | 34.20% | THR_37@O-SER_40@N-H |
| 4244 | 161 | 9.3827 | 39820 | 80.48% | VAL_32@O-ALA_42@N-H |
| 605 | 1068 | 80.7388 | 48847 | 98.72% | MET_59@O-VAL_43@N-H |
| 4740 | 107 | 9.2888 | 44029 | 88.99% | VAL_30@O-ALA_44@N-H |
| 5478 | 292 | 7.5329 | 41265 | 83.40% | VAL_30@O-THR_45@N-H |
| 850 | 604 | 57.16 | 48586 | 98.20% | LEU_28@O-ASN_47@N-H |
| 4461 | 9 | 1.1946 | 5329 | 10.77% | ASN_47@O-ALA_50@N-H |
| 1398 | 338 | 34.3276 | 47990 | 96.99% | ASN_47@O-ARG_51@N-H |
| 4665 | 135 | 9.2658 | 43225 | 87.36% | TYR_48@O-LYS_52@N-H |
| 9487 | 35 | 1.8615 | 17660 | 35.69% | GLU_49@O-LEU_53@N-H |
| 8643 | 74 | 1.6578 | 14328 | 28.96% | ALA_50@O-LEU_53@N-H |
| 7545 | 11 | 1.4869 | 11219 | 22.67% | ALA_50@O-GLY_54@N-H |
| 10581 | 42 | 2.7562 | 29163 | 58.94% | ARG_51@O-GLY_54@N-H |
| 6506 | 86 | 5.9175 | 38499 | 77.81% | ALA_50@O-VAL_55@N-H |
| 3973 | 121 | 11.2112 | 44542 | 90.02% | VAL_43@O-GLY_58@N-H |
| 9537 | 17 | 1.6655 | 15884 | 32.10% | LYS_56@O-MET_59@N-H |
| 7121 | 319 | 5.0232 | 35770 | 72.29% | GLY_41@O-ILE_61@N-H |
| 10433 | 14 | 1.7379 | 18131 | 36.64% | PRO_60@O-LYS_63@N-H |
| 2740 | 170 | 16.9255 | 46376 | 93.73% | PRO_60@O-ALA_64@N-H |
| 560 | 895 | 87.2982 | 48887 | 98.80% | ILE_61@O-MET_65@N-H |
| 8913 | 16 | 1.6247 | 14481 | 29.27% | ILE_62@O-GLN_66@N-H |
| 10717 | 48 | 1.942 | 20812 | 42.06% | LYS_63@O-GLN_66@N-H |
| 8330 | 49 | 4.1579 | 34635 | 70.00% | LYS_63@O-ILE_67@N-H |
| 6681 | 15 | 1.4308 | 9559 | 19.32% | ALA_64@O-ILE_67@N-H |
| 2255 | 295 | 20.8475 | 47011 | 95.01% | ALA_64@O-ALA_68@N-H |
| 11743 | 36 | 2.4281 | 28513 | 57.63% | ALA_68@O-ALA_71@N-H |
| 1184 | 532 | 40.3226 | 47742 | 96.49% | PRO_27@O-ILE_72@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 5012 | 148 | 7.4607 | 37393 | 75.57% | VAL_29@O-VAL_74@N-H |
| 11208 | 26 | 2.5545 | 28631 | 57.86% | SER_31@O-MET_76@N-H |
| 5591 | 20 | 1.8481 | 10333 | 20.88% | MET_76@O-LYS_78@N-H |
| 8715 | 92 | 4.0893 | 35638 | 72.03% | ARG_77@O-ILE_80@N-H |
| 7606 | 24 | 1.7782 | 13525 | 27.33% | ARG_77@O-TYR_81@N-H |
| 3349 | 59 | 1.5563 | 5212 | 10.53% | LYS_78@O-TYR_81@N-H |
| 2419 | 327 | 18.7565 | 45372 | 91.70% | LYS_78@O-GLU_82@N-H |
| 9681 | 33 | 2.9771 | 28821 | 58.25% | PRO_79@O-ALA_83@N-H |
| 3889 | 164 | 10.3474 | 40241 | 81.33% | ILE_80@O-PHE_84@N-H |
| 2601 | 264 | 16.6817 | 43389 | 87.69% | TYR_81@O-SER_85@N-H |
| 2595 | 325 | 16.9187 | 43904 | 88.73% | GLU_82@O-ASN_86@N-H |
| 6810 | 138 | 5.7206 | 38957 | 78.73% | ALA_83@O-ARG_87@N-H |
| 4430 | 181 | 9.7079 | 43006 | 86.92% | PHE_84@O-ILE_88@N-H |
| 1237 | 1051 | 38.7437 | 47926 | 96.86% | SER_85@O-MET_89@N-H |
| 4569 | 180 | 9.5835 | 43787 | 88.50% | ASN_86@O-ASN_90@N-H |
| 8720 | 39 | 3.1444 | 27419 | 55.42% | ARG_87@O-LEU_91@N-H |
| 5348 | 40 | 1.8407 | 9844 | 19.90% | ILE_88@O-LEU_91@N-H |
| 2713 | 244 | 16.9598 | 46012 | 92.99% | ILE_88@O-LEU_92@N-H |
| 2842 | 372 | 16.1038 | 45767 | 92.50% | MET_89@O-ASN_93@N-H |
| 8188 | 28 | 2.0561 | 16835 | 34.02% | ASN_90@O-LYS_94@N-H |
| 8724 | 38 | 1.9411 | 16934 | 34.22% | LEU_91@O-LYS_94@N-H |
| 6171 | 23 | 1.7751 | 10954 | 22.14% | LEU_91@O-HIE_95@N-H |
| 9563 | 38 | 2.47 | 23621 | 47.74% | LEU_92@O-HIE_95@N-H |
| 7918 | 69 | 4.4554 | 35278 | 71.30% | LEU_92@O-ALA_96@N-H |
| 2025 | 402 | 20.8247 | 42170 | 85.23% | ASP_110@O-ASP_97@N-H |
| 741 | 612 | 65.6559 | 48651 | 98.33% | TYR_108@O-GLU_100@N-H |
| 2618 | 322 | 11.2074 | 29341 | 59.30% | LYS_241@O-VAL_101@N-H |
| 5478 | 56 | 2.5694 | 14075 | 28.45% | GLU_106@O-ALA_102@N-H |
| 6377 | 45 | 3.6784 | 23457 | 47.41% | GLU_106@O-SER_103@N-H |
| 7987 | 177 | 4.1316 | 32999 | 66.69% | SER_103@O-GLU_106@N-H |
| 4907 | 280 | 4.1973 | 20596 | 41.63% | VAL_6@O-ALA_107@N-H |
| 525 | 837 | 65.8952 | 34595 | 69.92% | GLU_100@O-TYR_108@N-H |
| 2609 | 63 | 3.524 | 9194 | 18.58% | GLU_106@O-TYR_108@N-H |
| 1135 | 520 | 42.5207 | 48261 | 97.54% | ILE_4@O-LEU_109@N-H |
| 2229 | 344 | 20.5895 | 45894 | 92.75% | LYS_98@O-ASP_110@N-H |
| 7150 | 163 | 5.0649 | 36214 | 73.19% | ILE_2@O-VAL_111@N-H |
| 7931 | 38 | 1.7324 | 13740 | 27.77% | VAL_111@O-LYS_114@N-H |
| 8853 | 64 | 3.8677 | 34241 | 69.20% | VAL_111@O-VAL_115@N-H |
| 7301 | 13 | 1.418 | 10353 | 20.92% | THR_112@O-VAL_115@N-H |
| 9858 | 23 | 2.2697 | 22375 | 45.22% | ASN_113@O-GLU_116@N-H |
| 7452 | 78 | 5.2322 | 38990 | 78.80% | THR_112@O-GLY_117@N-H |
| 7635 | 60 | 5.1395 | 39240 | 79.31% | VAL_115@O-ASN_118@N-H |
| 6394 | 14 | 1.3924 | 8903 | 17.99% | ASN_118@O-ASN_121@N-H |
| 9714 | 38 | 3.2453 | 31525 | 63.71% | ASN_118@O-GLY_122@N-H |
| 5022 | 342 | 8.3598 | 41983 | 84.85% | PHE_119@O-ILE_123@N-H |
| 7647 | 76 | 4.0449 | 30931 | 62.51% | GLU_120@O-GLU_124@N-H |
| 7068 | 82 | 5.5805 | 39443 | 79.72% | ASN_121@O-LEU_125@N-H |
| 2608 | 206 | 17.8972 | 46676 | 94.33% | GLY_122@O-ALA_126@N-H |
| 2157 | 255 | 21.8215 | 47069 | 95.13% | ILE_123@O-ARG_127@N-H |
| 3416 | 221 | 13.3451 | 45587 | 92.13% | GLU_124@O-LYS_128@N-H |
| 697 | 1245 | 69.9555 | 48759 | 98.54% | LEU_125@O-ILE_129@N-H |
| 2931 | 313 | 15.5677 | 45629 | 92.22% | ALA_126@O-LYS_130@N-H |
| 3607 | 281 | 12.4303 | 44836 | 90.62% | ARG_127@O-GLN_131@N-H |
| 3774 | 507 | 11.9014 | 44916 | 90.78% | LYS_128@O-GLU_132@N-H |
| 4084 | 204 | 10.8244 | 44207 | 89.34% | ILE_129@O-ILE_133@N-H |
| 8719 | 89 | 3.2407 | 28256 | 57.11% | LYS_130@O-LEU_134@N-H |
| 6944 | 107 | 5.6326 | 39113 | 79.05% | GLN_131@O-GLU_135@N-H |
| 1520 | 413 | 31.3125 | 47595 | 96.19% | GLU_132@O-LYS_136@N-H |
| 9236 | 110 | 3.5225 | 32534 | 65.75% | ILE_133@O-GLU_137@N-H |
| 4423 | 21 | 1.3366 | 5912 | 11.95% | ILE_133@O-LYS_138@N-H |
| 5262 | 19 | 1.5785 | 8306 | 16.79% | ILE_133@O-ILE_139@N-H |
| 4712 | 11 | 1.3162 | 6202 | 12.53% | GLU_137@O-ILE_139@N-H |
| 849 | 489 | 11.371 | 9654 | 19.51% | ASP_7@O-THR_142@N-H |
| 4043 | 180 | 10.9644 | 44329 | 89.59% | GLY_163@O-VAL_143@N-H |
| 1726 | 474 | 27.6025 | 47642 | 96.29% | PHE_5@O-GLY_144@N-H |
| 7203 | 123 | 5.5531 | 39999 | 80.84% | GLY_165@O-VAL_145@N-H |
| 1171 | 610 | 41.1401 | 48175 | 97.36% | VAL_3@O-ALA_146@N-H |
| 6316 | 69 | 6.5249 | 41211 | 83.29% | ASN_148@O-ALA_152@N-H |
| 6661 | 111 | 5.0459 | 33611 | 67.93% | LYS_149@O-LYS_153@N-H |
| 3320 | 240 | 13.4961 | 44807 | 90.56% | ILE_150@O-ILE_154@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 915 | 522 | 52.9814 | 48478 | 97.98% | LEU_151@O-ILE_155@N-H |
| 10559 | 54 | 2.9208 | 30841 | 62.33% | ALA_152@O-ALA_156@N-H |
| 2191 | 316 | 20.885 | 45759 | 92.48% | LYS_153@O-ASP_157@N-H |
| 7060 | 153 | 5.2021 | 36727 | 74.23% | ILE_154@O-LYS_158@N-H |
| 6924 | 17 | 1.575 | 10905 | 22.04% | ILE_155@O-SER_159@N-H |
| 10114 | 40 | 2.7911 | 28229 | 57.05% | ALA_156@O-SER_159@N-H |
| 6778 | 45 | 2.8184 | 19103 | 38.61% | ALA_156@O-LYS_160@N-H |
| 5806 | 15 | 1.7375 | 10088 | 20.39% | ASP_157@O-LYS_160@N-H |
| 4596 | 13 | 1.5907 | 7311 | 14.78% | SER_159@O-GLY_163@N-H |
| 3396 | 371 | 13.2709 | 45068 | 91.09% | VAL_143@O-GLY_165@N-H |
| 7268 | 57 | 5.0645 | 36809 | 74.39% | VAL_145@O-ILE_167@N-H |
| 4802 | 192 | 8.8817 | 42650 | 86.20% | ARG_168@O-GLU_171@N-H |
| 6994 | 33 | 1.6538 | 11567 | 23.38% | ARG_168@O-VAL_172@N-H |
| 5324 | 13 | 1.284 | 6836 | 13.82% | PRO_169@O-VAL_172@N-H |
| 6278 | 25 | 1.6314 | 10242 | 20.70% | GLU_171@O-ASP_174@N-H |
| 9058 | 76 | 3.6006 | 32614 | 65.91% | GLU_171@O-PHE_175@N-H |
| 1789 | 311 | 26.1878 | 46850 | 94.69% | VAL_172@O-LEU_176@N-H |
| 7859 | 85 | 3.4446 | 27071 | 54.71% | GLN_173@O-ASN_177@N-H |
| 4681 | 33 | 1.8458 | 8640 | 17.46% | ASP_174@O-ASN_177@N-H |
| 8144 | 58 | 4.3238 | 35213 | 71.17% | PHE_175@O-GLU_178@N-H |
| 7256 | 241 | 2.4613 | 17859 | 36.09% | PHE_175@O-LEU_179@N-H |
| 8301 | 47 | 1.9275 | 16000 | 32.34% | LEU_176@O-LEU_179@N-H |
| 2573 | 541 | 17.0626 | 43902 | 88.73% | GLN_201@O-ILE_181@N-H |
| 5816 | 124 | 5.7044 | 33177 | 67.05% | ASP_180@O-GLU_183@N-H |
| 10713 | 29 | 2.4763 | 26529 | 53.62% | ILE_181@O-ILE_184@N-H |
| 4130 | 37 | 1.8191 | 7513 | 15.18% | ILE_184@O-ILE_187@N-H |
| 5921 | 25 | 1.5661 | 9273 | 18.74% | GLY_188@O-LEU_191@N-H |
| 3127 | 255 | 14.599 | 45651 | 92.26% | GLY_188@O-ALA_192@N-H |
| 5696 | 125 | 7.1419 | 40680 | 82.22% | SER_189@O-ARG_193@N-H |
| 9406 | 54 | 3.2077 | 30172 | 60.98% | VAL_190@O-ARG_194@N-H |
| 4251 | 156 | 10.4554 | 44446 | 89.83% | LEU_191@O-LEU_195@N-H |
| 5442 | 133 | 7.62 | 41468 | 83.81% | ALA_192@O-ASN_196@N-H |
| 7449 | 58 | 3.383 | 25200 | 50.93% | ARG_193@O-GLU_197@N-H |
| 5970 | 35 | 2.0022 | 11953 | 24.16% | ARG_194@O-GLU_197@N-H |
| 8270 | 27 | 2.2169 | 18334 | 37.05% | ARG_194@O-LEU_198@N-H |
| 8432 | 44 | 1.856 | 15650 | 31.63% | LEU_195@O-LEU_198@N-H |
| 9725 | 37 | 3.6056 | 35064 | 70.87% | ASN_196@O-GLY_199@N-H |
| 5261 | 112 | 7.865 | 41378 | 83.63% | LEU_195@O-ILE_200@N-H |
| 1168 | 567 | 38.4444 | 44903 | 90.75% | LEU_179@O-LEU_203@N-H |
| 5244 | 106 | 8.1766 | 42878 | 86.66% | LYS_202@O-ASP_205@N-H |
| 6976 | 173 | 5.3661 | 37434 | 75.66% | LEU_203@O-ILE_206@N-H |
| 8566 | 63 | 3.7582 | 32193 | 65.06% | ARG_204@O-LEU_207@N-H |
| 7561 | 100 | 4.2477 | 32117 | 64.91% | ASP_205@O-SER_208@N-H |
| 4827 | 105 | 5.5045 | 26570 | 53.70% | ASP_205@O-LYS_209@N-H |
| 4410 | 47 | 2.4757 | 10918 | 22.07% | ILE_206@O-LYS_209@N-H |
| 9948 | 65 | 2.8514 | 28366 | 57.33% | ASN_210@O-GLU_213@N-H |
| 8277 | 115 | 3.5074 | 29031 | 58.67% | ASN_210@O-LEU_214@N-H |
| 4200 | 22 | 1.4352 | 6028 | 12.18% | TYR_211@O-LEU_214@N-H |
| 2039 | 373 | 22.8901 | 46673 | 94.33% | TYR_211@O-GLU_215@N-H |
| 7794 | 79 | 3.2261 | 25144 | 50.82% | ASN_212@O-LYS_216@N-H |
| 4007 | 21 | 1.4567 | 5837 | 11.80% | GLU_213@O-LYS_216@N-H |
| 7388 | 70 | 3.0288 | 22377 | 45.23% | GLU_213@O-ILE_217@N-H |
| 7402 | 28 | 1.718 | 12717 | 25.70% | LEU_214@O-ILE_217@N-H |
| 4382 | 127 | 9.8791 | 43290 | 87.49% | LEU_214@O-THR_218@N-H |
| 8819 | 98 | 3.8626 | 34064 | 68.85% | GLU_215@O-GLY_219@N-H |
| 2525 | 188 | 18.478 | 46657 | 94.30% | GLY_219@O-ALA_223@N-H |
| 3068 | 258 | 14.5512 | 44643 | 90.23% | LYS_220@O-LEU_224@N-H |
| 6254 | 111 | 6.6337 | 41487 | 83.85% | ALA_221@O-TYR_225@N-H |
| 2367 | 491 | 19.7444 | 46735 | 94.45% | LYS_222@O-LEU_226@N-H |
| 2344 | 176 | 20.026 | 46941 | 94.87% | ALA_223@O-LEU_227@N-H |
| 1540 | 298 | 31.0325 | 47790 | 96.59% | LEU_224@O-LYS_228@N-H |
| 4573 | 110 | 9.6035 | 43917 | 88.76% | TYR_225@O-LEU_229@N-H |
| 1843 | 299 | 25.7721 | 47498 | 96.00% | LEU_226@O-ALA_230@N-H |
| 4116 | 144 | 10.1511 | 41782 | 84.44% | LEU_227@O-GLN_231@N-H |
| 6069 | 87 | 6.2829 | 38131 | 77.07% | LEU_229@O-ASN_232@N-H |
| 3612 | 210 | 10.4023 | 37573 | 75.94% | LYS_228@O-LYS_233@N-H |
| 3062 | 21 | 1.9745 | 6046 | 12.22% | LYS_228@O-TYR_234@N-H |
| 4111 | 151 | 3.7397 | 15374 | 31.07% | ILE_99@O-LYS_241@N-H |
| 2795 | 61 | 2.9256 | 8177 | 16.53% | GLU_239@O-LYS_241@N-H |
| 2384 | 525 | 6.5642 | 15649 | 31.63% | ILE_341@O-SER_242@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 2150 | 465 | 8.8451 | 19017 | 38.43% | VAL_101@O-LYS_243@N-H |
| 2724 | 275 | 13.1028 | 35692 | 72.14% | LEU_338@O-HIE_246@N-H |
| 3123 | 192 | 14.6987 | 45904 | 92.77% | VAL_336@O-ARG_248@N-H |
| 391 | 1063 | 125.4604 | 49055 | 99.14% | ILE_334@O-LEU_250@N-H |
| 1076 | 535 | 44.8132 | 48219 | 97.45% | ARG_332@O-LEU_252@N-H |
| 5029 | 194 | 7.3864 | 37146 | 75.07% | VAL_331@O-THR_256@N-H |
| 4017 | 122 | 11.0697 | 44467 | 89.87% | ASP_258@O-VAL_261@N-H |
| 7483 | 88 | 5.0843 | 38046 | 76.89% | ASP_258@O-ILE_262@N-H |
| 1155 | 450 | 41.4684 | 47896 | 96.80% | VAL_259@O-LEU_263@N-H |
| 7417 | 37 | 1.8738 | 13898 | 28.09% | VAL_261@O-TYR_265@N-H |
| 3683 | 142 | 12.3152 | 45357 | 91.67% | ILE_262@O-LEU_266@N-H |
| 873 | 458 | 55.6334 | 48568 | 98.16% | LEU_263@O-LYS_267@N-H |
| 4470 | 114 | 9.8993 | 44250 | 89.43% | PRO_264@O-LYS_268@N-H |
| 3278 | 148 | 13.9433 | 45706 | 92.37% | TYR_265@O-ALA_269@N-H |
| 1859 | 341 | 25.5557 | 47508 | 96.02% | LEU_266@O-ILE_270@N-H |
| 1866 | 497 | 25.2964 | 47203 | 95.40% | LYS_267@O-ASN_271@N-H |
| 2461 | 347 | 18.9045 | 46524 | 94.03% | LYS_268@O-GLU_272@N-H |
| 7119 | 110 | 5.536 | 39411 | 79.65% | ALA_269@O-ALA_273@N-H |
| 176 | 3196 | 280.0341 | 49286 | 99.61% | ILE_270@O-TYR_274@N-H |
| 7630 | 83 | 5.0484 | 38519 | 77.85% | ASN_271@O-ASN_275@N-H |
| 3244 | 31 | 1.8305 | 5938 | 12.00% | GLU_272@O-LYS_276@N-H |
| 5560 | 164 | 7.1293 | 39639 | 80.11% | ALA_273@O-LYS_276@N-H |
| 8419 | 85 | 3.9797 | 33505 | 67.72% | ALA_273@O-VAL_277@N-H |
| 8156 | 21 | 2.0249 | 16515 | 33.38% | VAL_277@O-GLY_279@N-H |
| 7634 | 93 | 2.6551 | 20269 | 40.96% | ILE_342@O-ILE_280@N-H |
| 6665 | 75 | 2.0104 | 13399 | 27.08% | LYS_345@O-ILE_280@N-H |
| 9859 | 66 | 3.2336 | 31880 | 64.43% | ASP_339@O-MET_282@N-H |
| 8337 | 21 | 1.5699 | 13088 | 26.45% | ASN_340@O-MET_282@N-H |
| 9256 | 56 | 3.8499 | 35635 | 72.02% | ASP_339@O-ARG_283@N-H |
| 1269 | 346 | 37.9046 | 48101 | 97.21% | LYS_300@O-ILE_284@N-H |
| 2642 | 232 | 17.6366 | 46596 | 94.17% | LYS_337@O-THR_285@N-H |
| 1285 | 460 | 37.4553 | 48130 | 97.27% | LYS_298@O-VAL_286@N-H |
| 741 | 625 | 65.7341 | 48709 | 98.44% | GLY_335@O-ILE_287@N-H |
| 508 | 685 | 96.3524 | 48947 | 98.92% | LEU_296@O-ALA_288@N-H |
| 1066 | 450 | 45.3715 | 48366 | 97.75% | ARG_333@O-ILE_289@N-H |
| 461 | 1130 | 106.1692 | 48944 | 98.92% | ASP_294@O-MET_290@N-H |
| 3430 | 361 | 13.2426 | 45422 | 91.80% | ASN_330@O-GLU_291@N-H |
| 8628 | 70 | 4.1633 | 35921 | 72.60% | MET_290@O-LEU_293@N-H |
| 1025 | 556 | 47.2283 | 48409 | 97.84% | ALA_288@O-LEU_296@N-H |
| 1092 | 584 | 44.2152 | 48283 | 97.58% | VAL_286@O-LYS_298@N-H |
| 784 | 601 | 62.0523 | 48649 | 98.32% | ILE_284@O-LYS_300@N-H |
| 2598 | 202 | 17.9126 | 46537 | 94.05% | MET_282@O-PHE_302@N-H |
| 1838 | 455 | 25.8107 | 47440 | 95.88% | PRO_281@O-ILE_306@N-H |
| 7575 | 73 | 2.935 | 22233 | 44.93% | SER_307@O-ASN_310@N-H |
| 3842 | 211 | 11.5643 | 44430 | 89.80% | SER_307@O-ALA_311@N-H |
| 5214 | 170 | 4.8506 | 25291 | 51.11% | ILE_308@O-TYR_312@N-H |
| 5705 | 39 | 2.4535 | 13997 | 28.29% | ASP_309@O-LYS_313@N-H |
| 4760 | 12 | 1.3609 | 6478 | 13.09% | ASN_310@O-LYS_313@N-H |
| 3323 | 305 | 13.3226 | 44271 | 89.47% | ASN_310@O-VAL_314@N-H |
| 4456 | 171 | 9.4253 | 41999 | 84.88% | ALA_311@O-ALA_315@N-H |
| 7211 | 105 | 4.1035 | 29590 | 59.80% | TYR_312@O-GLU_316@N-H |
| 4029 | 13 | 1.2946 | 5216 | 10.54% | LYS_313@O-GLU_316@N-H |
| 6758 | 130 | 2.7405 | 18520 | 37.43% | LYS_313@O-ASP_317@N-H |
| 4105 | 7 | 1.2129 | 4979 | 10.06% | VAL_314@O-ASP_317@N-H |
| 5830 | 179 | 7.071 | 41224 | 83.32% | VAL_314@O-LEU_318@N-H |
| 4286 | 106 | 10.3385 | 44311 | 89.56% | ALA_315@O-LEU_319@N-H |
| 5852 | 264 | 6.9725 | 40803 | 82.47% | GLU_316@O-ARG_320@N-H |
| 7479 | 123 | 4.8461 | 36244 | 73.25% | ASP_317@O-GLU_321@N-H |
| 8192 | 81 | 4.2063 | 34458 | 69.64% | LEU_318@O-LEU_322@N-H |
| 4060 | 20 | 1.499 | 6086 | 12.30% | LEU_319@O-LEU_322@N-H |
| 2553 | 317 | 17.698 | 45183 | 91.32% | LEU_319@O-LEU_323@N-H |
| 5903 | 153 | 5.1481 | 30389 | 61.42% | ARG_320@O-VAL_324@N-H |
| 4401 | 24 | 1.7866 | 7863 | 15.89% | GLU_321@O-VAL_324@N-H |
| 4315 | 40 | 2.2137 | 9552 | 19.31% | GLU_321@O-ARG_325@N-H |
| 7230 | 51 | 2.0035 | 14485 | 29.28% | LEU_322@O-ARG_325@N-H |
| 5817 | 30 | 2.3801 | 13845 | 27.98% | LEU_322@O-ASP_326@N-H |
| 7686 | 22 | 1.9863 | 15267 | 30.86% | LEU_323@O-ASP_326@N-H |
| 4368 | 171 | 8.413 | 36748 | 74.27% | ASP_326@O-ARG_328@N-H |
| 752 | 75 | 7.6476 | 5751 | 11.62% | LYS_327@O-ARG_329@N-H |
| 2926 | 247 | 15.7833 | 46182 | 93.34% | THR_256@O-VAL_331@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 962 | 1190 | 50.2214 | 48313 | 97.64% | ILE_289@O-ARG_332@N-H |
| 4962 | 208 | 8.605 | 42698 | 86.30% | ILE_289@O-ARG_333@N-H |
| 342 | 959 | 143.6199 | 49118 | 99.27% | LEU_250@O-ILE_334@N-H |
| 8445 | 51 | 4.4889 | 37909 | 76.62% | ILE_287@O-GLY_335@N-H |
| 699 | 480 | 69.7411 | 48749 | 98.52% | ARG_248@O-VAL_336@N-H |
| 910 | 484 | 53.3231 | 48524 | 98.07% | THR_285@O-LYS_337@N-H |
| 784 | 600 | 62.0625 | 48657 | 98.34% | HIE_246@O-LEU_338@N-H |
| 2352 | 253 | 19.8486 | 46684 | 94.35% | ARG_283@O-ASP_339@N-H |
| 885 | 918 | 54.5153 | 48246 | 97.51% | ILE_280@O-ILE_342@N-H |
| 1466 | 103 | 3.6357 | 5330 | 10.77% | ASN_240@O-ILE_343@N-H |
| 4465 | 157 | 7.9823 | 35641 | 72.03% | ASN_278@O-ASN_344@N-H |
| 7346 | 45 | 1.7985 | 13212 | 26.70% | ILE_342@O-LYS_345@N-H |
| 6942 | 72 | 2.1538 | 14952 | 30.22% | ILE_343@O-LYS_345@N-H |
| 5209 | 13 | 1.7015 | 8863 | 17.91% | ASN_344@O-ASN_347@N-H |
| 3816 | 22 | 2.0503 | 7824 | 15.81% | ASP_350@O-PHE_352@N-H |
| 1908 | 195 | 8.414 | 16054 | 32.45% | LYS_276@O-ASP_353@N-H |

**50°C, 500ns simulation of Dbh**

| # Lifetimes | Max Lifetime | Avg. Lifetime | # Frames | % Occupancy | Donor-Acceptor Pair |
|---|---|---|---|---|---|
| 1280 | 670 | 37.9898 | 48627 | 97.25% | ALA_146@O-VAL_3@N-H |
| 368 | 989 | 134.8587 | 49628 | 99.26% | LEU_109@O-ILE_4@N-H |
| 1032 | 629 | 47.4157 | 48933 | 97.87% | GLY_144@O-PHE_5@N-H |
| 4852 | 347 | 8.0387 | 39004 | 78.01% | ALA_107@O-VAL_6@N-H |
| 1334 | 389 | 36.4123 | 48574 | 97.15% | THR_142@O-ASP_7@N-H |
| 3240 | 102 | 4.0966 | 13273 | 26.55% | ASP_105@O-PHE_8@N-H |
| 4130 | 46 | 2.7242 | 11251 | 22.50% | ASP_7@O-ASP_9@N-H |
| 9254 | 53 | 3.0039 | 27798 | 55.60% | PHE_8@O-PHE_11@N-H |
| 11207 | 42 | 2.1991 | 24645 | 49.29% | TYR_10@O-ALA_13@N-H |
| 10593 | 54 | 2.9317 | 31055 | 62.11% | TYR_10@O-GLN_14@N-H |
| 7559 | 10 | 1.351 | 10212 | 20.42% | PHE_11@O-GLN_14@N-H |
| 5272 | 228 | 8.1959 | 43209 | 86.42% | PHE_11@O-VAL_15@N-H |
| 5808 | 94 | 7.1865 | 41739 | 83.48% | PHE_12@O-GLU_16@N-H |
| 9837 | 47 | 3.0845 | 30342 | 60.68% | ALA_13@O-GLU_17@N-H |
| 4082 | 12 | 1.2626 | 5154 | 10.31% | GLN_14@O-GLU_17@N-H |
| 8586 | 42 | 2.8452 | 24429 | 48.86% | GLN_14@O-VAL_18@N-H |
| 7599 | 55 | 2.0371 | 15480 | 30.96% | VAL_15@O-VAL_18@N-H |
| 6314 | 76 | 6.5969 | 41653 | 83.31% | VAL_15@O-LEU_19@N-H |
| 10675 | 20 | 1.8004 | 19219 | 38.44% | GLU_16@O-ASN_20@N-H |
| 6281 | 8 | 1.4117 | 8867 | 17.73% | GLU_17@O-ASN_20@N-H |
| 4191 | 123 | 10.3737 | 43476 | 86.95% | ASN_20@O-TYR_23@N-H |
| 8993 | 40 | 3.5301 | 31746 | 63.49% | PRO_21@O-LYS_24@N-H |
| 8767 | 50 | 3.3055 | 28979 | 57.96% | TYR_23@O-LYS_26@N-H |
| 1559 | 282 | 31.0359 | 48385 | 96.77% | ILE_72@O-VAL_29@N-H |
| 1026 | 407 | 47.7037 | 48944 | 97.89% | THR_45@O-VAL_30@N-H |
| 671 | 877 | 73.4814 | 49306 | 98.61% | VAL_74@O-SER_31@N-H |
| 1780 | 287 | 26.9826 | 48029 | 96.06% | ALA_42@O-VAL_32@N-H |
| 4841 | 44 | 3.1535 | 15266 | 30.53% | ARG_36@O-LYS_38@N-H |
| 3323 | 134 | 8.4141 | 27960 | 55.92% | THR_37@O-SER_40@N-H |
| 801 | 511 | 61.392 | 49175 | 98.35% | MET_59@O-VAL_43@N-H |
| 4586 | 91 | 9.7656 | 44785 | 89.57% | VAL_30@O-ALA_44@N-H |
| 10313 | 36 | 3.2206 | 33214 | 66.43% | VAL_30@O-THR_45@N-H |
| 1075 | 540 | 45.4828 | 48894 | 97.79% | LEU_28@O-ASN_47@N-H |
| 4818 | 7 | 1.1812 | 5691 | 11.38% | ASN_47@O-ALA_50@N-H |
| 2086 | 225 | 22.8993 | 47768 | 95.54% | ASN_47@O-ARG_51@N-H |
| 4958 | 193 | 8.6668 | 42970 | 85.94% | TYR_48@O-LYS_52@N-H |
| 9023 | 20 | 1.9705 | 17780 | 35.56% | GLU_49@O-LEU_53@N-H |
| 8833 | 40 | 1.7071 | 15079 | 30.16% | ALA_50@O-LEU_53@N-H |
| 7290 | 9 | 1.4379 | 10482 | 20.96% | ALA_50@O-GLY_54@N-H |
| 10684 | 28 | 2.8361 | 30301 | 60.60% | ARG_51@O-GLY_54@N-H |
| 7096 | 82 | 5.3875 | 38230 | 76.46% | ALA_50@O-VAL_55@N-H |
| 4546 | 87 | 9.6447 | 43845 | 87.69% | VAL_43@O-GLY_58@N-H |
| 9457 | 14 | 1.6111 | 15236 | 30.47% | LYS_56@O-MET_59@N-H |
| 2941 | 407 | 15.7375 | 46284 | 92.57% | GLY_41@O-ILE_61@N-H |
| 11575 | 15 | 1.8407 | 21306 | 42.61% | PRO_60@O-LYS_63@N-H |
| 2458 | 182 | 19.2559 | 47331 | 94.66% | PRO_60@O-ALA_64@N-H |
| 500 | 1361 | 98.986 | 49493 | 98.99% | ILE_61@O-MET_65@N-H |
| 7218 | 13 | 1.3962 | 10078 | 20.16% | ILE_62@O-GLN_66@N-H |
| 11261 | 26 | 2.3128 | 26044 | 52.09% | LYS_63@O-GLN_66@N-H |
| 9289 | 48 | 3.6953 | 34326 | 68.65% | LYS_63@O-ILE_67@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 7176 | 10 | 1.3612 | 9768 | 19.54% | ALA_64@O-ILE_67@N-H |
| 2816 | 266 | 16.6392 | 46856 | 93.71% | ALA_64@O-ALA_68@N-H |
| 12088 | 18 | 2.308 | 27899 | 55.80% | ALA_68@O-ALA_71@N-H |
| 698 | 517 | 70.5659 | 49255 | 98.51% | PRO_27@O-ILE_72@N-H |
| 2425 | 238 | 19.5072 | 47305 | 94.61% | VAL_29@O-VAL_74@N-H |
| 4959 | 21 | 1.7582 | 8719 | 17.44% | SER_31@O-MET_76@N-H |
| 4770 | 68 | 5.7698 | 27522 | 55.04% | TYR_33@O-MET_76@N-H |
| 3201 | 12 | 1.5711 | 5029 | 10.06% | MET_76@O-LYS_78@N-H |
| 9983 | 45 | 3.3895 | 33837 | 67.67% | ARG_77@O-ILE_80@N-H |
| 9229 | 50 | 2.2872 | 21109 | 42.22% | ARG_77@O-TYR_81@N-H |
| 1921 | 227 | 24.7881 | 47618 | 95.24% | LYS_78@O-GLU_82@N-H |
| 11095 | 26 | 2.7988 | 31053 | 62.11% | PRO_79@O-ALA_83@N-H |
| 4592 | 161 | 9.5379 | 43798 | 87.60% | ILE_80@O-PHE_84@N-H |
| 2874 | 163 | 16.2669 | 46751 | 93.50% | TYR_81@O-SER_85@N-H |
| 3077 | 286 | 15.0442 | 46291 | 92.58% | GLU_82@O-ASN_86@N-H |
| 8209 | 61 | 4.5781 | 37582 | 75.16% | ALA_83@O-ARG_87@N-H |
| 4154 | 125 | 10.7768 | 44767 | 89.53% | PHE_84@O-ILE_88@N-H |
| 1336 | 517 | 36.2672 | 48453 | 96.91% | SER_85@O-MET_89@N-H |
| 5828 | 75 | 7.2437 | 42216 | 84.43% | ASN_86@O-ASN_90@N-H |
| 9738 | 40 | 3.1142 | 30326 | 60.65% | ARG_87@O-LEU_91@N-H |
| 5075 | 15 | 1.415 | 7181 | 14.36% | ILE_88@O-LEU_91@N-H |
| 2855 | 261 | 16.2242 | 46320 | 92.64% | ILE_88@O-LEU_92@N-H |
| 3708 | 144 | 12.1713 | 45131 | 90.26% | MET_89@O-ASN_93@N-H |
| 9182 | 25 | 2.0753 | 19055 | 38.11% | ASN_90@O-LYS_94@N-H |
| 8917 | 72 | 1.9131 | 17059 | 34.12% | LEU_91@O-LYS_94@N-H |
| 6384 | 23 | 1.8235 | 11641 | 23.28% | LEU_91@O-HIE_95@N-H |
| 9512 | 46 | 2.4514 | 23318 | 46.64% | LEU_92@O-HIE_95@N-H |
| 7643 | 62 | 4.6745 | 35727 | 71.45% | LEU_92@O-ALA_96@N-H |
| 3189 | 367 | 13.0546 | 41631 | 83.26% | ASP_110@O-ASP_97@N-H |
| 5799 | 91 | 4.4941 | 26061 | 52.12% | ASP_110@O-LYS_98@N-H |
| 1058 | 732 | 41.7391 | 44160 | 88.32% | TYR_108@O-GLU_100@N-H |
| 920 | 117 | 9.5652 | 8800 | 17.60% | LYS_241@O-VAL_101@N-H |
| 3968 | 47 | 2.1381 | 8484 | 16.97% | GLU_106@O-ALA_102@N-H |
| 3861 | 61 | 3.4561 | 13344 | 26.69% | GLU_106@O-SER_103@N-H |
| 6855 | 143 | 4.6468 | 31854 | 63.71% | SER_103@O-GLU_106@N-H |
| 5177 | 167 | 3.3975 | 17589 | 35.18% | VAL_6@O-ALA_107@N-H |
| 2889 | 668 | 14.172 | 40943 | 81.89% | GLU_100@O-TYR_108@N-H |
| 1931 | 324 | 24.7954 | 47880 | 95.76% | ILE_4@O-LEU_109@N-H |
| 1438 | 372 | 30.9305 | 44478 | 88.96% | LYS_98@O-ASP_110@N-H |
| 8297 | 67 | 4.2567 | 35318 | 70.64% | ILE_2@O-VAL_111@N-H |
| 7530 | 30 | 1.6248 | 12235 | 24.47% | VAL_111@O-LYS_114@N-H |
| 7624 | 96 | 5.071 | 38661 | 77.32% | VAL_111@O-VAL_115@N-H |
| 5838 | 8 | 1.2845 | 7499 | 15.00% | THR_112@O-VAL_115@N-H |
| 10146 | 24 | 2.2274 | 22599 | 45.20% | ASN_113@O-GLU_116@N-H |
| 7791 | 77 | 4.9702 | 38723 | 77.45% | THR_112@O-GLY_117@N-H |
| 8316 | 43 | 4.6611 | 38762 | 77.52% | VAL_115@O-ASN_118@N-H |
| 5906 | 11 | 1.2843 | 7585 | 15.17% | ASN_118@O-ASN_121@N-H |
| 10105 | 31 | 3.2947 | 33293 | 66.59% | ASN_118@O-GLY_122@N-H |
| 4847 | 104 | 9.0404 | 43819 | 87.64% | PHE_119@O-ILE_123@N-H |
| 8432 | 63 | 3.8254 | 32256 | 64.51% | GLU_120@O-GLU_124@N-H |
| 6831 | 66 | 6.0767 | 41510 | 83.02% | ASN_121@O-LEU_125@N-H |
| 3172 | 139 | 14.675 | 46549 | 93.10% | GLY_122@O-ALA_126@N-H |
| 2588 | 219 | 18.1847 | 47062 | 94.12% | ILE_123@O-ARG_127@N-H |
| 3671 | 210 | 12.4634 | 45753 | 91.51% | GLU_124@O-LYS_128@N-H |
| 976 | 512 | 50.1936 | 48989 | 97.98% | LEU_125@O-ILE_129@N-H |
| 3116 | 210 | 14.775 | 46039 | 92.08% | ALA_126@O-LYS_130@N-H |
| 3787 | 193 | 11.9638 | 45307 | 90.61% | ARG_127@O-GLN_131@N-H |
| 4397 | 273 | 10.1542 | 44648 | 89.30% | LYS_128@O-GLU_132@N-H |
| 4311 | 131 | 10.3046 | 44423 | 88.85% | ILE_129@O-ILE_133@N-H |
| 8964 | 82 | 3.3569 | 30091 | 60.18% | LYS_130@O-LEU_134@N-H |
| 7395 | 81 | 5.2158 | 38571 | 77.14% | GLN_131@O-GLU_135@N-H |
| 3946 | 13 | 1.3163 | 5194 | 10.39% | GLU_132@O-GLU_135@N-H |
| 1601 | 471 | 29.9631 | 47971 | 95.94% | GLU_132@O-LYS_136@N-H |
| 9840 | 44 | 3.3048 | 32519 | 65.04% | ILE_133@O-GLU_137@N-H |
| 5138 | 25 | 1.4905 | 7658 | 15.32% | ILE_133@O-LYS_138@N-H |
| 5160 | 16 | 1.4758 | 7615 | 15.23% | ILE_133@O-ILE_139@N-H |
| 5660 | 14 | 1.3984 | 7915 | 15.83% | GLU_137@O-ILE_139@N-H |
| 1536 | 215 | 5.7826 | 8882 | 17.76% | ASP_7@O-THR_142@N-H |
| 4713 | 105 | 9.3866 | 44239 | 88.48% | GLY_163@O-VAL_143@N-H |
| 2035 | 442 | 23.4983 | 47819 | 95.64% | PHE_5@O-GLY_144@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 7540 | 82 | 5.2922 | 39903 | 79.81% | GLY_165@O-VAL_145@N-H |
| 923 | 617 | 53.1354 | 49044 | 98.09% | VAL_3@O-ALA_146@N-H |
| 6447 | 61 | 6.4793 | 41772 | 83.54% | ASN_148@O-ALA_152@N-H |
| 6570 | 85 | 5.6848 | 37349 | 74.70% | LYS_149@O-LYS_153@N-H |
| 3256 | 283 | 13.9982 | 45578 | 91.16% | ILE_150@O-ILE_154@N-H |
| 1498 | 442 | 32.1295 | 48130 | 96.26% | LEU_151@O-ILE_155@N-H |
| 11019 | 33 | 2.6293 | 28972 | 57.94% | ALA_152@O-ALA_156@N-H |
| 4530 | 8 | 1.1583 | 5247 | 10.49% | LYS_153@O-ALA_156@N-H |
| 1922 | 332 | 24.5281 | 47143 | 94.29% | LYS_153@O-ASP_157@N-H |
| 7075 | 113 | 5.3952 | 38171 | 76.34% | ILE_154@O-LYS_158@N-H |
| 6712 | 19 | 1.4184 | 9520 | 19.04% | ILE_155@O-SER_159@N-H |
| 10236 | 37 | 2.9004 | 29688 | 59.38% | ALA_156@O-SER_159@N-H |
| 7442 | 46 | 2.6432 | 19671 | 39.34% | ALA_156@O-LYS_160@N-H |
| 5958 | 15 | 1.5898 | 9472 | 18.94% | ASP_157@O-LYS_160@N-H |
| 4566 | 12 | 1.5009 | 6853 | 13.71% | SER_159@O-GLY_163@N-H |
| 4896 | 231 | 8.6842 | 42518 | 85.04% | VAL_143@O-GLY_165@N-H |
| 7839 | 100 | 4.6648 | 36567 | 73.13% | VAL_145@O-ILE_167@N-H |
| 5844 | 204 | 6.8441 | 39997 | 79.99% | ARG_168@O-GLU_171@N-H |
| 5939 | 41 | 1.5969 | 9484 | 18.97% | ARG_168@O-VAL_172@N-H |
| 5376 | 9 | 1.3021 | 7000 | 14.00% | PRO_169@O-VAL_172@N-H |
| 6254 | 15 | 1.6431 | 10276 | 20.55% | GLU_171@O-ASP_174@N-H |
| 8819 | 88 | 3.9976 | 35255 | 70.51% | GLU_171@O-PHE_175@N-H |
| 2466 | 250 | 17.0576 | 42064 | 84.13% | VAL_172@O-LEU_176@N-H |
| 8373 | 61 | 2.864 | 23980 | 47.96% | GLN_173@O-ASN_177@N-H |
| 5350 | 30 | 1.725 | 9229 | 18.46% | ASP_174@O-ASN_177@N-H |
| 3782 | 19 | 1.3773 | 5209 | 10.42% | ASP_174@O-GLU_178@N-H |
| 9061 | 68 | 3.7248 | 33750 | 67.50% | PHE_175@O-GLU_178@N-H |
| 8287 | 71 | 2.3164 | 19196 | 38.39% | PHE_175@O-LEU_179@N-H |
| 8393 | 24 | 1.7865 | 14994 | 29.99% | LEU_176@O-LEU_179@N-H |
| 4453 | 261 | 9.7673 | 43494 | 86.99% | GLN_201@O-ILE_181@N-H |
| 3730 | 223 | 6.0957 | 22737 | 45.47% | ASP_180@O-GLU_183@N-H |
| 7662 | 24 | 2.1397 | 16394 | 32.79% | ILE_181@O-ILE_184@N-H |
| 6145 | 16 | 1.4885 | 9147 | 18.29% | GLY_188@O-LEU_191@N-H |
| 2629 | 332 | 17.8167 | 46840 | 93.68% | GLY_188@O-ALA_192@N-H |
| 6204 | 140 | 6.6626 | 41335 | 82.67% | SER_189@O-ARG_193@N-H |
| 9066 | 107 | 3.6492 | 33084 | 66.17% | VAL_190@O-ARG_194@N-H |
| 2951 | 216 | 15.7835 | 46577 | 93.15% | LEU_191@O-LEU_195@N-H |
| 5148 | 139 | 8.3411 | 42940 | 85.88% | ALA_192@O-ASN_196@N-H |
| 7270 | 117 | 4.6014 | 33452 | 66.90% | ARG_193@O-GLU_197@N-H |
| 4128 | 37 | 1.6659 | 6877 | 13.75% | ARG_194@O-GLU_197@N-H |
| 9087 | 33 | 2.8797 | 26168 | 52.34% | ARG_194@O-LEU_198@N-H |
| 7148 | 42 | 1.661 | 11873 | 23.75% | LEU_195@O-LEU_198@N-H |
| 9918 | 38 | 3.5874 | 35580 | 71.16% | ASN_196@O-GLY_199@N-H |
| 5703 | 85 | 7.3277 | 41790 | 83.58% | LEU_195@O-ILE_200@N-H |
| 1808 | 368 | 25.0796 | 45344 | 90.69% | LEU_179@O-LEU_203@N-H |
| 8004 | 165 | 4.7777 | 38241 | 76.48% | LYS_202@O-ASP_205@N-H |
| 7619 | 59 | 4.9604 | 37793 | 75.59% | LEU_203@O-ILE_206@N-H |
| 1678 | 428 | 20.093 | 33716 | 67.43% | LEU_203@O-LEU_207@N-H |
| 2956 | 27 | 2.9185 | 8627 | 17.25% | ARG_204@O-LEU_207@N-H |
| 7669 | 20 | 2.2794 | 17481 | 34.96% | ARG_204@O-SER_208@N-H |
| 3042 | 50 | 3.3974 | 10335 | 20.67% | ASP_205@O-SER_208@N-H |
| 1504 | 49 | 4.1602 | 6257 | 12.51% | ASP_205@O-LYS_209@N-H |
| 7011 | 53 | 4.613 | 32342 | 64.68% | ILE_206@O-LYS_209@N-H |
| 10673 | 40 | 2.4849 | 26521 | 53.04% | ASN_210@O-GLU_213@N-H |
| 6527 | 125 | 5.8269 | 38032 | 76.06% | ASN_210@O-LEU_214@N-H |
| 3195 | 207 | 14.4 | 46008 | 92.02% | TYR_211@O-GLU_215@N-H |
| 8790 | 33 | 2.5677 | 22570 | 45.14% | ASN_212@O-LYS_216@N-H |
| 5181 | 16 | 1.5258 | 7905 | 15.81% | GLU_213@O-LYS_216@N-H |
| 5751 | 112 | 5.0134 | 28832 | 57.66% | GLU_213@O-ILE_217@N-H |
| 5711 | 18 | 1.7584 | 10042 | 20.08% | LEU_214@O-ILE_217@N-H |
| 5432 | 154 | 7.9523 | 43197 | 86.39% | LEU_214@O-THR_218@N-H |
| 10188 | 48 | 2.7531 | 28049 | 56.10% | GLU_215@O-GLY_219@N-H |
| 1868 | 223 | 25.705 | 48017 | 96.03% | GLY_219@O-ALA_223@N-H |
| 3185 | 182 | 14.1862 | 45183 | 90.37% | LYS_220@O-LEU_224@N-H |
| 10022 | 48 | 3.302 | 33093 | 66.19% | ALA_221@O-TYR_225@N-H |
| 1843 | 349 | 26.0239 | 47962 | 95.92% | LYS_222@O-LEU_226@N-H |
| 3480 | 136 | 13.2612 | 46149 | 92.30% | ALA_223@O-LEU_227@N-H |
| 7330 | 128 | 4.7513 | 34827 | 69.65% | LEU_224@O-LYS_228@N-H |
| 4234 | 19 | 1.5125 | 6404 | 12.81% | TYR_225@O-LYS_228@N-H |
| 8176 | 182 | 3.8158 | 31198 | 62.40% | TYR_225@O-LEU_229@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 5028 | 26 | 1.4093 | 7086 | 14.17% | LEU_226@O-LEU_229@N-H |
| 4729 | 123 | 9.1664 | 43348 | 86.70% | LEU_226@O-ALA_230@N-H |
| 3909 | 123 | 9.9872 | 39040 | 78.08% | LEU_227@O-GLN_231@N-H |
| 5736 | 45 | 4.6065 | 26423 | 52.85% | LEU_229@O-ASN_232@N-H |
| 6579 | 72 | 2.6635 | 17523 | 35.05% | LYS_228@O-LYS_233@N-H |
| 3332 | 57 | 5.2638 | 17539 | 35.08% | LYS_228@O-TYR_234@N-H |
| 2446 | 21 | 2.1345 | 5221 | 10.44% | ASN_232@O-TYR_234@N-H |
| 2552 | 98 | 4.3781 | 11173 | 22.35% | GLU_239@O-LYS_241@N-H |
| 3378 | 359 | 13.2315 | 44696 | 89.39% | VAL_336@O-ARG_248@N-H |
| 949 | 984 | 51.6407 | 49007 | 98.01% | ILE_334@O-LEU_250@N-H |
| 1274 | 359 | 37.3893 | 47634 | 95.27% | ARG_332@O-LEU_252@N-H |
| 4220 | 199 | 9.4526 | 39890 | 79.78% | VAL_331@O-THR_256@N-H |
| 6107 | 88 | 6.684 | 40819 | 81.64% | ASP_258@O-VAL_261@N-H |
| 7821 | 95 | 4.3953 | 34376 | 68.75% | ASP_258@O-ILE_262@N-H |
| 3467 | 67 | 1.4442 | 5007 | 10.01% | VAL_259@O-ILE_262@N-H |
| 2889 | 246 | 15.722 | 45421 | 90.84% | VAL_259@O-LEU_263@N-H |
| 5456 | 16 | 1.7311 | 9445 | 18.89% | VAL_261@O-TYR_265@N-H |
| 4606 | 7 | 1.2414 | 5718 | 11.44% | ILE_262@O-TYR_265@N-H |
| 3793 | 127 | 12.048 | 45698 | 91.40% | ILE_262@O-LEU_266@N-H |
| 1335 | 391 | 36.394 | 48586 | 97.17% | LEU_263@O-LYS_267@N-H |
| 5873 | 115 | 7.234 | 42485 | 84.97% | PRO_264@O-LYS_268@N-H |
| 4214 | 132 | 10.6972 | 45078 | 90.16% | TYR_265@O-ALA_269@N-H |
| 2975 | 226 | 15.7176 | 46760 | 93.52% | LEU_266@O-ILE_270@N-H |
| 2953 | 245 | 15.617 | 46117 | 92.23% | LYS_267@O-ASN_271@N-H |
| 3356 | 372 | 13.5837 | 45587 | 91.17% | LYS_268@O-GLU_272@N-H |
| 6351 | 160 | 6.2601 | 39758 | 79.52% | ALA_269@O-ALA_273@N-H |
| 615 | 603 | 80.213 | 49331 | 98.66% | ILE_270@O-TYR_274@N-H |
| 9597 | 71 | 3.4444 | 33056 | 66.11% | ASN_271@O-ASN_275@N-H |
| 4684 | 10 | 1.2882 | 6034 | 12.07% | GLU_272@O-ASN_275@N-H |
| 6900 | 79 | 2.818 | 19444 | 38.89% | GLU_272@O-LYS_276@N-H |
| 8753 | 28 | 2.3501 | 20570 | 41.14% | ALA_273@O-LYS_276@N-H |
| 8124 | 52 | 2.7312 | 22188 | 44.38% | ALA_273@O-VAL_277@N-H |
| 4264 | 34 | 1.4566 | 6211 | 12.42% | TYR_274@O-VAL_277@N-H |
| 2407 | 122 | 4.236 | 10196 | 20.39% | ILE_342@O-ILE_280@N-H |
| 2095 | 36 | 3.6368 | 7619 | 15.24% | LYS_345@O-ILE_280@N-H |
| 861 | 107 | 16.295 | 14030 | 28.06% | ASN_347@O-ILE_280@N-H |
| 9763 | 51 | 3.5602 | 34758 | 69.52% | ASP_339@O-MET_282@N-H |
| 7674 | 11 | 1.401 | 10751 | 21.50% | ASN_340@O-MET_282@N-H |
| 9852 | 43 | 3.5903 | 35372 | 70.74% | ASP_339@O-ARG_283@N-H |
| 1412 | 328 | 34.3534 | 48507 | 97.01% | LYS_300@O-ILE_284@N-H |
| 3422 | 170 | 13.4988 | 46193 | 92.39% | LYS_337@O-THR_285@N-H |
| 736 | 718 | 66.8872 | 49229 | 98.46% | LYS_298@O-VAL_286@N-H |
| 747 | 1013 | 65.905 | 49231 | 98.46% | GLY_335@O-ILE_287@N-H |
| 608 | 577 | 81.199 | 49369 | 98.74% | LEU_296@O-ALA_288@N-H |
| 1451 | 338 | 33.4087 | 48476 | 96.95% | ARG_333@O-ILE_289@N-H |
| 1078 | 1063 | 45.1178 | 48637 | 97.27% | ASP_294@O-MET_290@N-H |
| 1703 | 289 | 18.8015 | 32019 | 64.04% | ASN_330@O-GLU_291@N-H |
| 10423 | 28 | 2.7402 | 28561 | 57.12% | MET_290@O-LEU_293@N-H |
| 5588 | 17 | 1.4467 | 8084 | 16.17% | GLU_291@O-LEU_293@N-H |
| 1561 | 312 | 30.9488 | 48311 | 96.62% | ALA_288@O-LEU_296@N-H |
| 1147 | 717 | 42.476 | 48720 | 97.44% | VAL_286@O-LYS_298@N-H |
| 873 | 377 | 56.2451 | 49102 | 98.20% | ILE_284@O-LYS_300@N-H |
| 3017 | 192 | 15.4498 | 46612 | 93.22% | MET_282@O-PHE_302@N-H |
| 1694 | 372 | 28.4256 | 48153 | 96.31% | PRO_281@O-ILE_306@N-H |
| 8855 | 53 | 2.1211 | 18782 | 37.56% | SER_307@O-ASN_310@N-H |
| 3395 | 261 | 13.3962 | 45480 | 90.96% | SER_307@O-ALA_311@N-H |
| 4293 | 142 | 8.7107 | 37395 | 74.79% | ILE_308@O-TYR_312@N-H |
| 6754 | 30 | 2.4861 | 16791 | 33.58% | ASP_309@O-LYS_313@N-H |
| 5679 | 15 | 1.4172 | 8048 | 16.10% | ASN_310@O-LYS_313@N-H |
| 3240 | 264 | 13.9645 | 45245 | 90.49% | ASN_310@O-VAL_314@N-H |
| 4214 | 316 | 10.5389 | 44411 | 88.82% | ALA_311@O-ALA_315@N-H |
| 8576 | 64 | 3.2836 | 28160 | 56.32% | TYR_312@O-GLU_316@N-H |
| 4935 | 17 | 1.3386 | 6606 | 13.21% | LYS_313@O-GLU_316@N-H |
| 7847 | 86 | 3.1847 | 24990 | 49.98% | LYS_313@O-ASP_317@N-H |
| 5775 | 102 | 7.1479 | 41279 | 82.56% | VAL_314@O-LEU_318@N-H |
| 5959 | 89 | 7.1898 | 42844 | 85.69% | ALA_315@O-LEU_319@N-H |
| 6099 | 322 | 6.6091 | 40309 | 80.62% | GLU_316@O-ARG_320@N-H |
| 8096 | 70 | 4.0497 | 32786 | 65.57% | ASP_317@O-GLU_321@N-H |
| 8412 | 113 | 3.8581 | 32454 | 64.91% | LEU_318@O-LEU_322@N-H |
| 5279 | 19 | 1.577 | 8325 | 16.65% | LEU_319@O-LEU_322@N-H |

| | | | | | |
|---|---|---|---|---|---|
| 5203 | 257 | 7.9416 | 41320 | 82.64% | LEU_319@O-LEU_323@N-H |
| 7242 | 93 | 3.5301 | 25565 | 51.13% | ARG_320@O-VAL_324@N-H |
| 2784 | 44 | 2.3718 | 6603 | 13.21% | GLU_321@O-VAL_324@N-H |
| 2702 | 163 | 5.0259 | 13580 | 27.16% | ARG_320@O-ARG_325@N-H |
| 1289 | 42 | 4.3988 | 5670 | 11.34% | ASP_326@O-ARG_328@N-H |
| 2105 | 107 | 3.2128 | 6763 | 13.53% | LEU_323@O-ARG_329@N-H |
| 3012 | 35 | 2.1763 | 6555 | 13.11% | LYS_327@O-ARG_329@N-H |
| 4537 | 800 | 9.4022 | 42658 | 85.32% | THR_256@O-VAL_331@N-H |
| 1330 | 484 | 36.3173 | 48302 | 96.60% | ILE_289@O-ARG_332@N-H |
| 3947 | 170 | 11.1809 | 44131 | 88.26% | ILE_289@O-ARG_333@N-H |
| 801 | 678 | 61.3021 | 49103 | 98.21% | LEU_250@O-ILE_334@N-H |
| 7941 | 69 | 4.9746 | 39503 | 79.01% | ILE_287@O-GLY_335@N-H |
| 1537 | 377 | 31.2023 | 47958 | 95.92% | ARG_248@O-VAL_336@N-H |
| 1780 | 277 | 27.0393 | 48130 | 96.26% | THR_285@O-LYS_337@N-H |
| 313 | 278 | 34.9712 | 10946 | 21.89% | HIE_246@O-LEU_338@N-H |
| 2436 | 161 | 19.461 | 47407 | 94.81% | ARG_283@O-ASP_339@N-H |
| 634 | 505 | 53.4685 | 33899 | 67.80% | ILE_280@O-ILE_342@N-H |
| 1191 | 61 | 4.8405 | 5765 | 11.53% | ASN_278@O-ASN_344@N-H |
| 5872 | 38 | 1.689 | 9918 | 19.84% | ILE_342@O-LYS_345@N-H |
| 6575 | 27 | 1.8785 | 12351 | 24.70% | ILE_343@O-LYS_345@N-H |
| 2862 | 34 | 2.841 | 8131 | 16.26% | ILE_342@O-THR_346@N-H |
| 1224 | 129 | 7.8423 | 9599 | 19.20% | ASN_278@O-SER_349@N-H |
| 4294 | 50 | 3.8763 | 16645 | 33.29% | PHE_352@O-ILE_354@N-H |

# Appendix E: Selected AMBER scripts for MD simulations of Dbh

## Restrained Minimization:

```
Minimization of restrained Dbh
 &cntrl
  imin=1,maxcyc=2000,ncyc=1000,
  cut=8.0,ntb=1,
  ntc=2,ntf=2,
  ntpr=20,
  ntr=1,restraintmask=':1-34,39-344',
  restraint_wt=2.0
 /
```

## Restrained Heating:

```
Heat Dbh with restraints
 &cntrl
  imin=0,irest=0,ntx=1,
  nstlim=50000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0, ntb=1,
  ntpr=500, ntwx=500,
  ntt=3, gamma_ln=2.0,
  tempi=0.0, temp0=308.0,
  ntr=1, restraintmask=':1-354',
  restraint_wt=2.0,
  ntp=0,
  ig=-1
 /
 &wt TYPE='TEMP0', istep1=0, istep2=50000,
  value1=0.0, value2=308.0, /
 &wt TYPE='END' /
```

## Restrained Density Equilibration:

```
Restrained density equilibration of Dbh
 &cntrl
  imin=0,irest=1,ntx=5,
  nstlim=50000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0,ntb=2,ntp=1,taup=2.0,
  ntpr=5000,ntwx=5000,
  ntt=3,gamma_ln=2.0,
  temp0=308.0,
  ntr=1,restraintmask=':1-354',
  restraint_wt=2.0,
  ig=-1
 /
```

## Protein Equilibration, Gradual Decrease in Restraints:

```
Restraint back-off equilibration of Dbh
 &cntrl
  imin=0,irest=1,ntx=5,
  nstlim=500000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0,ntb=2,ntp=1,taup=2.0,
  ntpr=50000,ntwx=5000,
  ntt=3,gamma_ln=2.0,
  temp0=308.0,
  ntr-1, restraintmask=':1-354',
  restraint_wt=2.0,
  ig=-1,
 /
 &wt TYPE='REST', istep1=0, istep2=125000,
value1=2.0, value2=1.5, /
 &wt TYPE='REST', istep1=125001,
istep2=250000, value1=1.5, value2=1.0, /
 &wt TYPE='REST', istep1=250001,
istep2=375000, value1=1.0, value2=0.5, /
 &wt TYPE='REST', istep1=375001,
istep2=500000, value1=0.5, value2=0.0, /
 &wt TYPE='END' /
```

## Free Protein Equilibration:

```
Free equilibration of Dbh
 &cntrl
  imin=0,irest=1,ntx=5,
  nstlim=2000000,dt=0.002,
  ntc=2,ntf=2,
  cut=8.0,ntb=2,ntp=1,taup=2.0,
  ntpr=50000,ntwx=5000,
  ntt=3,gamma_ln=2.0,
  temp0=308.0,
  ig=-1,
 /
```

## Production Run:

```
NTV Production Dbh
 &cntrl
  irest=1,ntx=5,
  nstlim=5000000,dt=0.002,
  ntc=2,ntf=2,
  ntb=1,
  ntxo=2
  ntpr=5000,ntwx=5000,ntwr=500000,
  ntt=1, tautp=10.0,
  temp0=308.0, ig=-1, / & ewald /
```

# Appendix F: Temperature Coefficients of Amide Hydrogens of Dbh calculated from chemical shifts at 35°C, 45°C, 50°C, 55°C, and 65°C

| Res # | Res Type | $^1$Hδ, 35°C (ppm) | $^1$Hδ, 45°C (ppm) | $^1$Hδ, 50°C (ppm) | $^1$Hδ, 55°C (ppm) | $^1$Hδ, 65°C (ppm) | Temp. Coeff. (ppb/K) | Temp. Coeff. error |
|---|---|---|---|---|---|---|---|---|
| **2** | **Ile** | 8.351 | 8.258 | 8.250 | 8.241 | | -5.60 | 1.50 |
| **3** | **Val** | 8.378 | 8.351 | 8.358 | 8.321 | 8.293 | -2.85 | 0.51 |
| **4** | **Ile** | 9.067 | 9.030 | 9.021 | 9.010 | 8.986 | -2.64 | 0.21 |
| **5** | **Phe** | 9.821 | 9.802 | 9.783 | 9.770 | 9.745 | -2.61 | 0.13 |
| **6** | **Val** | 8.430 | 8.398 | 8.381 | 8.369 | 8.341 | -2.96 | 0.09 |
| **7** | **Asp** | 8.393 | 8.380 | 8.373 | 8.371 | 8.359 | -1.11 | 0.08 |
| **8** | **Phe** | 8.342 | 8.310 | 8.293 | 8.276 | 8.244 | -3.29 | 0.01 |
| **9** | **Asp** | 7.429 | 7.426 | 7.418 | 7.426 | 7.436 | 0.21 | 0.30 |
| **10** | **Tyr** | 8.876 | 8.812 | 8.778 | 8.726 | | -7.33 | 0.62 |
| **11** | **Phe** | 8.021 | 8.002 | 7.971 | 7.965 | 7.968 | -1.96 | 0.60 |
| **12** | **Phe** | 8.762 | 8.685 | 8.630 | 8.574 | 8.463 | -10.08 | 0.52 |
| **13** | **Ala** | 6.869 | 6.888 | 6.889 | 6.905 | 6.920 | 1.70 | 0.16 |
| **14** | **Gln** | 8.218 | 8.186 | 8.161 | 8.128 | 8.099 | -4.15 | 0.33 |
| **15** | **Val** | 8.137 | 8.108 | 8.092 | 8.078 | 8.044 | -3.09 | 0.07 |
| **16** | **Glu** | 7.127 | 7.114 | 7.108 | 7.108 | 7.098 | -0.93 | 0.11 |
| **17** | **Glu** | 7.912 | 7.893 | 7.868 | 7.857 | 7.851 | -2.19 | 0.40 |
| **18** | **Val** | 8.123 | 8.092 | 8.070 | 8.055 | 8.018 | -3.52 | 0.10 |
| **19** | **Leu** | 8.037 | 8.018 | 7.993 | 7.984 | 7.962 | -2.59 | 0.23 |
| **20** | **Asn** | 7.554 | 7.540 | 7.513 | 7.511 | 7.497 | -2.00 | 0.34 |
| **22** | **Gln** | 8.957 | 8.892 | 8.822 | 8.803 | | -8.10 | 1.04 |
| **23** | **Tyr** | 8.541 | 8.509 | 8.488 | 8.469 | 8.431 | -3.70 | 0.10 |
| **24** | **Lys** | 7.353 | 7.337 | 7.326 | 7.320 | 7.302 | -1.70 | 0.05 |
| **25** | **Gly** | 9.333 | 9.266 | 9.234 | 9.207 | | -6.35 | 0.20 |
| **26** | **Lys** | 7.963 | 7.937 | 7.921 | | | -2.77 | 0.15 |
| **28** | **Leu** | 8.141 | 8.122 | 8.096 | 8.094 | 8.069 | -2.44 | 0.26 |
| **29** | **Val** | 9.004 | 8.976 | 8.957 | 8.943 | 8.912 | -3.09 | 0.07 |
| **30** | **Val** | 8.683 | 8.634 | 8.602 | 8.578 | 8.525 | -5.30 | 0.10 |
| **31** | **Ser** | 9.438 | 9.421 | 9.415 | 9.411 | 9.393 | -1.45 | 0.09 |
| **32** | **Val** | 8.928 | 8.906 | 8.894 | 8.883 | 8.854 | -2.45 | 0.10 |
| **42** | **Ala** | 8.686 | 8.667 | 8.667 | 8.652 | 8.629 | -1.86 | 0.21 |
| **43** | **Val** | 8.738 | 8.692 | 8.664 | 8.637 | 8.606 | -4.51 | 0.29 |
| **44** | **Ala** | 9.949 | 9.887 | 9.866 | 9.834 | 9.783 | -5.51 | 0.16 |
| **45** | **Thr** | 7.728 | 7.722 | 7.709 | 7.706 | 7.695 | -1.15 | 0.14 |
| **47** | **Asn** | 8.360 | 8.328 | 8.313 | 8.307 | 8.293 | -2.22 | 0.31 |
| **49** | **Glu** | 8.826 | 8.701 | 8.642 | 8.583 | 8.467 | -11.95 | 0.13 |
| **50** | **Ala** | 7.761 | 7.725 | 7.708 | 7.692 | 7.660 | -3.36 | 0.06 |
| **51** | **Arg** | 8.619 | 8.581 | 8.557 | 8.540 | 8.499 | -4.01 | 0.08 |
| **52** | **Lys** | 8.046 | 8.028 | 8.009 | 7.992 | 7.953 | -3.15 | 0.30 |
| **53** | **Leu** | 6.933 | 6.938 | 6.937 | 6.940 | 6.939 | 0.20 | 0.08 |
| **54** | **Gly** | 7.599 | 7.588 | 7.574 | 7.565 | 7.542 | -1.94 | 0.17 |
| **55** | **Val** | 7.315 | 7.278 | 7.264 | 7.259 | 7.243 | -2.35 | 0.36 |
| **56** | **Lys** | 7.044 | 7.089 | 7.081 | 7.083 | 7.073 | 0.81 | 0.79 |
| **58** | **Gly** | 9.005 | 8.964 | 8.961 | 8.944 | 8.915 | -2.90 | 0.25 |
| **59** | **Met** | 7.625 | 7.574 | 7.561 | 7.544 | 7.515 | -3.60 | 0.32 |
| **61** | **Ile** | 6.826 | 6.813 | 6.809 | 6.802 | 6.793 | -1.10 | 0.05 |
| **62** | **Ile** | 8.357 | 8.317 | 8.259 | 8.226 | 8.154 | -7.00 | 0.59 |
| **63** | **Lys** | 6.712 | 6.716 | 6.712 | 6.716 | 6.715 | 0.09 | 0.09 |
| **64** | **Ala** | 7.992 | 7.962 | 7.944 | 7.930 | 7.897 | -3.17 | 0.05 |
| **65** | **Met** | 8.409 | 8.377 | 8.357 | 8.345 | 8.311 | -3.26 | 0.09 |
| **66** | **Gln** | 7.238 | 7.217 | 7.206 | 7.201 | 7.185 | -1.75 | 0.12 |
| **67** | **Ile** | 7.365 | 7.335 | 7.323 | 7.316 | 7.200 | -5.13 | 1.33 |
| **68** | **Ala** | 8.699 | 8.659 | 8.637 | 8.621 | 8.575 | -4.10 | 0.10 |
| **71** | **Ala** | 7.039 | 7.022 | 7.009 | 6.998 | 6.976 | -2.13 | 0.08 |
| **72** | **Ile** | 8.040 | 8.025 | 8.018 | 8.004 | 7.997 | -1.50 | 0.16 |
| **73** | **Tyr** | 8.386 | 8.308 | 8.271 | 8.238 | 8.167 | -7.27 | 0.14 |
| **74** | **Val** | 8.892 | 8.857 | 8.836 | 8.817 | 8.794 | -3.34 | 0.22 |
| **76** | **Met** | 8.239 | 8.205 | 8.174 | 8.156 | 8.110 | -4.36 | 0.20 |
| **77** | **Arg** | 7.403 | 7.396 | 7.386 | 7.386 | 7.378 | -0.85 | 0.11 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **78** | **Lys** | 8.190 | 8.135 | 8.082 | 8.040 | 7.964 | -7.73 | 0.44 |
| **80** | **Ile** | 7.451 | 7.423 | 7.422 | 7.410 | 7.416 | -1.18 | 0.44 |
| **81** | **Tyr** | 6.880 | 6.867 | 6.860 | 6.855 | 6.845 | -1.17 | 0.06 |
| **82** | **Glu** | 8.961 | 8.915 | 8.889 | 8.855 | 8.790 | -5.73 | 0.32 |
| **83** | **Ala** | 7.687 | 7.664 | 7.654 | 7.647 | 7.635 | -1.73 | 0.17 |
| **84** | **Phe** | 8.235 | 8.225 | 8.218 | 8.215 | 8.201 | -1.12 | 0.06 |
| **85** | **Ser** | 8.911 | 8.885 | 8.868 | 8.858 | 8.829 | -2.73 | 0.07 |
| **86** | **Asn** | 8.812 | 8.757 | 8.731 | 8.701 | 8.640 | -5.72 | 0.11 |
| **87** | **Arg** | 7.722 | 7.697 | 7.683 | 7.671 | 7.641 | -2.69 | 0.07 |
| **88** | **Ile** | 8.071 | 8.045 | 8.032 | 8.008 | 7.987 | -2.89 | 0.20 |
| **89** | **Met** | 9.434 | 9.388 | 9.359 | 9.336 | 9.272 | -5.38 | 0.24 |
| **90** | **Asn** | 7.855 | 7.830 | 7.818 | 7.809 | 7.784 | -2.34 | 0.06 |
| **91** | **Leu** | 7.593 | 7.568 | 7.555 | 7.545 | 7.524 | -2.30 | 0.08 |
| **92** | **Leu** | 8.177 | 8.167 | 8.159 | 8.147 | 8.124 | -1.79 | 0.21 |
| **93** | **Asn** | 7.878 | 7.866 | 7.862 | 7.856 | 7.845 | -1.09 | 0.03 |
| **94** | **Lys** | 7.223 | 7.201 | 7.191 | 7.186 | 7.174 | -1.62 | 0.17 |
| **95** | **His** | 7.411 | 7.382 | 7.377 | 7.362 | 7.346 | -2.15 | 0.18 |
| **96** | **Ala** | 7.067 | 7.057 | 7.054 | 7.051 | 7.044 | -0.75 | 0.05 |
| **98** | **Lys** | 7.535 | 7.524 | 7.519 | 7.512 | 7.498 | -1.23 | 0.06 |
| **99** | **Ile** | 8.635 | 8.555 | 8.525 | 8.506 | | -6.59 | 0.68 |
| **100** | **Glu** | 9.353 | 9.312 | 9.296 | 9.276 | 9.244 | -3.63 | 0.12 |
| **103** | **Ser** | 8.510 | 8.482 | 8.464 | 8.458 | 8.433 | -2.55 | 0.16 |
| **104** | **Ile** | 7.939 | 7.907 | 7.898 | 7.890 | 7.861 | -2.51 | 0.16 |
| **105** | **Asp** | 8.007 | 8.028 | 8.027 | 8.039 | 8.048 | 1.34 | 0.18 |
| **106** | **Glu** | 6.432 | 6.421 | 6.433 | 6.421 | 6.438 | 0.18 | 0.38 |
| **107** | **Ala** | 8.198 | 8.179 | 8.134 | 8.127 | 8.106 | -3.28 | 0.57 |
| **108** | **Tyr** | 9.074 | 9.021 | 9.009 | 8.988 | 8.951 | -4.02 | 0.25 |
| **109** | **Leu** | 10.111 | 10.075 | 10.060 | 10.050 | 10.020 | -2.99 | 0.15 |
| **110** | **Asp** | 8.673 | 8.652 | 8.647 | 8.632 | 8.602 | -2.33 | 0.21 |
| **114** | **Lys** | 9.218 | 9.153 | 9.131 | 9.086 | 9.003 | -7.12 | 0.44 |
| **115** | **Val** | 7.629 | 7.606 | 7.592 | 7.580 | 7.547 | -2.72 | 0.14 |
| **116** | **Glu** | 8.194 | 8.162 | 8.149 | 8.138 | 8.101 | -3.03 | 0.14 |
| **117** | **Gly** | 8.798 | 8.745 | 8.718 | 8.692 | 8.636 | -5.39 | 0.04 |
| **118** | **Asn** | 7.569 | 7.553 | 7.542 | 7.530 | 7.517 | -1.79 | 0.11 |
| **120** | **Glu** | 8.153 | 8.140 | 8.131 | 8.131 | 8.126 | -0.90 | 0.18 |
| **121** | **Asn** | 8.022 | 7.962 | 7.925 | 7.916 | 7.852 | -5.56 | 0.36 |
| **122** | **Gly** | 8.286 | 8.252 | 8.233 | 8.221 | 8.189 | -3.22 | 0.10 |
| **123** | **Ile** | 7.784 | 7.758 | 7.745 | 7.733 | 7.704 | -2.65 | 0.05 |
| **124** | **Glu** | 7.403 | 7.375 | 7.361 | 7.334 | 7.306 | -3.32 | 0.21 |
| **125** | **Leu** | 8.006 | 8.002 | 7.982 | - | - | -1.43 | 0.89 |
| **126** | **Ala** | 7.979 | 7.978 | 7.971 | 7.968 | 7.952 | -0.91 | 0.20 |
| **127** | **Arg** | 7.922 | 7.903 | 7.894 | 7.890 | 7.872 | -1.63 | 0.09 |
| **128** | **Lys** | 8.034 | 8.012 | 8.001 | 7.998 | 7.980 | -1.76 | 0.14 |
| **129** | **Ile** | 8.653 | 8.624 | 8.604 | 8.589 | 8.553 | -3.35 | 0.10 |
| **130** | **Lys** | 7.800 | 7.766 | 7.748 | 7.730 | 7.695 | -3.51 | 0.02 |
| **131** | **Gln** | 7.863 | 7.862 | 7.851 | 7.853 | 7.854 | -0.36 | 0.19 |
| **132** | **Glu** | 8.590 | 8.583 | 8.581 | 8.583 | 8.577 | -0.39 | 0.09 |
| **133** | **Ile** | 8.163 | 8.130 | 8.120 | 8.108 | 8.078 | -2.77 | 0.12 |
| **134** | **Leu** | 7.475 | 7.461 | 7.451 | 7.448 | 7.434 | -1.36 | 0.08 |
| **135** | **Glu** | 8.650 | 8.619 | 8.604 | 8.593 | 8.565 | -2.81 | 0.08 |
| **136** | **Lys** | 8.606 | 8.566 | 8.544 | 8.528 | 8.487 | -3.95 | 0.07 |
| **137** | **Glu** | 8.060 | 8.044 | 8.032 | 8.026 | 8.004 | -1.86 | 0.09 |
| **138** | **Lys** | 7.597 | 7.586 | 7.580 | 7.581 | 7.576 | -0.68 | 0.14 |
| **139** | **Ile** | 6.578 | 6.575 | 6.572 | 6.578 | 6.581 | 0.12 | 0.16 |
| **140** | **Thr** | 10.130 | 10.070 | 10.031 | 9.989 | 9.893 | -7.92 | 0.53 |
| **141** | **Val** | 7.341 | 7.328 | 7.320 | 7.319 | 7.309 | -1.05 | 0.11 |
| **142** | **Thr** | 9.032 | 9.006 | 8.987 | 8.973 | 8.934 | -3.27 | 0.17 |
| **143** | **Val** | 7.843 | 7.825 | 7.819 | 7.807 | 7.791 | -1.74 | 0.06 |
| **144** | **Gly** | 9.299 | 9.302 | 9.287 | 9.298 | 9.283 | -0.52 | 0.31 |
| **145** | **Val** | 8.791 | 8.752 | 8.738 | 8.728 | 8.704 | -2.85 | 0.25 |
| **146** | **Ala** | 9.231 | 9.208 | 9.192 | 9.173 | 9.139 | -3.11 | 0.19 |
| **149** | **Lys** | 7.389 | 7.384 | 7.374 | 7.362 | 7.359 | -1.12 | 0.21 |
| **150** | **Ile** | 7.448 | 7.423 | 7.425 | 7.410 | 7.416 | -1.09 | 0.40 |
| **151** | **Leu** | 8.061 | 8.036 | 8.015 | 8.002 | 7.980 | -2.77 | 0.17 |
| **152** | **Ala** | 7.602 | 7.571 | 7.554 | 7.547 | 7.525 | -2.55 | 0.21 |
| **153** | **Lys** | 7.354 | 7.332 | 7.289 | 7.261 | 7.240 | -4.13 | 0.60 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **154** | **Ile** | 7.993 | 7.938 | 7.938 | | | -1.83 | 1.06 |
| **155** | **Ile** | 8.367 | 8.357 | 8.342 | 8.327 | 8.314 | -1.89 | 0.22 |
| **156** | **Ala** | 7.784 | 7.758 | 7.744 | 7.732 | 7.705 | -2.63 | 0.02 |
| **157** | **Asp** | 8.360 | 8.328 | 8.313 | 8.307 | 8.293 | -2.22 | 0.31 |
| **159** | **Ser** | 7.843 | 7.823 | 7.811 | 7.800 | 7.774 | -2.30 | 0.08 |
| **160** | **Lys** | 7.080 | 7.055 | 7.036 | 7.033 | 7.009 | -2.35 | 0.19 |
| **162** | **Asn** | 9.974 | 9.957 | 9.936 | 9.932 | 9.896 | -2.59 | 0.26 |
| **163** | **Gly** | 8.209 | 8.201 | 8.224 | 8.215 | | 0.31 | 0.50 |
| **165** | **Gly** | 9.320 | 9.284 | 9.269 | 9.248 | 9.209 | -3.69 | 0.10 |
| **166** | **Val** | 8.503 | 8.432 | 8.395 | 8.361 | 8.302 | -6.74 | 0.20 |
| **167** | **Ile** | 8.638 | 8.600 | 8.565 | 8.543 | 8.509 | -4.44 | 0.30 |
| **168** | **Arg** | 8.457 | 8.462 | 8.454 | | | -0.10 | 0.52 |
| **171** | **Glu** | 7.843 | 7.810 | 7.799 | 7.778 | 7.745 | -3.26 | 0.12 |
| **172** | **Val** | 6.946 | 6.935 | 6.929 | 6.928 | 6.923 | -0.76 | 0.12 |
| **174** | **Asp** | 8.277 | 8.170 | 8.126 | | | -10.16 | 0.47 |
| **175** | **Phe** | 7.974 | 7.956 | 7.942 | 7.915 | | -2.81 | 0.56 |
| **176** | **Leu** | 8.810 | 8.753 | 8.728 | 8.690 | 8.615 | -6.48 | 0.35 |
| **177** | **Asn** | 7.940 | 7.898 | 7.876 | 7.859 | 7.821 | -3.96 | 0.10 |
| **178** | **Glu** | 7.050 | 7.049 | 7.048 | 7.045 | 7.042 | -0.28 | 0.05 |
| **179** | **Leu** | 7.062 | 7.038 | 7.029 | 7.030 | 7.033 | -0.95 | 0.44 |
| **180** | **Asp** | 9.131 | 9.060 | 9.030 | 8.986 | 8.902 | -7.61 | 0.29 |
| **181** | **Ile** | 8.225 | 8.193 | 8.180 | 8.169 | 8.145 | -2.64 | 0.13 |
| **182** | **Asp** | 8.281 | 8.279 | 8.263 | 8.274 | 8.270 | -0.38 | 0.30 |
| **183** | **Glu** | 7.551 | 7.535 | 7.524 | 7.527 | | -1.33 | 0.34 |
| **184** | **Ile** | 7.284 | 7.260 | 7.243 | 7.239 | 7.224 | -2.01 | 0.23 |
| **188** | **Gly** | 8.377 | 8.310 | 8.285 | | | -6.21 | 0.42 |
| **191** | **Leu** | 8.209 | 8.132 | 8.084 | 8.030 | 7.977 | -7.98 | 0.53 |
| **192** | **Ala** | 8.511 | 8.460 | 8.430 | 8.394 | 8.329 | -6.12 | 0.25 |
| **193** | **Arg** | 7.831 | 7.811 | 7.796 | 7.788 | 7.764 | -2.24 | 0.08 |
| **194** | **Arg** | 7.750 | 7.718 | 7.707 | 7.699 | 7.668 | -2.65 | 0.15 |
| **195** | **Leu** | 8.428 | 8.395 | 8.375 | 8.342 | 8.304 | -4.25 | 0.28 |
| **196** | **Asn** | 8.740 | 8.711 | 8.687 | 8.655 | 8.616 | -4.28 | 0.32 |
| **197** | **Glu** | 8.039 | 8.015 | 7.980 | | | -3.71 | 1.14 |
| **198** | **Leu** | 7.365 | 7.352 | 7.344 | 7.343 | 7.333 | -1.05 | 0.11 |
| **199** | **Gly** | 7.900 | 7.888 | 7.877 | 7.872 | 7.856 | -1.48 | 0.07 |
| **200** | **Ile** | 7.914 | 7.868 | 7.847 | 7.828 | 7.797 | -3.91 | 0.22 |
| **201** | **Gln** | 9.016 | 9.008 | 9.008 | 8.998 | 8.983 | -1.09 | 0.18 |
| **202** | **Lys** | 8.570 | 8.554 | 8.541 | 8.531 | 8.509 | -2.06 | 0.09 |
| **203** | **Leu** | 8.541 | 8.519 | 8.509 | 8.506 | 8.491 | -1.63 | 0.16 |
| **204** | **Arg** | 8.470 | 8.419 | 8.386 | 8.342 | 8.263 | -6.98 | 0.48 |
| **205** | **Asp** | 7.962 | 7.929 | 7.918 | 7.908 | 7.893 | -2.28 | 0.26 |
| **212** | **Asn** | 7.937 | 7.907 | 7.898 | 7.890 | 7.861 | -2.45 | 0.15 |
| **207** | **Leu** | 7.187 | 7.169 | 7.158 | 7.151 | 7.139 | -1.62 | 0.12 |
| **213** | **Glu** | 7.684 | 7.623 | 7.603 | 7.571 | | -5.56 | 0.27 |
| **214** | **Leu** | 7.553 | 7.527 | 7.491 | 7.470 | 7.439 | -3.99 | 0.37 |
| **215** | **Glu** | 8.374 | 8.341 | 8.315 | 8.275 | 8.244 | -4.56 | 0.42 |
| **216** | **Lys** | 7.298 | 7.266 | 7.250 | 7.241 | 7.236 | -2.11 | 0.43 |
| **217** | **Ile** | 7.442 | 7.440 | 7.437 | 7.438 | 7.435 | -0.23 | 0.04 |
| **218** | **Thr** | 8.421 | 8.390 | 8.368 | 8.354 | 8.318 | -3.45 | 0.10 |
| **223** | **Ala** | 7.629 | 7.638 | 7.638 | 7.644 | 7.638 | 0.33 | 0.20 |
| **224** | **Leu** | 7.881 | 7.898 | 7.897 | 7.897 | 7.898 | 0.50 | 0.25 |
| **225** | **Tyr** | 7.587 | 7.567 | 7.560 | 7.548 | 7.530 | -1.90 | 0.05 |
| **226** | **Leu** | 8.091 | 8.051 | 8.034 | 8.015 | 7.972 | -3.93 | 0.08 |
| **228** | **Lys** | 8.357 | 8.332 | 8.323 | 8.311 | 8.291 | -2.19 | 0.07 |
| **229** | **Leu** | 8.177 | 8.138 | 8.072 | | | -3.50 | 0.52 |
| **230** | **Ala** | 7.600 | 7.571 | 7.555 | 7.547 | 7.525 | -2.49 | 0.16 |
| **231** | **Gln** | 7.697 | 7.653 | 7.630 | 7.631 | 7.616 | -2.65 | 0.59 |
| **252** | **Leu** | 8.028 | 7.996 | 7.978 | 7.967 | 7.942 | -2.87 | 0.15 |
| **254** | **Tyr** | 6.767 | 6.722 | 6.719 | 6.715 | 6.698 | -2.14 | 0.48 |
| **256** | **Thr** | 8.380 | 8.334 | 8.317 | 8.297 | 8.263 | -3.88 | 0.16 |
| **258** | **Asp** | 8.729 | 8.668 | 8.666 | 8.625 | 8.599 | -4.33 | 0.52 |
| **259** | **Val** | 8.764 | 8.667 | 8.631 | | | -8.99 | 0.62 |
| **260** | **Lys** | 8.326 | 8.314 | 8.310 | 8.300 | 8.281 | -1.49 | 0.14 |
| **261** | **Val** | 7.258 | 7.257 | 7.254 | 7.247 | 7.230 | -0.94 | 0.25 |
| **262** | **Ile** | 7.684 | 7.652 | 7.633 | 7.623 | 7.594 | -2.99 | 0.13 |
| **265** | **Tyr** | 7.380 | 7.351 | 7.341 | 7.336 | 7.325 | -1.80 | 0.29 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **266** | **Leu** | 8.053 | 8.061 | 8.060 | 8.061 | | 0.27 | 0.11 |
| **267** | **Lys** | 8.745 | 8.706 | 8.682 | 8.666 | 8.624 | -4.03 | 0.08 |
| **268** | **Lys** | 7.412 | 7.408 | 7.394 | 7.400 | 7.390 | -0.74 | 0.21 |
| **269** | **Ala** | 8.248 | 8.212 | 8.191 | 8.183 | 8.155 | -3.08 | 0.22 |
| **270** | **Ile** | 8.397 | 8.344 | 8.324 | 8.310 | 8.302 | -3.19 | 0.67 |
| **271** | **Asn** | 8.727 | 8.687 | 8.649 | 8.625 | 8.587 | -4.82 | 0.32 |
| **272** | **Glu** | 8.360 | 8.350 | 8.341 | 8.333 | 8.308 | -1.73 | 0.21 |
| **273** | **Ala** | 8.395 | 8.378 | 8.375 | 8.370 | 8.367 | -0.92 | 0.19 |
| **274** | **Tyr** | 8.966 | 8.917 | 8.891 | 8.866 | 8.806 | -5.31 | 0.16 |
| **275** | **Asn** | 7.737 | 7.720 | 7.710 | 7.702 | 7.682 | -1.83 | 0.04 |
| **276** | **Lys** | 7.389 | 7.384 | 7.367 | 7.362 | 7.359 | -1.12 | 0.26 |
| **277** | **Val** | 7.286 | 7.272 | 7.261 | 7.260 | 7.252 | -1.14 | 0.17 |
| **279** | **Gly** | 7.286 | 7.275 | 7.263 | 7.262 | 7.247 | -1.30 | 0.12 |
| **280** | **Ile** | 8.233 | 8.228 | 8.221 | 8.215 | 8.187 | -1.51 | 0.32 |
| **282** | **Met** | 8.076 | 8.067 | 8.053 | 8.053 | 8.041 | -1.19 | 0.15 |
| **283** | **Arg** | 8.023 | 7.992 | 7.992 | 7.977 | 7.966 | -1.86 | 0.27 |
| **284** | **Ile** | 8.182 | 8.163 | 8.155 | 8.152 | 8.139 | -1.40 | 0.13 |
| **285** | **Thr** | 8.381 | 8.378 | 8.373 | 8.373 | 8.370 | -0.38 | 0.06 |
| **286** | **Val** | 8.798 | 8.771 | 8.761 | 8.753 | 8.734 | -2.10 | 0.14 |
| **287** | **Ile** | 8.560 | 8.539 | 8.529 | 8.524 | 8.511 | -1.62 | 0.15 |
| **288** | **Ala** | 9.388 | 9.357 | 9.340 | 9.328 | 9.298 | -2.99 | 0.07 |
| **289** | **Ile** | 8.608 | 8.611 | 8.614 | | | 0.39 | 0.07 |
| **290** | **Met** | 9.028 | 9.020 | 9.010 | 9.001 | 8.972 | -1.87 | 0.29 |
| **291** | **Glu** | 8.463 | 8.451 | 8.419 | 8.409 | 8.390 | -2.61 | 0.38 |
| **293** | **Leu** | 8.287 | 8.258 | 8.253 | 8.249 | 8.231 | -1.77 | 0.23 |
| **294** | **Asp** | 7.654 | 7.620 | 7.599 | 7.588 | 7.561 | -3.11 | 0.18 |
| **295** | **Ile** | 8.025 | 7.997 | 7.981 | 7.954 | 7.926 | -3.40 | 0.21 |
| **296** | **Leu** | 9.171 | 9.133 | 9.113 | 9.098 | 9.063 | -3.59 | 0.09 |
| **297** | **Ser** | 8.386 | 8.337 | 8.302 | 8.288 | 8.281 | -3.64 | 0.77 |
| **298** | **Lys** | 8.559 | 8.527 | 8.511 | 8.503 | 8.483 | -2.52 | 0.23 |
| **300** | **Lys** | 8.596 | 8.544 | 8.532 | 8.521 | 8.494 | -3.29 | 0.41 |
| **301** | **Lys** | 8.123 | 8.081 | 8.057 | | | -4.37 | 0.15 |
| **302** | **Phe** | 9.179 | 9.126 | 9.104 | 9.078 | 9.025 | -5.10 | 0.07 |
| **306** | **Ile** | 8.457 | 8.421 | 8.373 | 8.357 | 8.335 | -4.30 | 0.63 |
| **307** | **Ser** | 8.823 | 8.748 | 8.704 | | | -7.87 | 0.32 |
| **308** | **Ile** | 8.740 | 8.684 | 8.632 | 8.610 | 8.592 | -5.18 | 0.86 |
| **309** | **Asp** | 7.521 | 7.524 | 7.524 | 7.528 | 7.544 | 0.73 | 0.22 |
| **310** | **Asn** | 7.312 | 7.289 | 7.279 | 7.273 | 7.259 | -1.75 | 0.15 |
| **311** | **Ala** | 8.632 | 8.588 | 8.562 | 8.544 | 8.501 | -4.37 | 0.10 |
| **312** | **Tyr** | 7.518 | 7.514 | 7.509 | 7.513 | 7.506 | -0.37 | 0.11 |
| **313** | **Lys** | 7.194 | 7.172 | 7.159 | 7.151 | 7.132 | -2.07 | 0.08 |
| **314** | **Val** | 8.351 | 8.331 | 8.313 | 8.300 | 8.266 | -2.86 | 0.19 |
| **315** | **Ala** | 8.902 | 8.858 | 8.822 | 8.807 | 8.746 | -5.19 | 0.26 |
| **316** | **Glu** | 7.627 | 7.604 | 7.592 | 7.579 | 7.561 | -2.23 | 0.08 |
| **317** | **Asp** | 7.936 | 7.909 | 7.888 | 7.871 | 7.831 | -3.53 | 0.18 |
| **318** | **Leu** | 8.418 | 8.417 | 8.395 | 8.398 | 8.384 | -1.21 | 0.30 |
| **319** | **Leu** | 8.429 | 8.381 | 8.352 | 8.323 | 8.284 | -4.93 | 0.21 |
| **320** | **Arg** | 8.481 | 8.450 | 8.429 | 8.416 | 8.383 | -3.28 | 0.09 |
| **321** | **Glu** | 8.018 | 7.977 | 7.956 | 7.942 | 7.910 | -3.59 | 0.18 |
| **322** | **Leu** | 8.348 | 8.298 | 8.270 | 8.248 | 8.200 | -4.94 | 0.09 |
| **323** | **Leu** | 7.639 | 7.638 | 7.635 | 7.641 | 7.642 | 0.12 | 0.12 |
| **324** | **Val** | 7.537 | 7.546 | 7.550 | 7.556 | 7.547 | 0.40 | 0.27 |
| **325** | **Arg** | 7.507 | 7.494 | 7.488 | 7.479 | 7.471 | -1.23 | 0.08 |
| **326** | **Asp** | 7.915 | 7.853 | 7.824 | 7.797 | 7.744 | -5.69 | 0.14 |
| **328** | **Arg** | 8.358 | 8.336 | 8.320 | 8.304 | 8.270 | -2.96 | 0.18 |
| **329** | **Arg** | 8.550 | 8.538 | 8.527 | 8.516 | 8.486 | -2.14 | 0.25 |
| **330** | **Asn** | 9.010 | 8.982 | 8.966 | 8.955 | 8.926 | -2.79 | 0.06 |
| **331** | **Val** | 8.487 | 8.448 | 8.432 | 8.415 | 8.380 | -3.54 | 0.07 |
| **332** | **Arg** | 9.216 | 9.192 | 9.162 | 9.144 | 9.100 | -3.96 | 0.31 |
| **333** | **Arg** | 8.383 | 8.357 | 8.339 | 8.326 | 8.297 | -2.89 | 0.07 |
| **334** | **Ile** | 8.969 | 8.933 | 8.916 | 8.899 | 8.858 | -3.67 | 0.09 |
| **335** | **Gly** | 9.290 | 9.265 | 9.250 | 9.243 | 9.222 | -2.26 | 0.13 |
| **336** | **Val** | 8.309 | 8.230 | 8.188 | 8.167 | 8.118 | -6.36 | 0.57 |
| **337** | **Lys** | 8.724 | 8.697 | 8.684 | 8.680 | 8.660 | -2.09 | 0.19 |
| **339** | **Asp** | 8.549 | 8.518 | 8.496 | 8.479 | 8.446 | -3.48 | 0.10 |