INVESTIGATING COLD ADAPTATION OF INTRINSICALLY DISORDERED PROTEINS IN P53 HOMOLOGS

by

Leona A. Martin, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Biochemistry
December 2016

Committee Members:

Steven T. Whitten, Chair

Rachell E. Booth

Karen A. Lewis

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

DEDICATION

To my husband, Blake E. Martin. Without your endless support, this would not have been possible.

ACKNOWLEDGEMENTS

I would like to thank my graduate advisor, Dr. Steven T. Whitten, for providing me with an excellent opportunity to further my education. I would like to thank my committee members, Dr. Rachell E. Booth and Dr. Karen A. Lewis, for always assisting me with my efforts and troubleshooting. A special thank you to Dr. Ronald Walter and Dr. John Postlethwait, who provided the DNA sequences. I would also like to thank my lab mates, particularly Lance R. English and Erin C. Tilton.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

I. INTRODUCTION

1.1 Intrinsically Disordered Proteins

Intrinsically disordered proteins (IDPs) are proteins, or proteins with regions of at least 30 amino acids, that lack stable tertiary structure under native, biological conditions. This lack of stable tertiary structure gives IDPs structural flexibility and increased conformational freedom[1,2]. IDPs are found to be more common in eukaryotes than prokaryotes with approximately 33 % of all eukaryotic proteins being disordered[3,4]. The importance of IDPs is expressed by the high incidence of IDPs in cancer-associated proteins, 79 % of which are classified as IDPs[5]. Many IDPs are transcription factors or cell-signaling proteins and often are found to be related to diseases[6]. More evolutionary change in amino acid sequences over time is observed than with stable, globular proteins, though disorder is highly conserved, indicating that disorder is indeed important for protein function[7–9]. Perhaps because these proteins lack stable tertiary structure, there is less need to conserve the sequence as long as disorder is conserved[10].

As efforts have been made to better elucidate IDPs, patterns characteristic of IDPs have been discovered that help predict a protein's propensity for disorder. Regions with little diversity in amino acid composition are considered to have low sequence complexity, which is associated with disorder; a low number of hydrophobic residues, as well as a high net charge have also been shown to positively correlate with disorder[11,12]. Hydrophobic amino acids such as tryptophan, phenylalanine, isoleucine, and leucine are the least frequent amino acids found in proteins classified as IDPs. Meanwhile, alanine, arginine, glycine, glutamine, serine, proline, glutamic acid, and lysine are disorder-

promoting amino acids and are in greater abundance in IDPs[11]. These amino acid compositions for IDPs result in low chain hydrophobicity and is thought to lead to the inability of the disordered protein to form a hydrophobic core, which is a strong driving force for protein folding into stable, globular structures[10]. Net electrostatic repulsions resulting from high net charge may also promote disorder in IDPs by preventing a collapse into stable and folded tertiary structures[12]. High net charge and electrostatic repulsions may lead to a decrease in conformational freedom, or increased rigidity, when compared to a random coil[13].

Though IDPs are more rigid than a random coil, IDPs can often retain more flexibility than a globular protein, which may be important in biology. This flexibility in IDPs, which are often transcription factors and cell-signaling proteins, may allow them to be more capable to handle responses for a diverse set of environmental conditions[3]. A trademark of IDPs is also that they are more extended, with a larger hydrodynamic radius (R_h), than globular proteins[14]. These traits of IDPs may allow for promiscuous binding, defined as being capable of binding multiple different ligands with high affinity[15]. Many disordered proteins undergo a local conformational change from disordered to ordered upon binding of target ligands[9,15].

Though IDPs lack stable tertiary structures, a common structural motif in disordered and unfolded proteins has been identified as a poly-L-proline type II (PPII) helix[16]. PPII helices are different from α-helices in that they have a left-handed turn with 3 residues/turn and a 3.1 Å unit height[17]. Whereas α-helices have a right-handed turn with 3.6 residues per turn and a 1.5 Å unit height[18]. As the name indicates, proline residues are common in PPII structure, though they are not necessary. Poly-lysine and poly-alanine

peptides have been shown to have strong propensities for PPII structure. Both glutamine and arginine have a higher prevalence in PPII structures over other secondary structures[16,17,19]. Amino acids with small side chains as well as amino acids with flexible side chains tend to be present in PPII structures, allowing the peptide backbone to be better solvated[19]. A loss of PPII is observed with increasing temperature, when measured by circular dichroism (CD) and NMR spectroscopy[16,20]. Consistent with the idea that PPII is common in IDP structures, increases in temperature cause the hydrodynamic size of IDPs to decrease and spectral signals for PPII also decrease[13,21].

Perhaps due to the extended, solvent-exposed structures of IDPs, post-translational modifications are found frequently on IDPs. Post-translational modifications are especially important for cell-signaling proteins, of which many are IDPs[5]. Phosphorylation, an important regulatory mechanism in cells, occurs more frequently in IDPs and on residues near disorder-promoting amino acids[10,22]. IDPs may be more susceptible to degradation and shorter half-lives due to their exposed cleavage sites; post-translational modifications can both increase the half-life of the protein as well as target it for degradation through ubiquitin-mediated degradation[10,23].

1.2 The Importance of p53

One example of an IDP is the transcription factor p53, a tetrameric protein with large regions of disorder[24]. p53 is linked to many types of cancer as a transcriptional regulator and tumor suppressor, and mutations in the p53 gene are common in many cancers[25]. As a transcription factor, the p53 protein binds specific DNA sequences in target genes to increase transcription of regulatory proteins involved in cell cycle arrest

and apoptosis[25]. It has also been shown that p53 has an important role in cell differentiation and in development[25].

The p53 protein has four main domains: the transactivational domain, the proline-rich domain, DNA-binding domain, and tetramerization domain. The transactivational domain (TAD) is located at the N-terminus, residues 1 – 63[26]. The TAD consists of activation domain 1 (AD1), residues 1 – 42, and activation domain 2 (AD2), residues 43 – 63. These residues tend to be highly acidic with a net negative charge equal to -15[26]. The TAD is primarily disordered when not bound to regulatory proteins, as determined by NMR[27,28]. Upon the binding of the regulatory protein MDM2, a small segment of residues of the p53 TAD become ordered, adopting an α-helical conformation[29]. MDM2 is an antagonist regulator of p53 that binds residues 17 – 27 of p53 TAD[30]. MDM2 binds p53 tightly with high specificity allowing for efficient activity as an inhibitor to the transcriptional activity of p53[25,30]. The binding of MDM2 forms a negative feedback loop with p53. As p53 is increased, transcription of MDM2 increases, which leads to decreased p53 through targeted degradation, then leading to decreased MDM2[31]. Degradation of p53 is increased by MDM2 by triggering ubiquination, leading p53 to the ubiquitin-proteosome pathway[25,32].

The proline-rich domain is located at residues 64 – 93 and has high proline content with five repeats of a PXXP motif[33]. This PXXP motif is conserved across multiple vertebral species[33,34]. The proline-rich domain is thought to have substantial PPII structure, owing to twelve proline and ten alanine residues, that cause chain rigidity and helps to orient the N-terminus away from the core of the p53 tetramer[35]. Both the proline-rich domain and the TAD are linked to cell growth control and suppression, aiding in

tumor suppression[33]. The proline-rich region may not be required for cell cycle arrest, however, it is needed for cell apoptosis[36].

While ubiquitination targets p53 for degradation, other post-translational modifications help stabilize p53, resulting in increased longevity[32]. N-terminal phosphorylation activates p53 as a tumor suppressor and helps prevent ubiquination[25,32]. Phosphorylation of p53 also alters the selectivity of binding to promoters[37]. Within the TAD, there are ten potential phosphorylation sites, including many of the serine and threonine residues[31,34]. Ser-15, Ser-33, and Ser-37 are phosphorylated upon exposure to DNA damaging stimuli such as UV light and ionizing radiation[38,39]. These phosphorylation events induce C-terminal acetylation, increasing transcriptional activity. Phosphorylation events on the N-terminus of p53 prevent the binding of MDM2 to the TAD, inhibiting the negative feedback loop[40]. Without MDM2 binding, there is increased transcriptional activity of p53, increasing the concentrations of p53 in the cell, which allows regulatory events such as apoptosis to occur as a result of damaging stimuli.

1.3 Methods to Measure Structure

There are many methods utilized to study protein structure, however, many of them are not well-suited for use with IDPs. X-ray crystallography is perhaps the most common method for studying the secondary and tertiary structures of globular proteins. Due to the high flexibility of IDPs, they do not crystallize well; when crystals do form, it is likely that the structure is altered from its natively unfolded state[35]. Also, when X-ray crystallography is performed on IDPs, the intrinsically disordered regions do not exhibit clear electron density regions because they do not scatter the X-rays diffusely due to the

varying conformational states[14]. With X-ray crystallography it is possible to identify potential regions of disorder due to the lack of electron density[41]. Even with this valuable technology, IDPs cannot be defined with certainty. Though X-ray crystallography is suitable for studying the folded domains of p53, the p53 N-terminal structure, when unbound, cannot be determined. X-ray crystallography is not a suitable method for studying the disordered structures associated with p53[14,42].

Using nuclear magnetic resonance spectroscopy (NMR), secondary structure can be elucidated for many proteins by examining proton-proton interactions. One of the advantages of this technique is that the protein is in solution, so it does not need to be crystallized[14]. Though NMR provides advantages for IDPs, and is often used to study IDPs, accurate measurements are difficult due to the conformational fluctuations in IDPs which create a varied conformational ensemble of multiple, fluid conformational states[34].

A beneficial tool in measuring secondary structure is circular dichroism spectroscopy (CD). As with NMR, the protein structure is determined in solution rather than from crystals[14]. Traditional secondary structures of proteins, such as α-helices and β-sheets have long been determined by CD[43]. In addition to more traditional secondary structures, PPII structure can also be determined by CD[44]. CD has been useful for studying the N-terminus of human p53 (residues 1 – 93). Human p53(1-93) exhibits a local maximum at 221 nm, which is a hallmark of PPII structure[44,45]. Temperature trends for human p53(1-93) are also consistent with those seen for PPII. The minima (approximately 200 nm) increases as temperatures increase, while the maxima (221 nm) decreases as temperatures increase[45,46]. There is also an isochromatic point observed for human p53(1-93) around 210 nm[45].

Another indicator of protein structure is by size or hydrodynamic radius (R_h) using Stokes radius determination. This can be measured by size exclusion chromatography (SEC). It has been observed with both SEC and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) that human p53(1-93) is larger than folded globular proteins of the same molecular weight[43]. R_h can also be determined using dynamic light scattering (DLS). An advantage to DLS is that it can easily be performed over a range of temperatures to monitor thermal effects on IDP structures such as human p53(1-93). In previous studies, the R_h of human p53(1-93) was shown to decrease as temperature increases, decreasing from 35 Å at 5 °C to 26 Å at 75 °C[21]. An overall decreased R_h was observed in human p53(1-93) when all of the prolines were mutated to glycine residues, indicating that proline residues are important for the increased R_h[47]. Though the R_h was decreased in the mutants, the thermal sensitivity of R_h remained[47].

1.4 Sequence Effects on the Structure of the p53 N-terminus

One way to study the effects that amino acid sequence have on IDPs, particularly the p53 N-terminus due to its high degree of disorder and PPII structure, is to turn to homologous proteins in differing species. Previously, multiple eukaryotic species' p53 N-termini were studied and found to vary in sequence while maintaining intrinsic disorder[48]. p53 homologues in different species maintain functionality at a wide variety of conditions, including the extreme temperatures of the arctic. With the noted strong temperature effects on human p53(1-93) size and structure, how will temperature effect the p53 homologues from cold-adapted species?

p53 homologs from icefish (*Chaenocephalus aceratus*) and stickleback (*Gasterosteus aculeatus*) were chosen, as these species are vertebrates that live in cold environments and do not regulate their physiological temperatures. Icefish is from Antarctica and lives at a constant -2 °C[49]. Stickleback is found in varying climates throughout the world. The subspecies in this study is from Alaska and lives at a temperature of 0 – 5 °C[50]. As a control, a homolog of p53 from Beluga whale (*Delphinapterus leucas*), which maintain a constant body temperature similar to humans but lives in cold temperatures, will be used. The amino acid sequences of both icefish p53(1-89) and stickleback p53(1-80) contain lower proline content and stickleback has a lower net charge than that of human p53(1-93). Whereas whale p53(1-86) has high homology with human p53(1-93) and contains a high proline content (Figure 1.1).

Figure 1.1 The Amino Acid Sequences for p53 N-terminus Homologs. Amino acid sequences were aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo).

Table 1.1 Comparison of p53 N-terminus homologs to human p53(1-93).

Protein	% identity	% homology
Whale p53(1-86)	67	77.8
Icefish p53(1-89)	48	21.2
Stickleback p53(1-80)	45	20.0

Each amino acid sequence was compared to human p53(1-93) using NCBI Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

The goal of this project is to investigate the differences in structural properties due to sequence variations in p53 homologs from species that live at different temperatures. Because IDP structures are generally temperature sensitive, our hypothesis is that structural adaptation to cold temperatures is needed to maintain disorder-associated properties[21,45]. Human, whale, icefish, and stickleback p53 N-terminus homologs will be compared. In pursuit of this goal, the hydrodynamic radii and CD spectrum of the four p53 N-terminus homologues will be measured. Based on previous studies, the reduced proline content of ice fish and stickleback p53 N-termini may result in a lower R_h. The effect that temperature has on p53 N-terminus R_h and secondary structure also correlates with proline content and net charge, indicating that there may be a lesser temperature effect on the structure of icefish and stickleback p53. PPII structure will be monitored over a range of temperatures using CD.

II. MATERIALS AND METHODS

2.1 Reagents and Equipment

All reagents used were of ACS grade or higher and were purchased from EMD Millipore (Billerica, MA), unless otherwise noted. An EMD Millipore Milli-Q Integral 3 water purification system was used to filter and deionize all water used for reagents and growth media. To weigh materials and reagents, an A&D GH-200 analytical balance (Tokyo, Japan) was used. The pH meter used to adjust the pH of solutions was a Beckman Coulter Φ 510 pH meter (Brea, CA). A Welch DryFast 2032 Ultra Diaphragm Pump (Niles, IL) was used for vacuum filtration and the degassing of solutions. A Thermo Fisher Scientific Sorvall LYNX 6000 Superspeed centrifuge (Waltham, MA) was used for all centrifugation steps. Dialysis tubing from Spectrum Labs, Spectra/Por tubing with a 12-14 kDa molecular weight cut-off, was used for all dialysis (Rancho Dominguez, CA).

A Hirayama HICLAVE HV-50 autoclave (Kasukabe-Shi, Japan) was used for all sterilization of medium and equipment. Incubators used for cell cultures were a VWR 120 V forced air microbiological incubator (Radnor, PA) and a Thermo Fisher Scientific MaxQ 5000 floor-model shaker (Waltham, MA). Cell lysis was performed with a Branson Sonifier S-450A (Danbury, CT). A Bio-Rad Biologic LP low pressure chromatography system (Hercules, CA) was used for all liquid chromatography: nickel affinity chromatography, ion-exchange chromatography, and size exclusion chromatography. A Jasco J-710 spectropolarimeter was used to perform circular dichroism experiments, along with a Jasco spectropolarimeter power supply and a Jasco

PFD-425S peltier (Easton, MD). Dynamic light scattering was performed using a Zetasizer Nano ZS with peltier temperature control (Malvern Instruments, Malvern, UK).

2.2 Plasmid and Transformation

Plasmid DNA was provided by DNA 2.0 (Menlo Park, CA). The plasmid contained the DNA sequence for one each of the p53 N-terminus homologs with a 6x-histidine tag at the N-terminus as well as a thrombin cleavage site for downstream removal of the 6x-His tag. A *lac* operon and T5 promoter preceded the target gene to control expression. The plasmid also contained the β-lactamase gene, providing ampicillin resistance.

To transform the plasmid DNA into competent cells, 10 ng plasmid DNA were added to 100 µL *Escherichia coli* BL21(DE3)pLysS calcium chloride competent cells from Novagen (EMD; Billerica, MA). The cells and DNA were incubated for 5 minutes on ice, for 0.5 minutes at 42 °C, followed by 2 minutes on ice. SOC medium, 250 µL, was added to the cell/DNA mixture and incubated for 30 minutes at 37 °C to allow for cell growth and β-lactamase expression, providing ampicillin resistance. A lysogeny broth (LB) agar plate with 100 µg/mL ampicillin (AMP) was inoculated with 100 µL of transformed cells and incubated overnight at 37 °C.

A single colony of transformed cells was used to inoculate 5 mL LB/AMP. The culture was incubated overnight at 30 °C with orbital rotation. To prepare glycerol stocks, sterile 50 % glycerol was combine with the overnight culture at a ratio of 1:1, for a final concentration of 25 % glycerol. Glycerol stocks were stored at -80 °C.

2.3 Expression and Purification of Recombinant p53 N-terminus Homologs

Glycerol stocks were used to inoculate an LB/AMP agar plate by streaking for colony isolation; the plate was then incubated overnight at 37 °C. A single colony was aseptically transferred to 20 mL of LB/AMP and incubated overnight at 30 °C with orbital rotation. The overnight culture was transferred to four flasks with 500 mL LB/AMP each for growth and expression. Cultures were grown to $O.D._{600}$ = 0.6 at 37 °C with orbital rotation. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the cultures for a final concentration of 1 mM to induce expression of the p53 N-terminus, and was then incubated for four hours at 37 ° C with orbital rotation. The cells from each 500 mL culture were pelleted using centrifugation at 10,000 *g* for 15 minutes at 4 °C; the pellets were stored at -80 °C until lysis and purification.

Two (each from 500 mL culture) *E. coli* cell pellets were resuspended in 20 mL Lysis Buffer (6 M guanidine hydrochloride, 10 mM Tris hydrochloride, 100 mM sodium phosphate, pH 8.0) on ice. Denaturing conditions were used to prevent proteases from cleaving the target proteins. The resuspended pellets were sonicated on ice four times 1.5 minutes, with a Duty Cycle of 80 % and output of 5. Lysed cells were centrifuged at 14,000 *g* for 60 minutes at 4 °C to pellet cell debris.

Purification was monitored by absorbance at 280 nm (A_{280}) using a Bio-Rad Biologic LP low pressure chromatography system. Freshly regenerated HIS-Select nickel affinity gel from Sigma-Aldrich (St. Louis, MO) was equilibrated with 30 mL Lysis Buffer prior to loading the entire volume of cell lysate supernatant at room temperature. Wash 1 was performed with 30 mL Lysis Buffer. The second wash was performed under native conditions using 30 mL Wash Buffer 2 (10 mM Tris hydrochloride, 100 mM

sodium phosphate, pH 8.0) to remove the guanidine hydrochloride from the target protein prior to thrombin digest. The third wash was performed with 25 mL Wash Buffer 3 (10 mM Tris hydrochloride, 100 mM sodium phosphate, 10 mM imidazole, pH 8.0). Elution of p53 N-terminus was performed using 25 mL Elution Buffer (10 mM Tris hydrochloride, 100 mM sodium phosphate, 350 mM imidazole, pH 4.3). The wash and elution profile was monitored by A_{280} using a Biologic system. The Ni-NTA media was washed with water, 1 M sodium hydroxide, followed by water prior to purifying the other two pellets' lysate. Eluates were dialyzed twice at 4 °C against 4 L of 20 mM Tris hydrochloride, 100 mM sodium chloride, pH 8.0. Figures 2.1 – 2.3 show representative chromatograms for nickel affinity purification.

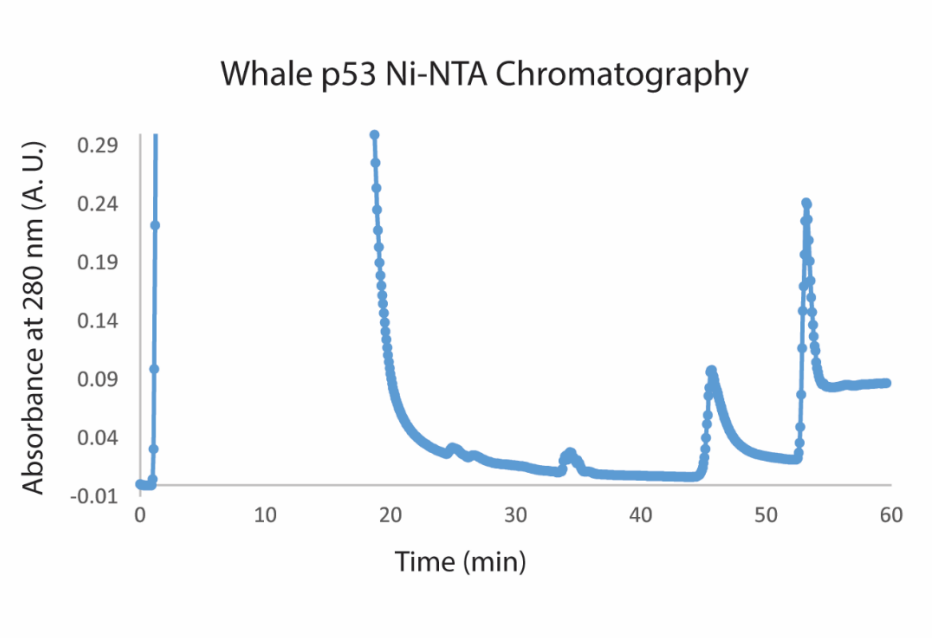

Figure 2.1 Whale p53(1-86) Ni-NTA Chromatogram. An example chromatogram from nickel affinity purification of Whale p53 N-terminus. The increase in A_{280} from 0 – 20 minutes represents the flow-through and Wash 1. The small peak at approximately 35 minutes was Wash 2. The peak at approximately 47 minutes was Wash 3. The eluate was collected from the peak at approximately 55 min.

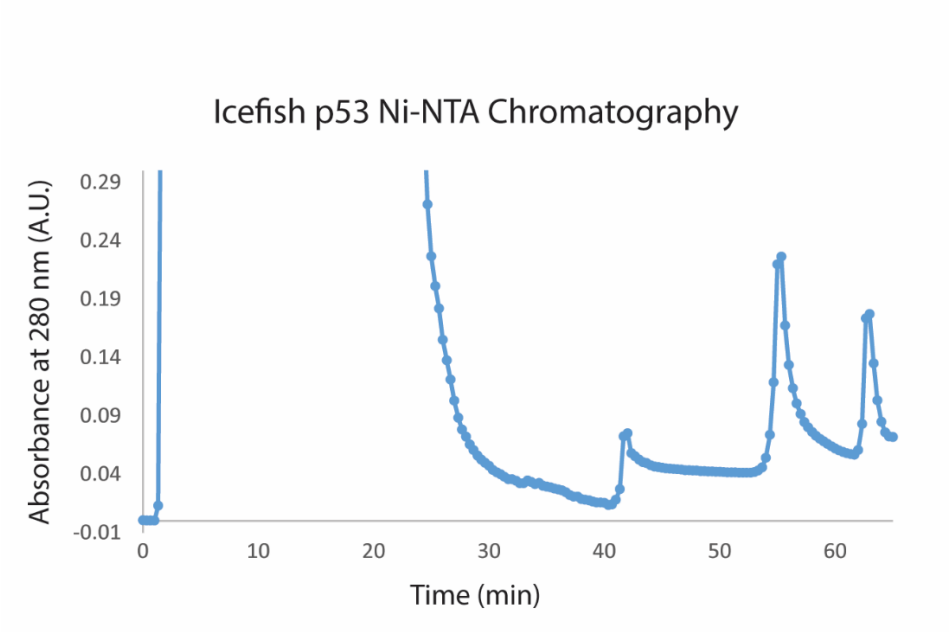

Figure 2.2 Icefish p53(1-89) Ni-NTA Chromatogram. An example chromatogram from the nickel affinity purification of Icefish p53 N-terminus. The increase in A_{280} from 0 – 28 minutes represents the flow-through and Wash 1. The small peak at approximately 42 minutes was Wash 2. The peak at approximately 55 minutes was Wash 3. The eluate was collected from the peak at approximately 63 min.

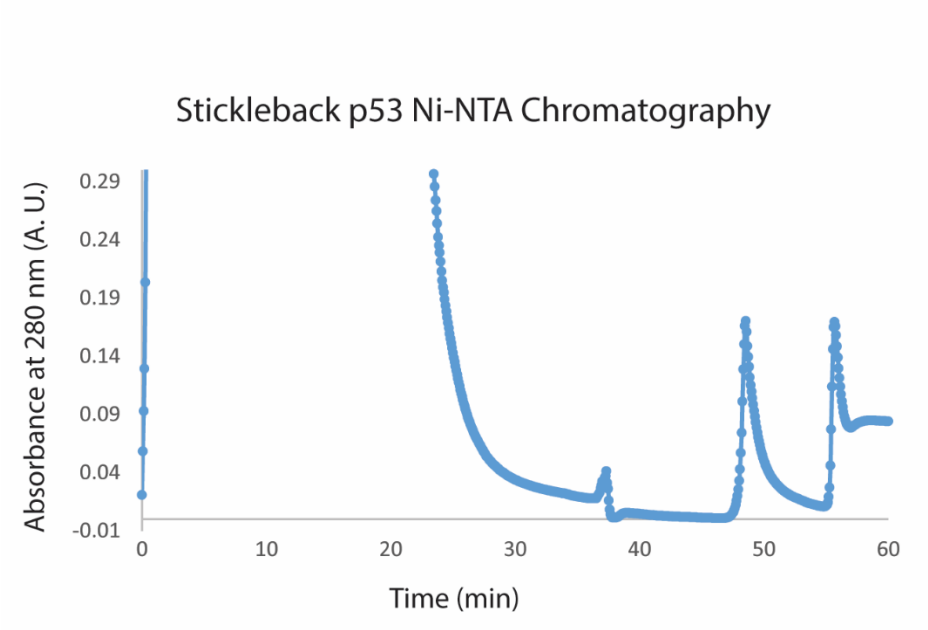

Figure 2.3 Stickleback p53(1-80) Ni-NTA Chromatogram. An example chromatogram from the nickel affinity purification of Stickleback p53 N-terminus. The increase in A_{280} from 0 – 30 minutes represents the flow-through and Wash 1. The small peak at approximately 37 minutes was Wash 2. The peak at approximately 49 minutes was Wash 3. The eluate was collected from the peak at approximately 57 min.

The 6x-His tag was removed by adding approximately 10 U of thrombin the dialyzed protein, followed by incubation protected from light, for maximal activity and to protect the thrombin from degradation, for two hours. Diethylaminoethyl (DEAE) weak anion exchange was performed using approximately 10 mL degassed DEAE Sephacel media (GE Healthcare; Chicago, IL) that was equilibrated with Wash Buffer 1 (20 mM sodium acetate, 25 mM sodium chloride, pH 4.8). Digested protein was applied to the equilibrated DEAE and allowed to flow by gravity. The column was washed with 30 mL Wash Buffer 1, 25 mL Wash Buffer 2 (20 mM sodium acetate, 50 mM sodium chloride, pH 4.8), and 25 mL Wash Buffer 3 (20 mM sodium acetate, 100 mM sodium chloride, pH 4.8). Protein was eluted using Elution Buffer (20 mM sodium acetate, 400 mM sodium chloride, pH 4.8). Wash and elution profiles were monitored by A_{280}. Eluted protein was dialyzed at 4 °C against 4 L of 10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0. Protein samples were concentrated in a Thermo Fisher Savant DNA 120 SpeedVac (Waltham, MA) to a concentration of 0.5 – 0.6 mg/mL; samples were again dialyzed against 4 L of Sodium Phosphate Buffer (10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0) at 4 °C. Example chromatograms for anion exchange chromatography are shown in Figures 2.4 – 2.6.

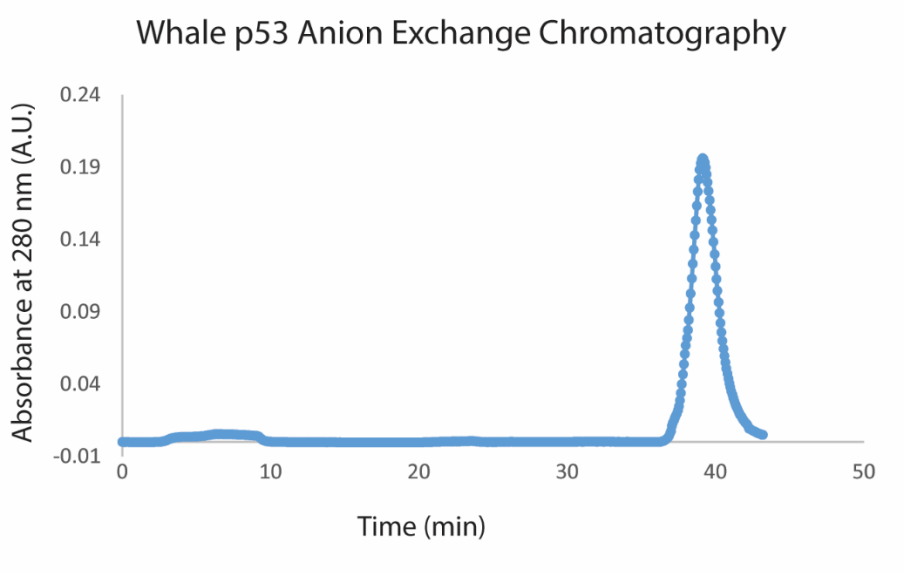

Figure 2.4 Whale p53(1-86) Anion Exchange Chromatogram. An example chromatogram from the weak anion exchange purification of Whale p53 N-terminus. The eluate was collected from the peak from approximately 38 – 42 minutes.

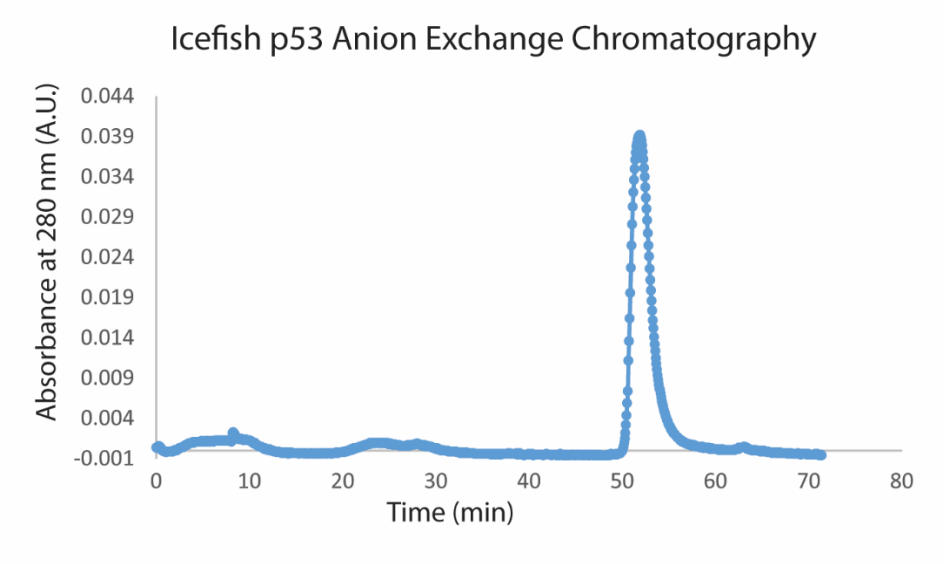

Figure 2.5 Icefish p53(1-89) Anion Exchange Chromatogram. An example chromatogram from the weak anion exchange purification of Icefish p53 N-terminus. The eluate was collected from the peak, approximately 51 - 55 minutes.

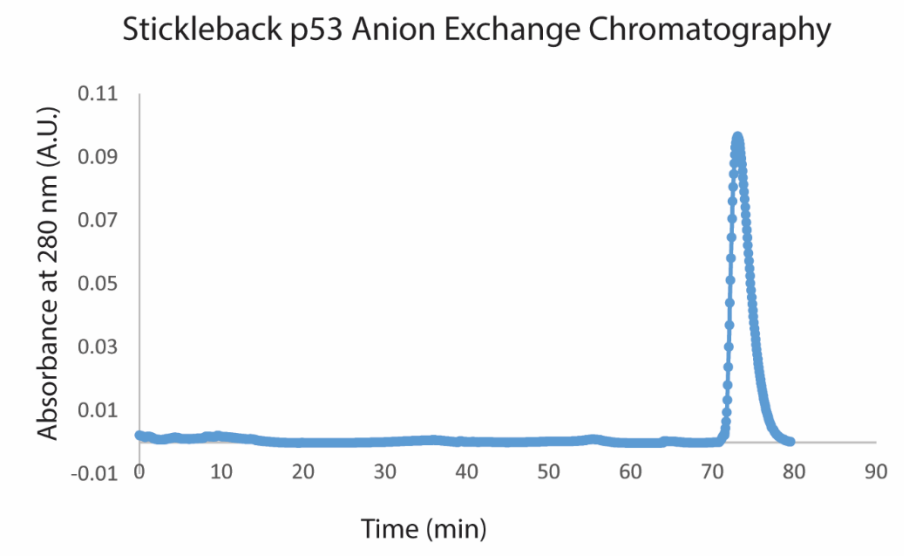

Figure 2.6 Stickleback p53(1-80) Anion Exchange Chromatogram. An example chromatogram from the weak anion exchange purification of Stickleback p53 N-terminus. The eluate was collected from the peak, approximately 72 - 78 minutes.

2.4 SDS-PAGE

Samples throughout purification were analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to identify p53 N-terminus homologs as well as sample purity. Tris-hydrochloride SDS-PAGE gels were hand-cast with a 10 % resolving gel and a 4 % stacking gel. Samples were combined 1:1 with 2x Laemmli Buffer from Bio-Rad (Hercules, CA) with β-mercaptoethanol added (950:50), and boiled for 8 minutes prior to loading on the gel. Precision Plus Protein All Blue Standards from Bio-Rad were used to approximate molecular weight of proteins of interest. Gels were electrophoresed using a Bio-Rad PowerPac HV power supply at 4 °C in tris-glycine SDS buffer (25 mM Tris-base, 190 mM glycine, 0.1 % (w/v) SDS) at 100 V for 10 min followed by approximately 45 minutes at 200 V.

Silver staining was performed after rinsing the electrophoresed gel three times with ultra-pure water. The gel was incubated with Fixer Solution (50 % (v/v) methanol, 2 M acetic acid, 0.05 % (v/v) formaldehyde) for 30 minutes with gentle rocking. The gel was then rinsed twice for 15 minutes in 50 % (v/v) ethanol. The gel was then incubated for 1 minute with Pretreatment Solution (5 mM sodium thiosulfate), followed by three 20 second washes with ultra-pure water. The gel was incubated at 4 °C in Silver Solution (12 mM silver nitrate, 0.05 % (v/v) formaldehyde) for 20 minutes with gentle rocking, followed by two 20 second rinses with ultra-pure water. The gel was then reduced using Reducing Solution (300 mM sodium carbonate, 0.15 mM sodium thiosulfate, 0.05 % (v/v) formaldehyde) for 30 seconds followed by approximately 1 minute. The reaction was stopped by rinsing with ultra-pure water followed by rinsing and storage in 5 % (v/v)

methanol, 5 % (v/v) acetic acid. A Bio-Rad Molecular Imager ChemiDoc XRS+ imaging system was used to image silver stained SDS-PAGE gels.

2.5 Mass Spectrometry

Each of the samples after purification were analyzed using electrospray ionization mass spectrometry (ESI-MS) to confirm the protein of interest and the successful removal of the 6x-His tag by measuring m/z and calculating molecular weight. Protein samples were diluted to 10 – 15 μM with ultra-pure water and dialyzed against 4 L ultra-pure water overnight at 4 °C prior to being submitted for mass spectrometry to Dr. Xiaopeng Li's lab. Samples were combined with 1 % formic acid in acetonitrile (1:1) and infused on a Waters Synapt G2 ESI-Q-TOF mass spectrometer and analyzed using MassLynx software (Milford, MA). An ESI capillary voltage of 3.5 kV was used, sample cone voltage was 35 V, and the extraction cone voltage was 3.5 V. The source temperature was 150 °C with a cone gas flow of 10 L/h nitrogen gas. The desolvation temperature was 180 °C with a desolvation gas flow of 700 L/h nitrogen gas. The molecular weight of each sample was calculated from m/z values using Equation 2.1, where m is measured m/z, M is molecular weight, z(H) is the charge state multiplied by the mass of a proton, and z is the charge state.

Equation 2.1: $m = \frac{(M + z(H))}{z}$

2.6 Circular Dichroism

Circular dichroism (CD) was performed on purified p53 N-terminus homologs to elucidate secondary structure as temperature increases. Protein samples were diluted to approximately 0.170 mg/mL with filtered dialysate (Sodium Phosphate Buffer) prior to analysis. Measurements were acquired from 197 – 245 nm, 0.5 nm increments using a Jasco J-710 spectropolarimeter, a Jasco spectropolarimeter power supply, and a Jasco PFD-425S peltier. Absorption range was decreased when the HT voltage was greater than 700 V for the previous temperature scan. Scans were completed at temperatures of 5 – 85 °C, in 10 °C increments. Eight scans were performed at each temperature at a rate of 20 nm/minute; raw CD data (millidegrees) reported was the average of the eight scans. Raw data was converted to molar residue ellipticity (MRE) using Equation 2.2 and plotted using Microsoft Excel.

Equation 2.2: $$MRE = \frac{(\text{CD } millidegrees)*(0.1)*(\text{mean residue weight } \frac{g}{mol})}{(\text{pathlength } cm)*(\text{concentration} \frac{mg}{mL})}$$

2.7 Size Exclusion Chromatography

Size exclusion chromatography (SEC) was performed using Sephadex G-75 gel filtration media from GE Healthcare (Little Chalfont, UK) to determine the distribution coefficients (K_d) and calculate the hydrodynamic radii (R_h) of the p53 N-terminus homologs. All samples were combined with Blue Dextran and DNP Aspartate to determine the column void volume and the column total volume for each trial. Five trials were performed for each sample using Sodium Phosphate Buffer.

G-75 media was prepared using approximately 4 g of media and allowing the media to swell in Sodium Phosphate Buffer for three hours prior to degassing the media. The column was allowed to pack by gravity. A minimum of three column volumes of Sodium Phosphate Buffer were passed through the column for equilibration. The column was kept at room temperature for swelling, degassing, equilibration, and experimental trials. Media was stored in 20 % ethanol at 4 °C when not in use.

A stock blue dextran and N-(2,4-dinitrophenyl)-L-aspartic acid solution (BDA) was prepared in Sodium Phosphate Buffer at 3.0 mg/mL and 0.3 mg/mL, respectively, and stored at 4 °C.

Equine myoglobin, chicken albumin, and staphylococcal nuclease were used as protein standards to compare the size exclusion chromatography K_D with R_h. Equine myoglobin (Sigma-Aldrich M1882) was reconstituted in Sodium Phosphate Buffer to a concentration of 20 mg/mL, filtered with a 0.2 µm polyvinylidene (PVDF) syringe filter, aliquoted, and stored at -80 °C until use. Prior to use, myoglobin was diluted to 10 mg/mL using the same buffer. The myoglobin was prepared by Lance R. English.

Chicken egg albumin (Thermo-Fisher Scientific Acros 40045) was reconstituted in Equilibration Buffer (10 mM sodium phosphate, pH 7.0) at 1 mg/mL. Albumin was applied to a DE52 diethylaminoethyl weak anion exchange column (GE Whatman), degassed and equilibrated with Equilibration Buffer, and allowed to flow by gravity. The elution profile was monitored by A_{280} using a Bio-Rad Biologic LP low pressure chromatography system. The column was washed with 30 mL Equilibration Buffer, followed by 30 mL Wash Buffer 2 (10 mM sodium phosphate, 50 mM sodium chloride, pH 7.0). The albumin was eluted with Sodium Phosphate Buffer. Purity was verified using SDS-PAGE followed by silver staining.

Recombinant staphylococcus nuclease (SNase) was expressed and purified by David Englehardt following previously published protocols[51,52].

Equine myoglobin (10 mg/mL), chicken albumin (1.0 mg/mL), and SNase (1 mg/mL), and purified p53 N-terminus homologs (0.4 – 0.6 mg/mL each) in Sodium Phosphate Buffer were each combined with BDA stock solution, 9:1. For each of five trials, 100 µL of the protein/BDA solution was applied to the column bed and allowed to flow by gravity using Sodium Phosphate Buffer. Elution times for void volume (V_0), elution volume (V_e), and total volume (V_t) were monitored at A_{280} using a Bio-Rad Biologic LP chromatography system. K_d was calculated using Equation 2.3. K_d and experimental R_h values were compared using previously reported R_h values determined by crystallographic measurements of the protein standards and plotted versus K_d using Microsoft Excel to determine R_h values of the p53 N-terminus proteins.

Equation 2.3: $K_d = \frac{V_e - V_0}{V_t - V_0}$

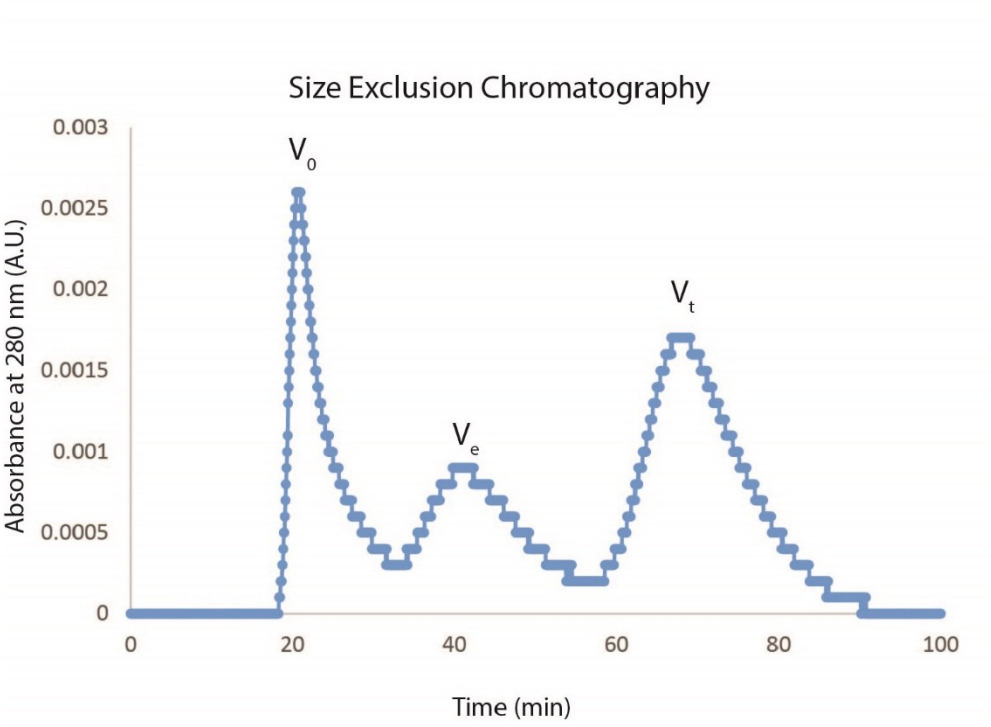

Figure 2.7 SEC Chromatogram. An example chromatogram using absolute absorption from size exclusion chromatography using SNase. The first peak (void volume, V_0) is blue dextran, the second peak (elution volume, V_e) is SNase, and the third peak (total volume, V_t) is DNP-aspartate.

2.8 Dynamic Light Scattering

Dynamic light scattering (DLS) was used to determine the hydrodynamic radius of the p53 N-terminus homologs. After concentration, protein samples were diluted if necessary to 0.5 – 0.6 mg/mL using dialysate (Sodium Phosphate Buffer) and filtered using a 0.2 μm PVDF syringe filter. An average of five scans was recorded per twelve trials from 5 – 75 °C, at 10 °C intervals using a Zetasizer Nano ZS with peltier temperature control. Data were analyzed using GraphPad Prism 6. Outliers were determined using Grubb's Test.

III. RESULTS AND DISCUSSION

3.1 Introduction

There is a structure-function relationship describing that the function of a protein is dictated by its folded structure. However, IDPs have no defined stable tertiary structure, but do have important functionalities[2]. Many IDPs function as transcription factors or cell-signaling proteins[6]. Amino acid sequence dictates structure in folded proteins as well as the lack of structure in IDPs[52]. Across multiple species, sequences of IDPs vary considerably, but disorder and function are conserved[10].

Studies have been performed using human p53 N-terminus, a known IDP, to elucidate secondary structure as well as change in structure as a function of temperature. Continuing these studies may help provide a better understanding of how disorder is important for function. Previous studies have shown that along with other IDPs, human p53(1-93) structure is highly temperature sensitive[21,45,53]. As temperatures increase, PPII content and R_h decrease significantly[21,45]. CD has been a powerful tool in measuring secondary structure as a function of temperature; PPII structure can be monitored from 197 nm – 245 nm. Temperature effects on the size of human p53(1-93) have been demonstrated with SDS-PAGE and DLS[21,45].

One common method of studying sequence effects on protein structure is by performing mutations. However, it is difficult to maintain protein function with amino acid substitutions. In an effort to study sequence effects on IDPs and their temperature dependence, p53 homologs were chosen that are functional at physiological temperatures differing from that of humans. Non-mammalian vertebrates that live in cold

environments, icefish and stickleback, were selected to elucidate the structure-temperature dependence of p53 N-terminus in cold-adapted species. Whale p53 N-terminus was chosen to act as a control; whales are cold-adapted mammals that regulate body temperature near 37 °C. These data were compared to that of the established human p53 N-terminus data to determine the effects of temperature on the different p53 N-terminus homologs.

3.2 Assessing the Purity of p53 N-terminus Homologs

Recombinant protein consisting of the intrinsically disordered N-terminal region of p53 from human, whale, icefish and stickleback was expressed in *E. coli* and isolated using nickel affinity chromatography as well as anion exchange chromatography. The purity of recombinant protein obtained in this manner was assessed using SDS-PAGE and ESI-MS. All three of the p53 N-terminus homologs had multiple bands visualized after nickel affinity chromatography, indicating the possible presence of other protein species (Figures 3.1 – 3.3). Following anion exchange, gel images indicated that most of the impurities were removed, however there were two distinct bands visible in the whale p53(1-86) and icefish p53(1-89) silver-stained SDS-PAGE gels. A second band was also observed in some of the stickleback p53(1-80) anion exchange eluates of various purifications (image not shown). Based on the gel image intensities, it appeared that whale, icefish, and stickleback p53 N-terminus homologs migrated as would approximately 18 kDa, 17 kDa, and 27 kDa proteins, respectively. IDPs often show anomalous electrophoretic mobility that is not indicative of molecular weight[45].

To validate the identities of each protein, as well as to verify that the 6x His tags were successfully removed, ESI-MS was performed on each sample prior to use in experiments. Figures 3.4 – 3.6 are spectra of p53 N-terminus homolog samples that were used for further experiments. The calculated molecular weights for each protein were determined for samples both with and without the 6x His tag using ExPASy ProtParam (Table 3.1). Whale p53(1-86) had a calculated molecular weight of 9,413 Da and an experimental molecular weight from ESI-MS of 9,469 Da. The calculated molecular weight of icefish p53(1-89) was 10,256 Da, and the measured molecular weight of icefish p53(1-89) was 10,253 Da. Stickleback p53(1-80) had a calculated molecular weight of 8,780 Da and a measured molecular weight of 8,810 Da. For each homolog, the calculated molecular weights from each of the charge states were consistent with each other, indicating the presence of a single, pure protein in each sample. The calculated molecular weights for all three samples indicated that the correct proteins were overexpressed and purified as well as successful removal of the 6x His tag with thrombin digestion.

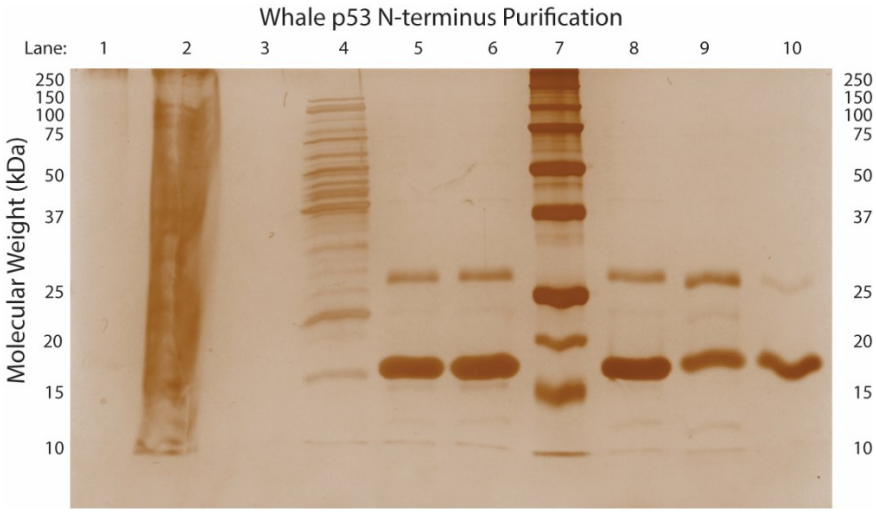

Figure 3.1 Whale p53(1-86) SDS-PAGE. An example silver-stained SDS-PAGE gel, Lane1: Ni-NTA initial flow through; Lane 2: Ni-NTA Wash 1; Lane 3: Ni-NTA Wash 2; Lane 4: Ni-NTA Wash 3; Lane 5: Ni-NTA eluate from first two pellets; Lane 6: Ni-NTA eluate from second two pellets; Lane 7: Protein standards; Lane 8: combined Ni-NTA eluates; Lane 9: Ni-NTA eluate after thrombin digest; Lane 10: Anion exchange eluate after dialysis.

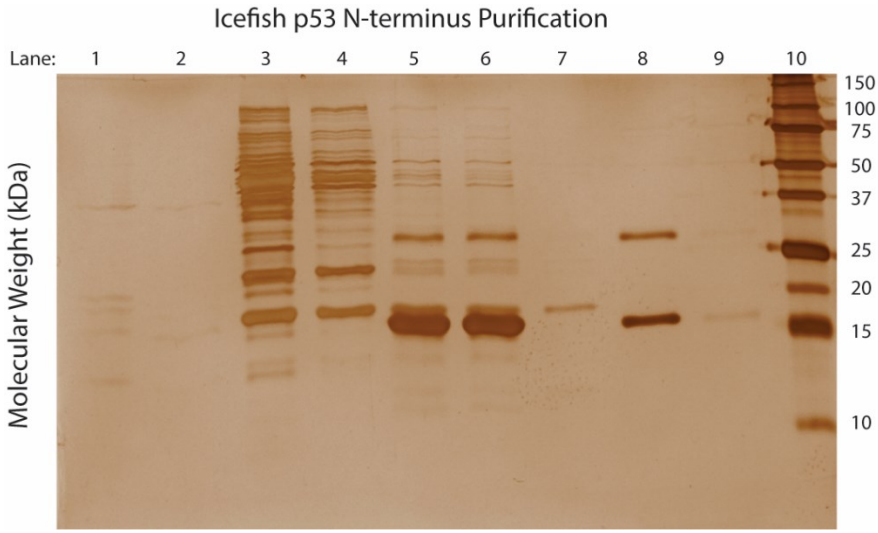

Figure 3.2 Icefish p53(1-89) SDS-PAGE. An example silver-stained SDS-PAGE gel from icefish p53(1-89) purification, Lane1: Ni-NTA Wash 2 (from first two pellets); Lane 2: Ni-NTA Wash 2 (from second two pellets); Lane 3: Ni-NTA Wash 3 (from first two pellets); Lane 4: Ni-NTA Wash 3 (from second two pellets); Lane 5: Ni-NTA eluate from first two pellets; Lane 6: Ni-NTA eluate from second two pellets; Lane 7: Anion exchange flow through; Lane 8: Anion exchange eluate (initial ~80 % of the peak); Lane 9: Anion exchange eluate (remaining portion of the peak); Lane 10: Protein standards.

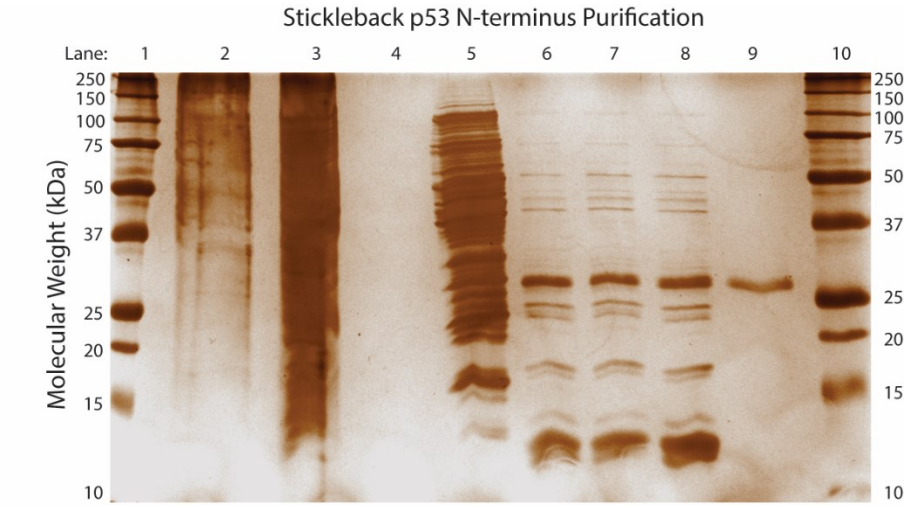

Figure 3.3 Stickleback p53(1-80) SDS-PAGE. An example silver-stained SDS-PAGE gel from stickleback p53(1-80) purification, Lane 1: Protein standards; Lane 2: Ni-NTA initial flow through; Lane 3: Ni-NTA Wash 1; Lane 4: Ni-NTA Wash 2; Lane 5: Ni-NTA Wash 3; Lane 6: Ni-NTA eluate from first two pellets; Lane 7: Ni-NTA eluate from second two pellets; Lane 8: combined Ni-NTA eluates; Lane 9: anion exchange eluate (after thrombin digestion); Lane 10: Protein standards.

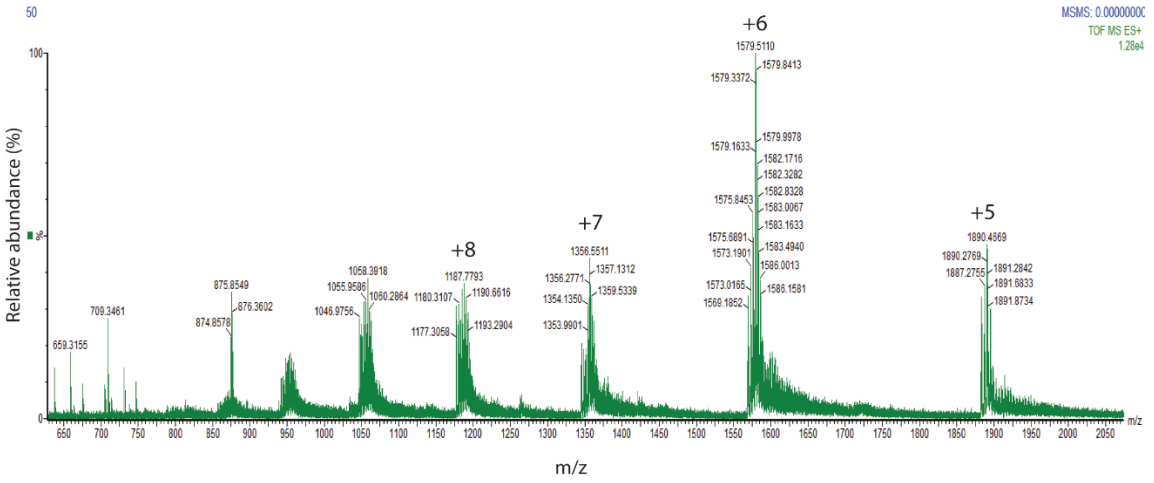

Figure 3.4 Whale p53(1-86) ESI-MS Chromatogram

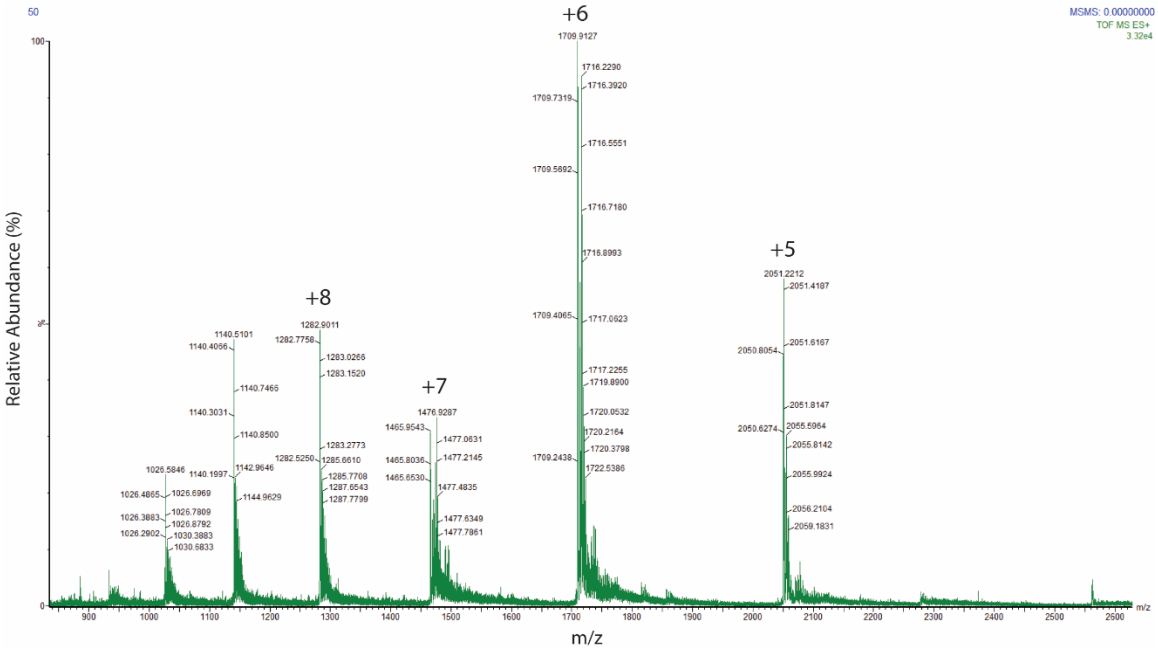

Figure 3.5 Icefish p53(1-89) ESI-MS Chromatogram.

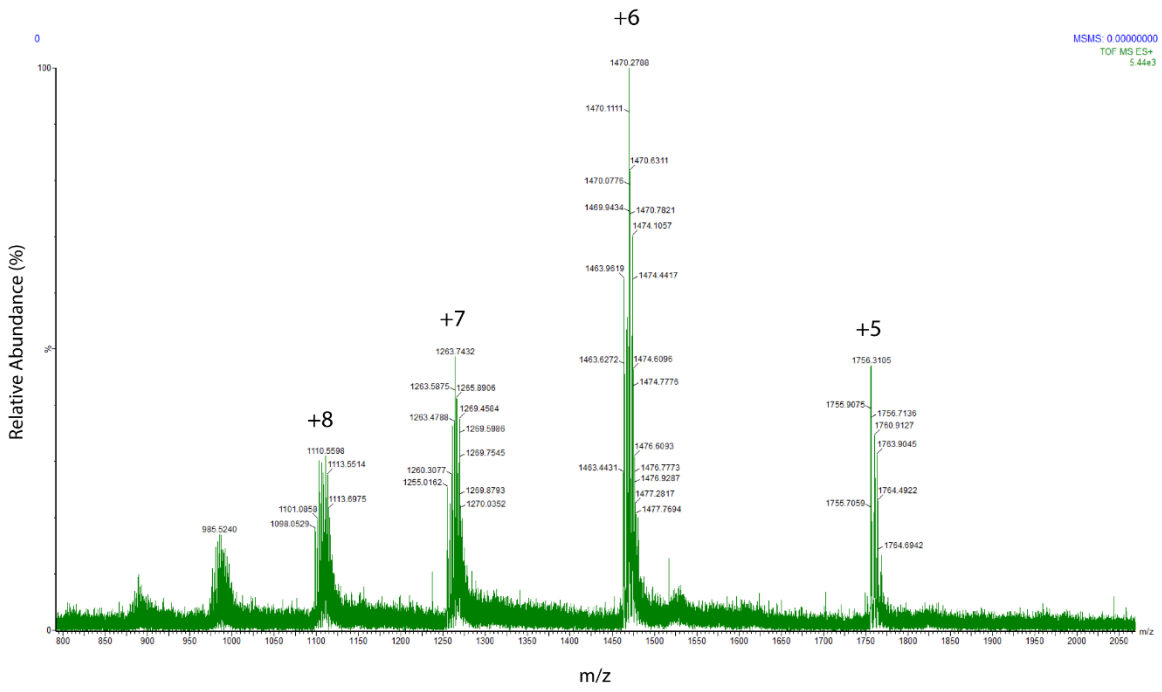

Figure 3.6 Stickleback p53(1-80) ESI-MS Chromatogram.

Table 3.1 Molecular Weight Verification of p53 N-terminus Homologs.

Protein	Calculated MW (Da) with 6x His	Calculated MW (Da)	Measured MW (Da)
Whale p53(1-86)	11,364.57	9,413.41	9,469
Icefish p53(1-89)	12,207.54	10,256.38	10,253
Stickleback p53(1-80)	10,730.90	8,779.73	8,810

3.3 Circular Dichroism and Secondary Structure

CD uses right and left circularly polarized light that is absorbed differentially by chiral molecules. Biological small molecules like amino acids generally have at least one chiral center, with the exception of glycine. As a result, protein structure exhibits chirality that is evident in CD spectra. Different secondary structures, such as α-helices, β-sheets, and even PPII, absorb the polarized light differently, allowing identification of secondary structures present in a protein. The trademarks of CD spectra for PPII were elucidated as a prominent local maxima at 221 nm and a local minima around 200 nm[54]. Human p53(1-93) CD spectra exhibits these characteristic properties of PPII and shows a decrease in PPII structure as temperature increases[45].

CD was performed on each sample to investigate the secondary structural properties of the p53 homologs. The CD spectra for human p53(1-93) was consistent with the previously published results, with the exception of the results for 25 °C (Figure 3.7)[45]. There was a local maximum at 221 nm, associated with PPII structure, which decreased as the temperature increased. The isochromatic point was approximately 210 nm. The minima at approximately 200 nm became less pronounced as the temperature increased.

The spectra for whale p53(1-86) closely resembled that of human p53(1-93) (Figure 3.8). The local maximum at 221 nm was approximately the same in whale as in human when comparing MRE values at each temperature to the MRE values at 85 °C (Figures 3.7 B and 3.8 B). The minimum at approximately 200 nm also behaved similarly to that of human.

Icefish p53(1-89) CD spectra exhibited several differences from that of human p53(1-93) or whale p53(1-86) (Figure 3.9). The local maximum was much less prominent

than either human or whale. The minima at approximately 200 nm was much less negative than that of either human or whale. Though there appeared to be less PPII structure in icefish p53(1-89), there was still a correlation between temperature and PPII content. As the temperature increased, the local maxima decreased, and therefore the PPII content decreased.

The spectra for stickleback p53(1-80) displayed a local maxima at approximately 219 nm (Figure 3.10). The MRE difference at the maxima was surprisingly more pronounced than that of icefish, and was instead more similar to that of human and whale. Similarly to icefish p53(1-89), stickleback p53(1-80) had a less intense minima than either human or whale.

The MRE difference for each temperature from that of 85 °C was averaged between 220 – 222 nm to generate a plot comparing all four p53 homologs (Figure 3.11). Each homolog demonstrated linearity, indicating a non-cooperative unfolding of p53 N-terminus as temperature increases[45]. The human p53(1-93) 25 °C data point does not follow the linear trend well, but was included in the linear regression calculation. Human p53(1-93) and whale p53(1-86) were consistent with each other, with overlapping lines and data points. The slope for stickleback p53(1-80) was slightly lower than that of either human or whale, and the slope for icefish was lower than all of the other homologs. Through CD, human and whale p53 N-terminus homologs demonstrated the most PPII structure as well as the most temperature dependence on PPII.

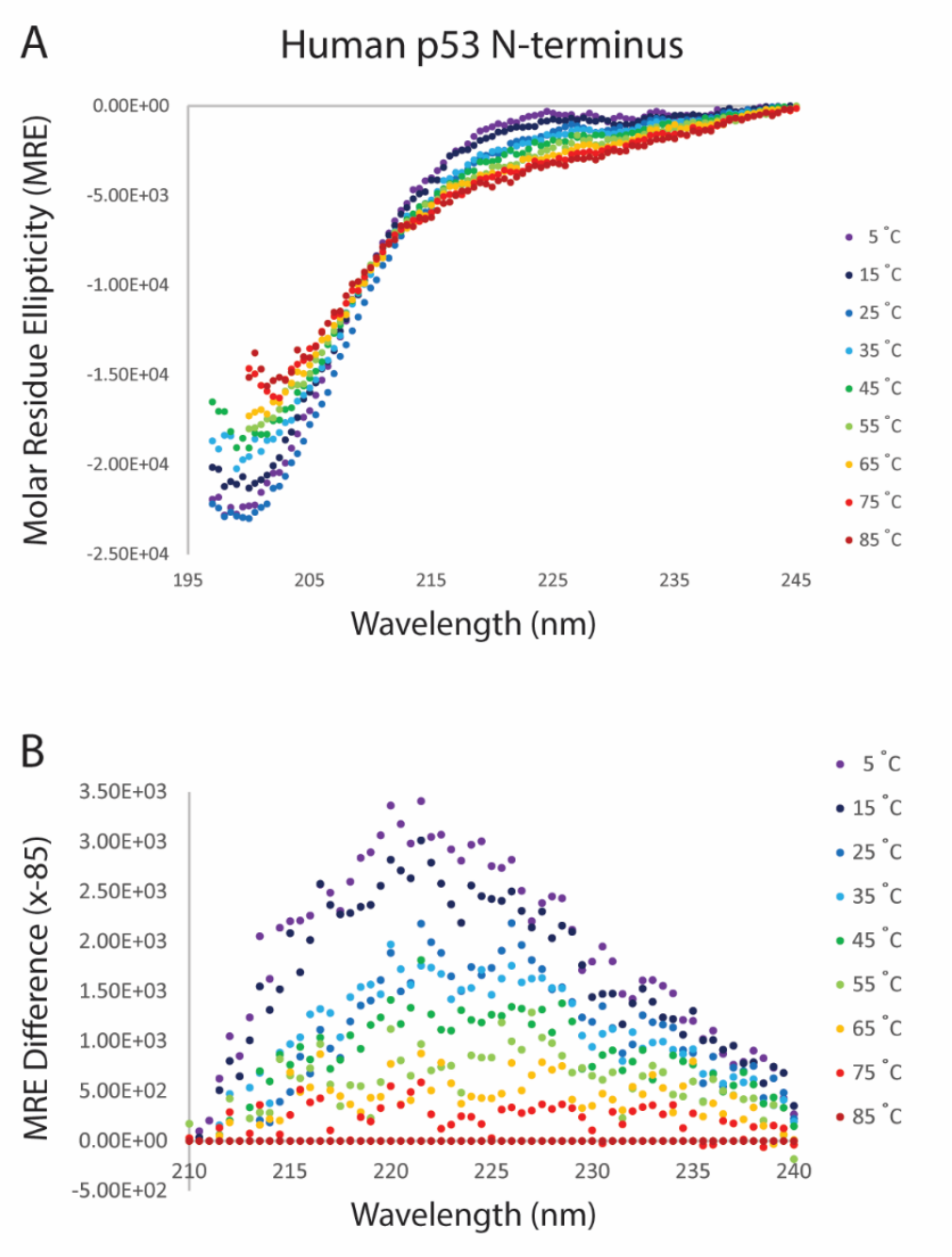

Figure 3.7 Human p53(1-93) CD Results. A) CD spectra is shown for human p53 N-terminus at nine temperatures: 5 °C (purple), 15 °C (dark blue), 25 °C (blue), 35 °C (cyan), 45 °C (green), 55 °C (light green), 65 °C (orange), 75 °C (red), and 85 °C (dark red). B) The MRE for each spectra is shown relative to that of the 85 °C MRE values.

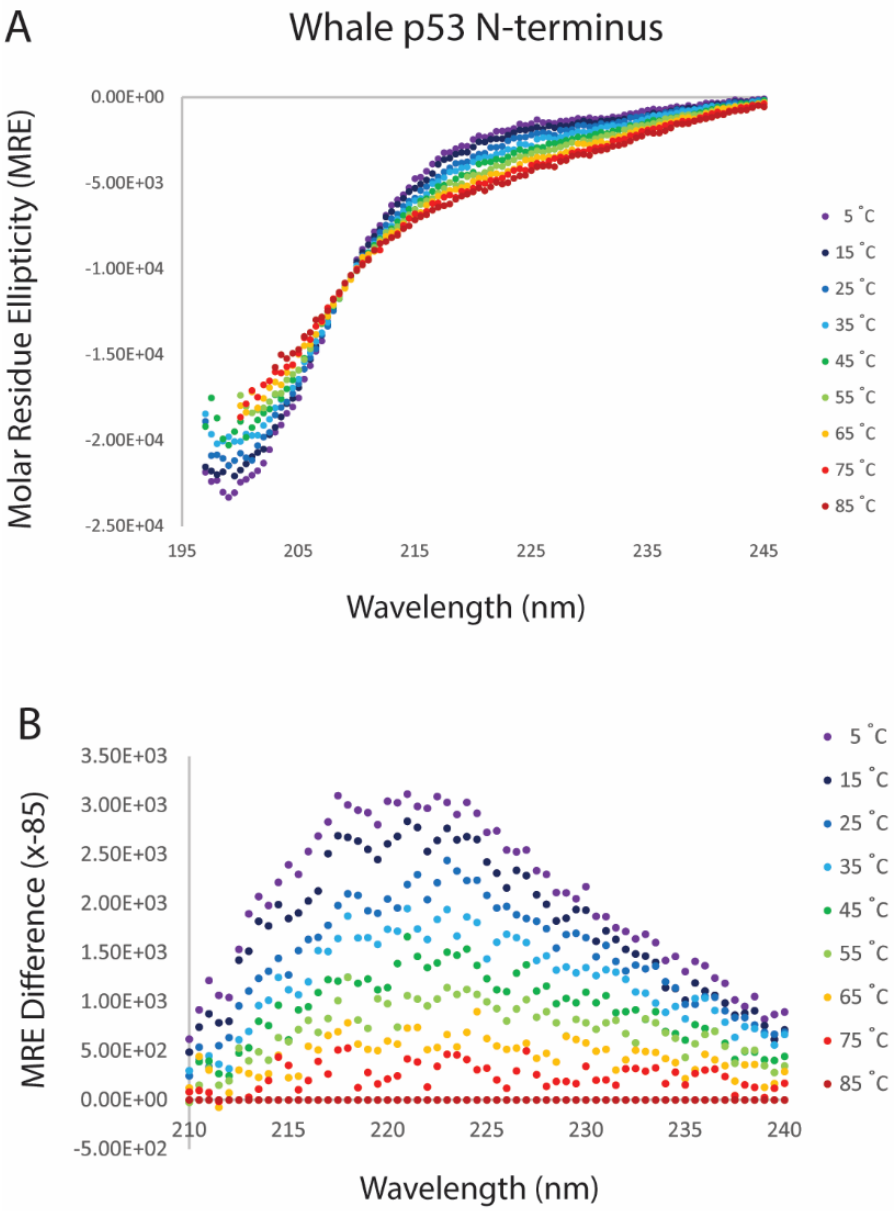

Figure 3.8 Whale p53(1-86) CD Results. A) CD spectra is shown for whale p53 N-terminus at nine temperatures: 5 °C (purple), 15 °C (dark blue), 25 °C (blue), 35 °C (cyan), 45 °C (green), 55 °C (light green), 65 °C (orange), 75 °C (red), and 85 °C (dark red). B) The MRE for each spectra is shown relative to that of the 85 °C MRE values.

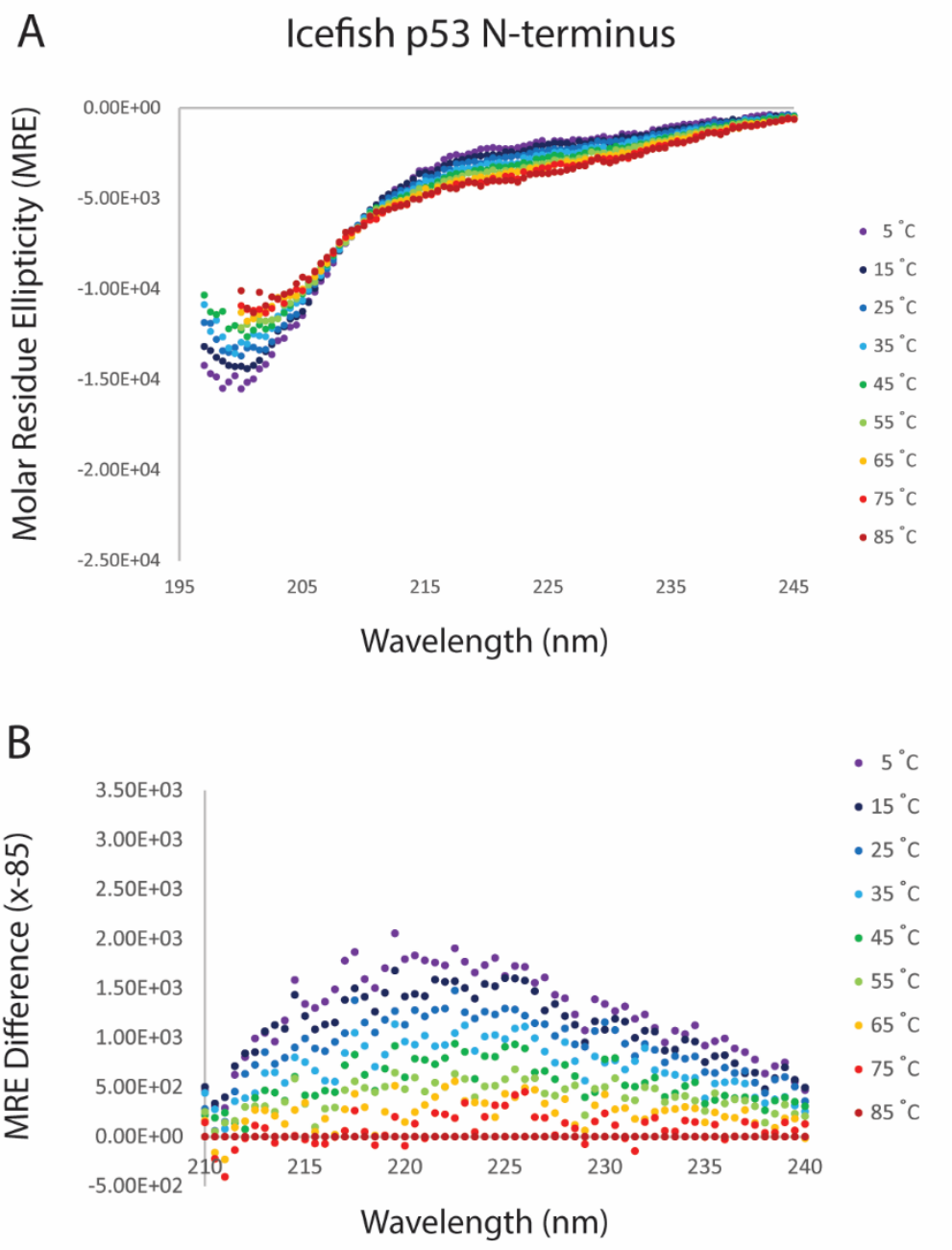

Figure 3.9 Icefish p53(1-89) CD Results. A) CD spectra is shown for icefish p53 N-terminus at nine temperatures: 5 °C (purple), 15 °C (dark blue), 25 °C (blue), 35 °C (cyan), 45 °C (green), 55 °C (light green), 65 °C (orange), 75 °C (red), and 85 °C (dark red). B) The MRE for each spectra is shown relative to that of the 85 °C MRE values.

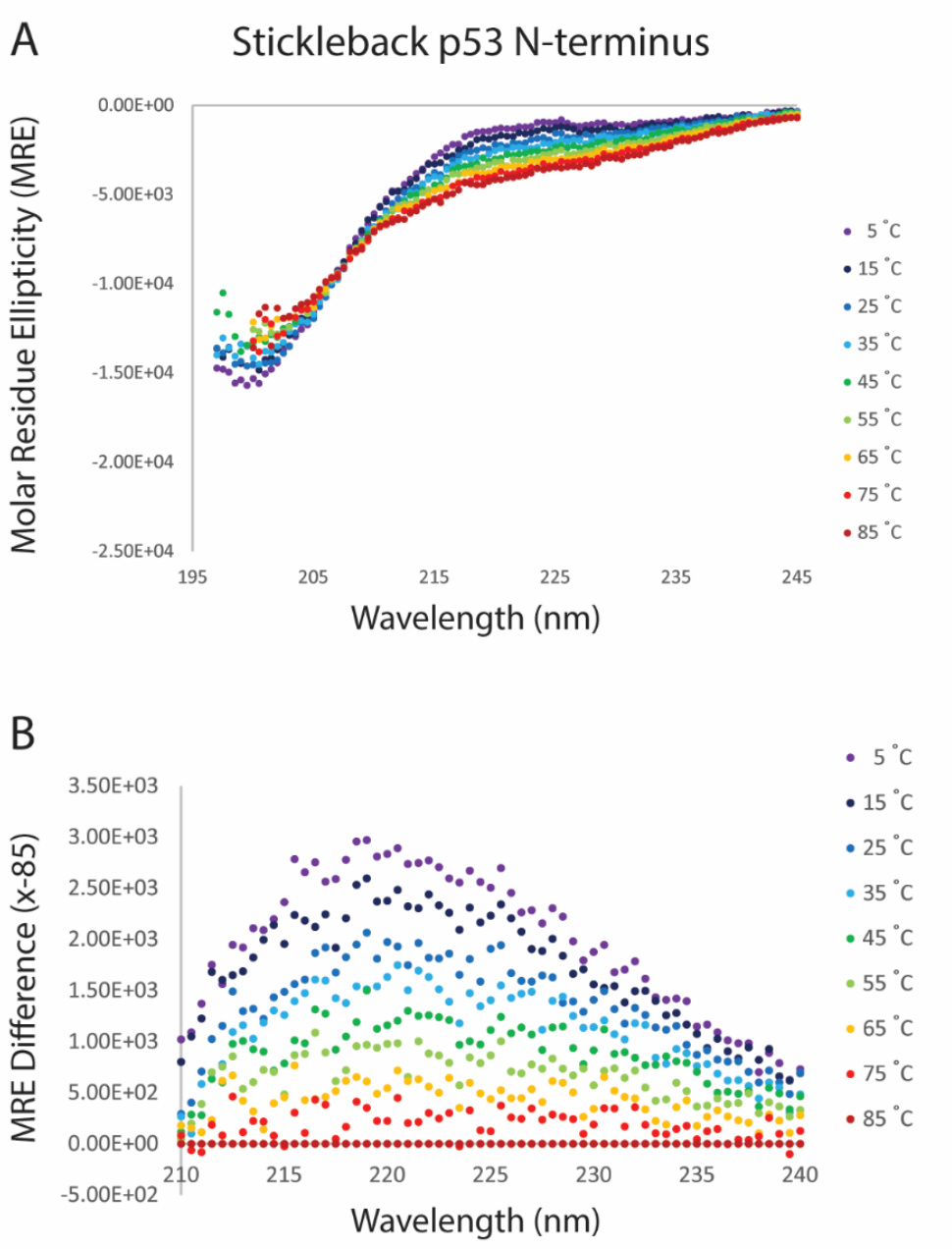

Figure 3.10 Stickleback p53(1-80) CD Results. A) CD spectra is shown for stickleback p53 N-terminus at nine temperatures: 5 °C (purple), 15 °C (dark blue), 25 °C (blue), 35 °C (cyan), 45 °C (green), 55 °C (light green), 65 °C (orange), 75 °C (red), and 85 °C (dark red). B) The MRE for each spectra is shown relative to that of the 85 °C MRE values.

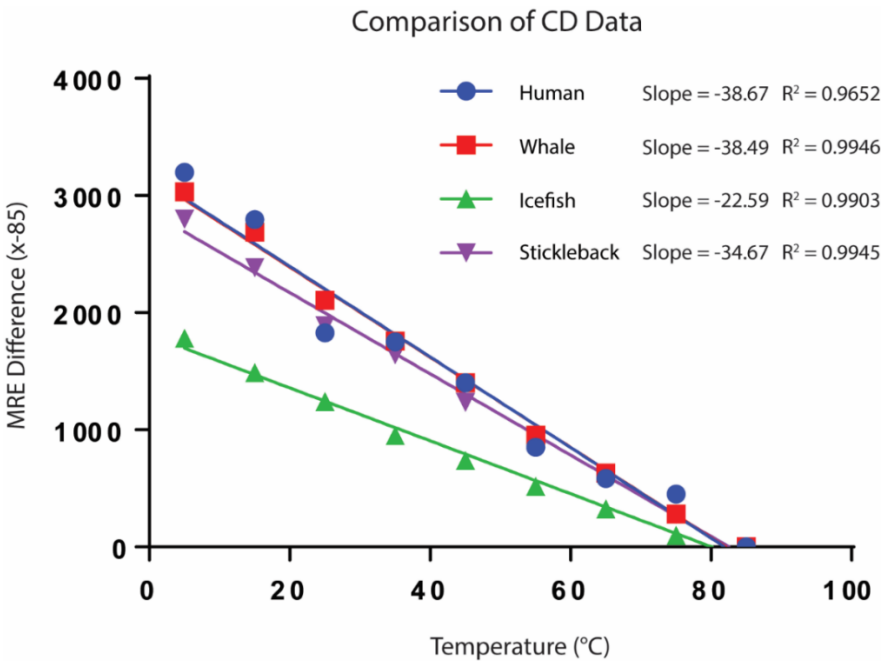

Figure 3.11 Comparison of CD Results. The average MRE values were calculated for each protein from 220 – 222 nm. The 85 °C mean value was subtracted from all averages to create a linear plot as a function of temperature.

3.4 Size Exclusion Chromatography

SEC is a chromatographic method taking advantage of particle pore size to separate molecules by size. Larger molecules enter the pores poorly and elute sooner. Molecules enter the pores more easily as their size decreases, eluting later due to an extended path of travel through the column. For proteins, constant monitoring of the eluate using a UV spectrophotometer, allows one to visualize when the proteins elute based on absorbance at 280 nm. The radius of the molecule, in this case the protein, greatly affects the way the molecule is able to travel through the column; larger, more extended proteins with a greater R_h elute earlier than a folded protein of similar molecular weight (Hong 2002). The R_h of a protein can be calculated by measuring the distribution coefficient (K_d) along with a series of folded protein standards that have known R_h values through other techniques such as X-ray crystallography.

SEC was performed to determine the R_h of each p53 N-terminus homolog. R_h was calculated by comparing the measured K_d values to those of a set of folded protein standards: chicken albumin, staphylococcal nuclease (SNase), and equine myoglobin. K_d and standard deviation values were calculated from five trials of each protein. R_h calculations were made by generating a standard curve using known R_h values, by means of X-ray crystallography, of the folded protein standards and the measured K_d values. Two separate experiments were performed at different times. The first included albumin, SNase, myoglobin, and human p53(1-93). The second included albumin, SNase, myoglobin, whale p53(1-86), icefish p53(1-89), and stickleback p53(1-80). Separate graphs were generated for each experiment.

The molecular weight values used for the p53 N-terminus homologs and SNase were calculated using ExPASy ProtParam. The molecular weights for albumin and myoglobin were provided by the manufacturers. Standard curves were generated using the protein standards to compare molecular weight with K_d (Figures 3.12 – 3.13). The protein standards represent typical globular proteins and have a linear trend between the logarithm of molecular weight (log MW) and K_d (Figures 3.12 – 3.13). All four of the p53 homologs, human, whale, icefish, and stickleback, are below the standard curve generated and have smaller K_d values, representative of proteins with much higher molecular weights than those that were calculated by sequence using ExPASy ProtParam.

The hydrodynamic radii values for the protein standards were measured using X-ray crystallography; R_h was calculated by measuring the greatest distance between two alpha carbons in one subunit and reported in PDB. Based on the linear regression from R_h v. K_d of the protein standards, experimental R_h values of the p53 N-terminus homologs were calculated (Figures 3.14 – 3.15, Table 3.2). The experimental R_h for human p53(1-93) was 32.5 Å, which is similar to the previously reported SEC measurement, 31.95 Å (Perez 2014). Whale p53(1-86) had an experimental R_h of 30.6 Å; the experimental R_h for icefish p53(1-89) was 29.8 Å; and the experimental R_h for stickleback p53(1-80) was 26.8 Å.

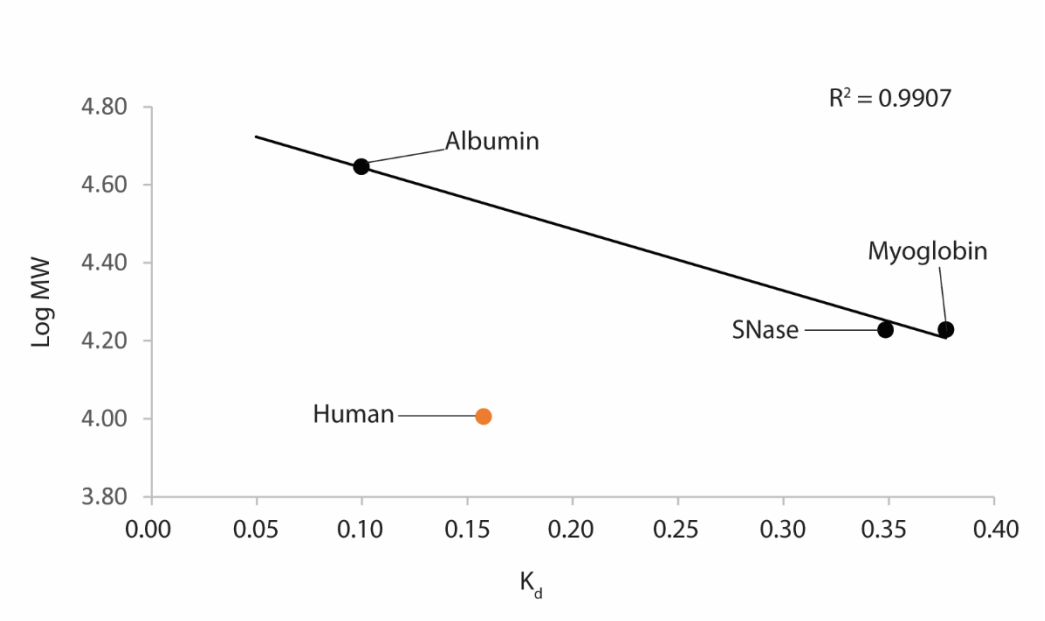

Figure 3.12 SEC Data for Human p53(1-93). The linear curve was determined using protein standards: chicken albumin, SNase, and equine myoglobin. K_d was calculated using Equation 2.2. Molecular weights of human p53(1-93) and SNase were calculated using ExPASy ProtParam. Molecular weights for chicken albumin and equine myoglobin were provided by the manufacturers.

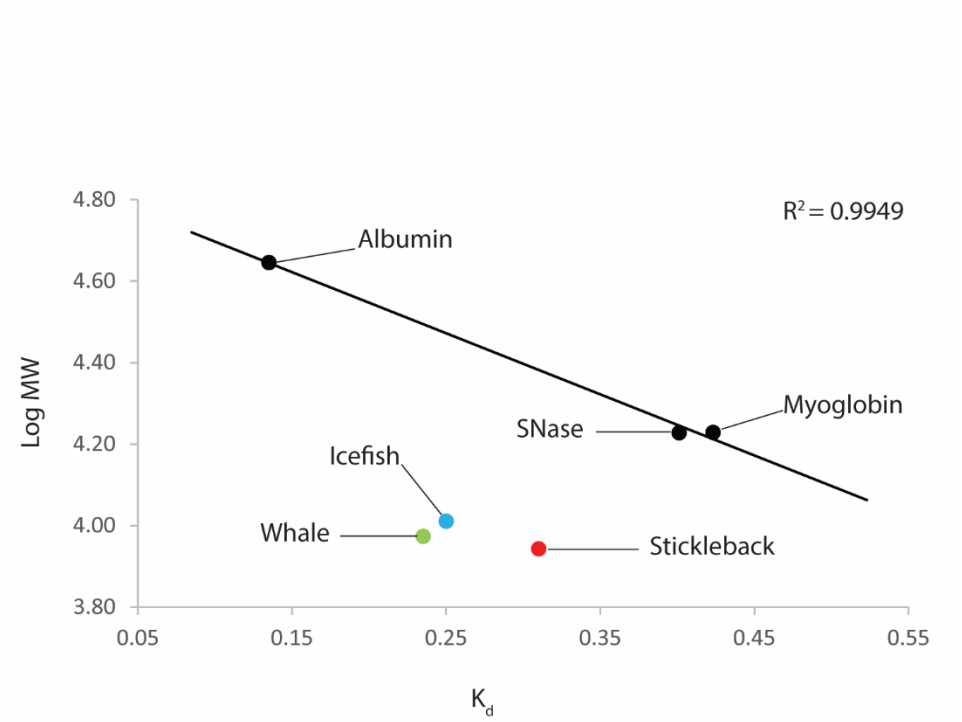

Figure 3.13 SEC Data for p53 N-terminus Homologs. The linear curve was determined using protein standards: chicken albumin, SNase, and equine myoglobin. K_d was calculated using Equation 2.2. Molecular weights of the p53 homologs and SNase were calculated using ExPASy ProtParam. Molecular weights for chicken albumin and equine myoglobin were provided by the manufacturers.

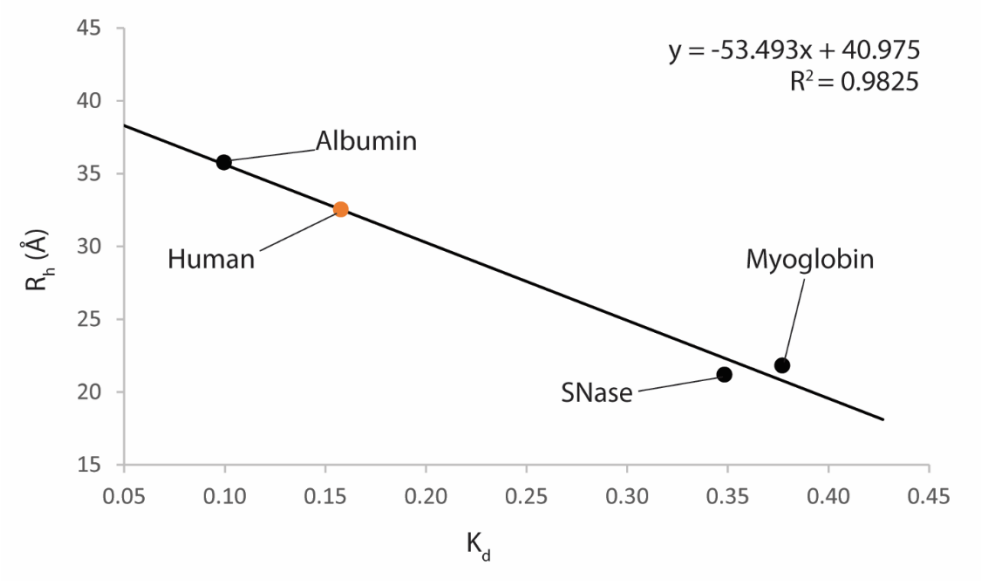

Figure 3.14 Calculation of Human p53(1-93) R_h Values. R_h of the protein standards (from X-ray crystallography) and the measured K_d values from SEC were used to generate a linear curve. This curve was used to calculate the R_h value of human p53 N-terminus using the measured K_d value (Table 3.2).

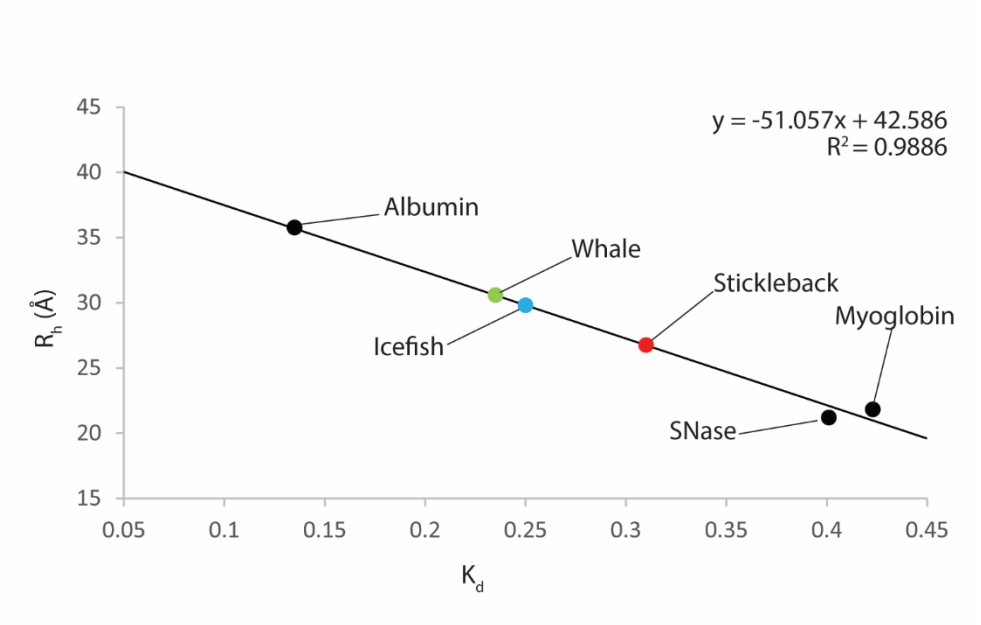

Figure 3.15 Calculation of p53 N-terminus R_h Values. The R_h of the protein standards (from X-ray crystallography) and the measured K_d values from SEC were used to generate a linear curve. This curve was used to calculate the R_h values of the p53 homologs using their measured K_d values (Table 3.2).

Table 3.2 Calculated K_d and R_h values from SEC

Protein	K_d	Standard Deviation	MW (Da)	Log MW	R_h experimental (Å)
Human p53(1-93)	0.158*	0.00208	10,123	4.005	32.5 ± 0.047
Whale p53(1-86)	0.235	0.000777	9,413	3.974	30.6 ± 0.018
Icefish p53(1-89)	0.250	0.00184	10,256	4.011	29.8 ± 0.042
Stickleback p53(1-80)	0.310	0.000465	8,780	3.943	26.8 ± 0.011

*Rh experimental value was calculated using the linear regression in figure 3.11. All other Rh experimental values were calculated using the linear regression in figure 3.12.

3.5 Dynamic Light Scattering

DLS measures R_h by measuring light scatter off of particles tumbling through solution. The Brownian motion of the particle movement in solution is measured through the scatter of the light. The larger the molecule is, the slower the Brownian motion will be. R_h can then be calculated from the light scatter caused by the Brownian motion using the Stokes-Einstein relationship. The advantage to this technique is that the R_h can be measured as a function of temperature.

DLS was performed on whale, icefish, and stickleback p53 N-terminus homologs to measure R_h as a function of temperature from 5 °C – 75 °C. Data from twelve trials per protein were analyzed by constructing box and whiskers plots using GraphPad Prism. It has been previously reported that human p53(1-93) has a R_h of 32.8 Å at 25 °C when measured by DLS; it also had a trend characteristic of IDPs wherein there was a linear decrease in R_h as the temperature increased[47].

DLS measurements for whale p53(1-86) demonstrated a mild decrease in mean R_h as the temperature increased (Figure 3.16). The mean R_h values decreased from 33.4 Å at 5 °C to 29.2 Å at 75 °C. Icefish p53(1-89) mean R_h values decreased more than those for whale p53(1-86) as temperatures increased (Figure 3.17). Mean R_h values for icefish p53(1-89) decreased from 34.5 Å at 5 °C to 27.3 Å at 75 °C. Stickleback p53(1-80) mean R_h values were relatively unchanged with the change in temperature (Figure 3.18). The mean R_h values at 25 °C for whale, icefish, and stickleback p53 N-terminus homologs were 33.5 Å, 31.6Å, and 34.5 Å, respectively. Mean R_h values measured by DLS at 25 °C were compared to the R_h values measured using SEC at room temperature (approximately 21 – 22 °C) in Table 3.3.

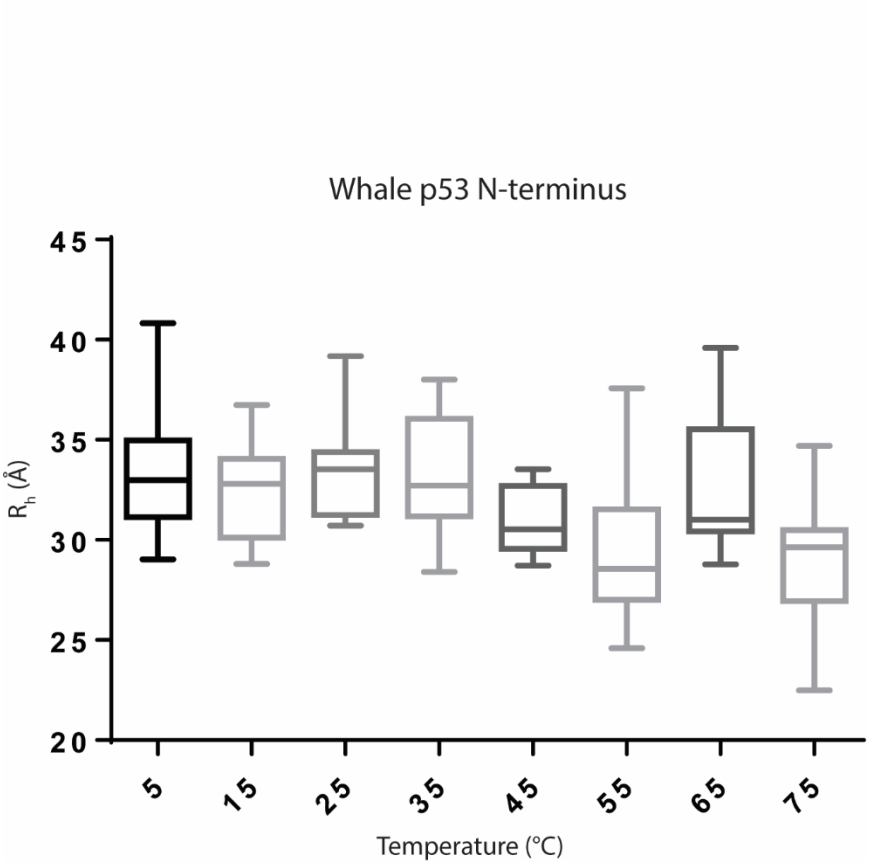

Figure 3.16 DLS Data for Whale p53(1-86). Box and whisker plots were generated from DLS data (n=12) to visualize the change in R_h as a function of temperature.

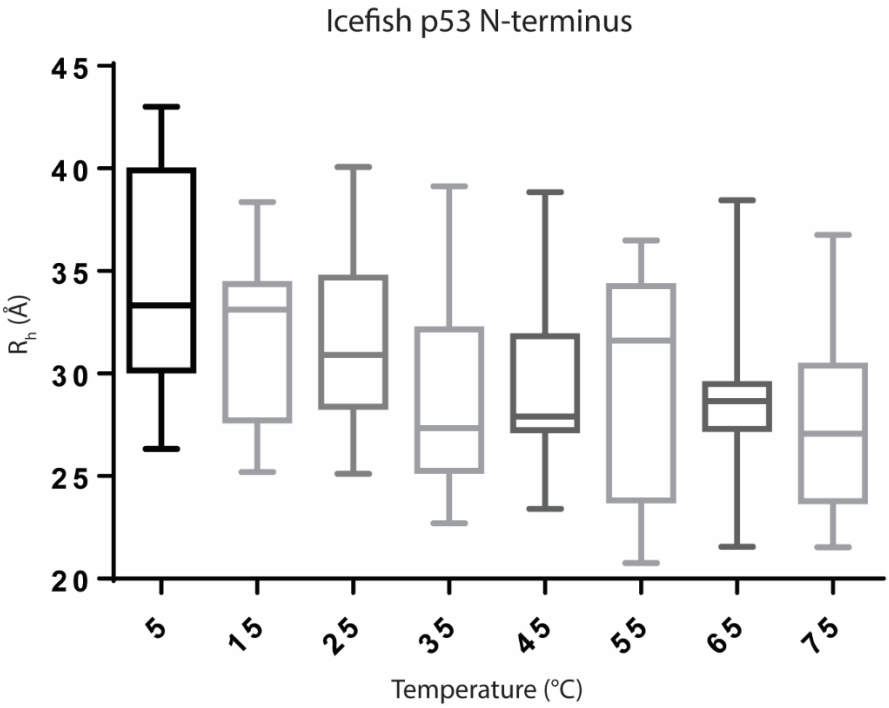

Figure 3.17 DLS Data for Icefish p53(1-89). Box and whisker plots were generated from DLS data (n=12) to visualize the change in R_h as a function of temperature.

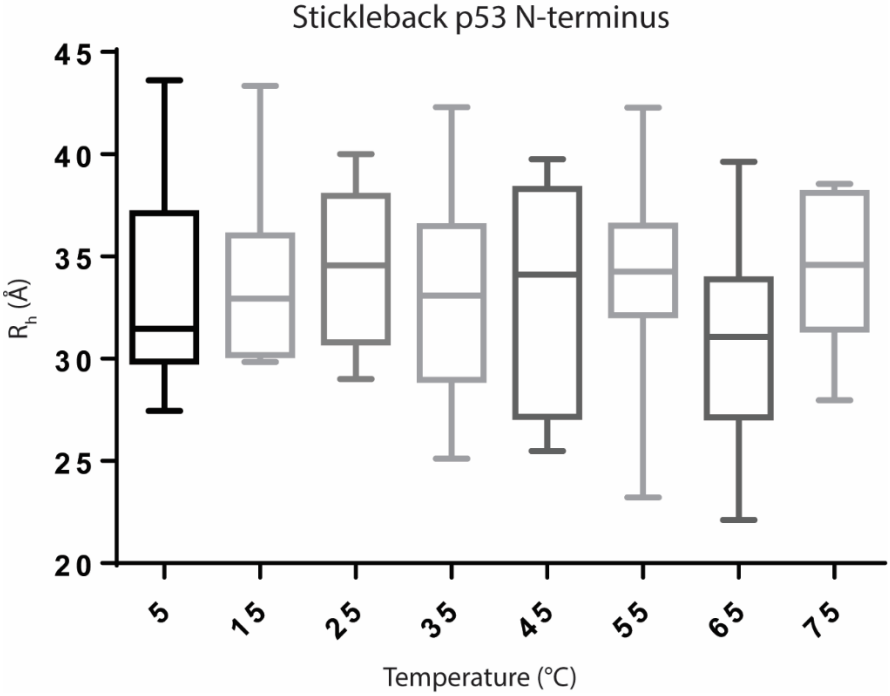

Figure 3.18 DLS Data for Stickleback p53(1-80). Box and whisker plots were generated from DLS data (n=12) to visualize the change in R_h as a function of temperature.

Table 3.3 Comparison of Measured R_h Values.

Protein	MW (Da)	$R_{h,\,SEC}$ (Å)	$R_{h,\,DLS}$ (Å)
Human p53(1-93)	10,123	32.5 ± 0.047	32.8 ± 1.2[46]
Whale p53(1-86)	9,413	30.6 ± 0.018	33.5 ± 0.82
Icefish p53(1-89)	10,256	29.8 ± 0.042	31.6 ± 1.4
Stickleback p53(1-80)	8,780	26.8 ± 0.011	34.5 ± 1.1

$R_{h,\,SEC}$ values were at 21 – 22 °C. $R_{h,\,DLS}$ values are averaged from the 25 °C measurements.

3.6 Discussion

Human p53(1-93) has a larger R_h than a folded protein of the same molecular weight[47]. This is evidenced by the way it migrated in SDS-PAGE; it electrophoresed at a higher molecular weight than its calculated molecular weight[45]. SEC and DLS are two techniques that were used to support the hypothesis that human p53(1-93) has a greater R_h than that of a traditional folded protein of the same molecular weight[47]. Similarly to human p53(1-93), whale, icefish, and stickleback p53 N-terminus homologs each also demonstrated electrophoretic mobility as would larger proteins, based on calculated molecular weight. Stickleback p53(1-80) had the largest difference between apparent molecular weight and calculated molecular weight in SDS-PAGE. This was surprising as the SEC results indicated that stickleback p53(1-80) had the smallest R_h. Since the identity of the homologs could not be determined by SDS-PAGE, ESI-MS was performed to confirm identity and to confirm that the 6x His tag had been removed. Although the measured molecular weights varied slightly from the calculated molecular weights, the data confirm that the proteins expressed and purified had consistent molecular weights with the target proteins.

With the fewest prolines of the four homologs, icefish p53(1-89) exhibited the least PPII structure and the least temperature dependence when measured by CD, as evidenced by the lower local maximum. Perhaps temperatures are required to be lower, at -2 °C, the environmental temperature of icefish, for icefish p53(1-89) to exhibit comparable PPII structure to that of whale p53(1-86). The reduced number of amino acids with PPII propensity may have adapted to have maximal PPII structure at its physiological temperature. Stickleback p53(1-80) had almost as much PPII structure as whale p53(1-86). Even with little homology and fewer proline residues, PPII structure

was conserved. Stickleback subspecies had evolved to thrive in a wide range of temperatures, perhaps stickleback p53 is able to retain PPII structure at warmer temperatures than icefish p53.

SEC demonstrated that human, whale, icefish, and stickleback p53 N-terminus homologs had a higher R_h than folded proteins of the same molecular weight. There appeared to be a correlation between decreased proline content and decreased R_h values using SEC, in order of decreasing proline content and R_h: human p53(1-93), whale p53(1-86), icefish p53(1-89), and stickleback p53(1-80). Though net charges were different amongst the homologs, the small differences did not seem to impact R_h or PPII content. In contrast to the manner in which it electrophoresed, stickleback p53(1-80) had the smallest R_h of the four proteins. The lower R_h was expected due to its smaller molecular weight and decreased proline content.

The K_d values of the folded protein standards varied slightly between the two SEC experiments, but had the same trend and linear relationship to R_h in the two experiments as well as in previous experiments[47]. Subtle changes in the column media, such as packing and flow rate, may have affected results between the two experiments. Within an individual experiment, standard deviations were low, indicating that there was good reproducibility within a given experiment, as long as the column flow was not disrupted.

One of the advantages of using DLS to measure R_h of proteins is that temperature dependence can be studied across a broad range of temperatures. This allows one to observe the temperature dependence of R_h. However, R_h measured using DLS at 25 °C had a much higher standard error than that measured using SEC. In contrast to the

traditional folded protein model, human p53(1-93) and other IDPs decrease in size, with smaller R_h values, as temperature increases[21,47]. Unexpectedly, the temperature dependence of whale p53(1-86) was not as prevalent as that with human p53(1-93), possibly indicating less disorder in whale p53(1-86). Icefish p53(1-89) demonstrated the greatest temperature dependence on R_h of the three homologs, similar to that of human p53(1-93), confirming that disorder is maintained even with the numerous sequence variations. Stickleback p53(1-80) appeared to have no temperature dependence on R_h, behaving more as a random coil than an IDP. However, there was a large discrepancy between measured R_h values from stickleback p53(1-80) DLS experiments compared to SEC experiments. With the variability present in DLS when measuring such relatively small R_h values, the data is probably not as reliable as with other methods of determining R_h, such as SEC. In future experiments, it may be possible to observe this temperature dependence of R_h using SEC at different temperatures in controlled surroundings.

3.7 Future Directions

The ability to perform SEC at a several different temperatures would provide a more accurate means to measure the temperature dependence of R_h in IDPs without relying so heavily on DLS. As stickleback p53(1-80) results varied from that of the other homologs, it would be interesting to compare the p53 N-termini of various subspecies of stickleback adapted to a variety of temperatures. Stickleback p53(1-80) contains PPII structure with only nine prolines; perhaps a study could be done using p53 N-termini sequences from nature with lower net negative charges to determine the effects of charge on PPII structure and temperature dependence. Contributing to negative charge are

phosphorylation events. Icefish p53(1-89) has more sites for phosphorylation than human p53(1-93) or whale p53(1-86). Studies on phosphorylated icefish p53(1-89) could be performed to determine if intrinsic disorder and PPII structure is increased upon phosphorylation.

REFERENCES

(1) Eliezer, D. *Curr. Opin. Struct. Biol.* **2009**, *19* (1), 23–30.

(2) Tompa, P. *Trends Biochem. Sci.* **2012**, *37* (12), 509–516.

(3) Dunker, A. K.; Obradovic, Z. *Nat. Biotechnol.* **2001**, *19* (9), 805–806.

(4) Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T. *J. Mol. Biol.* **2004**, *337* (3), 635–645.

(5) Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradović, Z.; Dunker, A. K. *J. Mol. Biol.* **2002**, *323* (3), 573–584.

(6) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215–246.

(7) Dunker, A. K.; Brown, C. J.; Obradovic, Z. *Adv. Protein Chem.* **2002**, *62*, 25–49.

(8) Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradović, Z. *Biochemistry (Mosc.)* **2002**, *41* (21), 6573–6582.

(9) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6* (3), 197–208.

(10) Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L. *Curr. Opin. Struct. Biol.* **2008**, *18* (6), 756–764.

(11) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins* **2000**, *41* (3), 415–427.

(12) Lise, S.; Jones, D. T. *Proteins* **2005**, *58* (1), 144–150.

(13) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (33), 14609–14614.

(14) Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z. *J. Mol. Graph. Model.* **2001**, *19* (1), 26–59.

(15) Dunker, A. K.; Garner, E.; Guilliot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J. E. *Pac. Symp. Biocomput. Pac. Symp. Biocomput.* **1998**, 473–484.

(16) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106* (5), 1877–1897.

(17) A. A. Adzhubei; Sternberg, M. J. *J. Mol. Biol.* **1993**, *229* (2), 472–493.

(18) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37* (4), 205–211.

(19) Rucker, A. L.; Creamer, T. P. *Protein Sci. Publ. Protein Soc.* **2002**, *11* (4), 980–985.

(20) Kjaergaard, M.; Nørholm, A.-B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. *Protein Sci. Publ. Protein Soc.* **2010**, *19* (8), 1555–1564.

(21) Langridge, T. D.; Tarver, M. J.; Whitten, S. T. *Proteins* **2014**, *82* (4), 668–678.

(22) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. *Nucleic Acids Res.* **2004**, *32* (3), 1037–1049.

(23) Hagai, T.; Azia, A.; Tóth-Petróczy, Á.; Levy, Y. *J. Mol. Biol.* **2011**, *412* (3), 319–324.

(24) Bell, S.; Klein, C.; Müller, L.; Hansen, S.; Buchner, J. *J. Mol. Biol.* **2002**, *322* (5), 917–927.

(25) Oren, M.; Rotter, V. *Cell. Mol. Life Sci. CMLS* **1999**, *55* (1), 9–11.

(26) Fields, S.; Jang, S. K. *Science* **1990**, *249* (4972), 1046–1049.

(27) Chang, J.; Kim, D. H.; Lee, S. W.; Choi, K. Y.; Sung, Y. C. *J. Biol. Chem.* **1995**, *270* (42), 25014–25019.

(28) Lee, H.; Mok, K. H.; Muhandiram, R.; Park, K. H.; Suk, J. E.; Kim, D. H.; Chang, J.; Sung, Y. C.; Choi, K. Y.; Han, K. H. *J. Biol. Chem.* **2000**, *275* (38), 29426–29432.

(29) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293* (2), 321–331.

(30) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274* (5289), 948–953.

(31) Levine, A. J.; Hu, W.; Feng, Z. *Cell Death Differ.* **2006**, *13* (6), 1027–1036.

(32) Oren, M. *J. Biol. Chem.* **1999**, *274* (51), 36031–36034.

(33) Walker, K. K.; Levine, A. J. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (26), 15335–15340.

(34) Lowry, D. F.; Stancik, A.; Shrestha, R. M.; Daughdrill, G. W. *Proteins* **2008**, *71* (2), 587–598.

(35) Wells, M.; Tidow, H.; Rutherford, T. J.; Markwick, P.; Jensen, M. R.; Mylonas, E.; Svergun, D. I.; Blackledge, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (15), 5762–5767.

(36) Sakamuro, D.; Sabbatini, P.; White, E.; Prendergast, G. C. *Oncogene* **1997**, *15* (8), 887–898.

(37) Oren, M. *Cell Death Differ.* **2003**, *10* (4), 431–442.

(38) Siliciano, J. D.; Canman, C. E.; Taya, Y.; Sakaguchi, K.; Appella, E.; Kastan, M. B. *Genes Dev.* **1997**, *11* (24), 3471–3481.

(39) Sakaguchi, K.; Herrera, J. E.; Saito, S.; Miki, T.; Bustin, M.; Vassilev, A.; Anderson, C. W.; Appella, E. *Genes Dev.* **1998**, *12* (18), 2831–2841.

(40) Shieh, S. Y.; Ikeda, M.; Taya, Y.; Prives, C. *Cell* **1997**, *91* (3), 325–334.

(41) Huber, R. *Nature* **1979**, *280* (5723), 538–539.

(42) Joerger, A. C.; Fersht, A. R. *Oncogene* **2007**, *26* (15), 2226–2242.

(43) Dawson, R.; Müller, L.; Dehner, A.; Klein, C.; Kessler, H.; Buchner, J. *J. Mol. Biol.* **2003**, *332* (5), 1131–1141.

(44) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6* (12), 1767–1770.

(45) Schaub, L. J.; Campbell, J. C.; Whitten, S. T. *Protein Sci. Publ. Protein Soc.* **2012**, *21* (11), 1682–1688.

(46) Tiffany, M. L.; Krimm, S. *Biopolymers* **1972**, *11* (11), 2309–2316.

(47) Perez, R. B.; Tischer, A.; Auton, M.; Whitten, S. T. *Proteins* **2014**, *82* (12), 3373–3384.

(48) Xue, B.; Brown, C. J.; Dunker, A. K.; Uversky, V. N. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2013**, *1834* (4), 725–738.

(49) Ruud, J. T. *Nature* **1954**, *173* (4410), 848–850.

(50) *Biology of the three-spined stickleback*; Östlund-Nilsson, S., Mayer, I., Huntingford, F., Eds.; Marine biology series; CRC Press: Boca Raton, 2007.

(51) Shortle, D.; Meeker, A. K. *Biochemistry (Mosc.)* **1989**, *28* (3), 936–944.

(52) Whitten, S. T.; García-Moreno E, B. *Biochemistry (Mosc.)* **2000**, *39* (46), 14292–14304.

(53) Uversky, V. N. *Protein J.* **2009**, *28* (7–8), 305–325.

(54) R. W. Woody. *Advances in biophysical chemistry: a research annual. Volume 2 Volume 2*; L. A. Bush, Ed.; Jai Press: Greenwich (Connecticut); London, 1992.