COLD ADAPTATION OF INTRINSICALLY DISORDERED STRUCTURE

STUDIED IN P53 HOMOLOGS

by

Erin C. Tilton, B.S.

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

Committee Members:

Steven T. Whitten, Chair

L. Kevin Lewis

Karen A. Lewis

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

I would like to thank my research advisor, Dr. Steven T. Whitten, for giving me the opportunity to work in his lab. I would also like to thank Dr. Karen A. Lewis and Dr. L. Kevin Lewis for taking the time to serve on my committee. Additionally, I would like to thank my lab colleagues, Lance R. English and Leona A. Martin, for their ideas, assistance, and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF EQUATIONS

Equation **Page**

CHAPTER 1

INTRODUCTION

1.1 Intrinsically disordered protein

Structural descriptions of functional biology, commonly known as the protein structure-function paradigm, state that protein function is dependent upon its well defined three-dimensional shape.[1] In contrast, a relatively new class of proteins described as intrinsically disordered provide evidence that a stable three-dimensional structure is not a prerequisite for protein function.

Intrinsically disordered proteins (IDPs) are biologically active proteins that do not form stable tertiary structures at physiological conditions.[2] They are common with 10-35% of prokaryotic and 15-45% of eukaryotic proteins considered disordered or containing disordered regions consisting of 30 contiguous residues or longer.[3,4] In humans, approximately a third of proteins are classified as intrinsically disordered and roughly 10% are fully disordered.[5] Typically, IDPs perform regulatory and signaling functions, including signal transduction, regulation of transcription[6] and translation,[7] self-assembly,[8] and post-translational modification.[9–11]

Examples of IDPs and their functions are now commonly found throughout the literature. For instance, the disordered oncoprotein E7 from human papilloma virus (HPV) interacts with cell regulation proteins within its host organism in order to hijack the host's replication machinery.[12,13] When E7 binds to transcription coactivators CBP/p300 and pRb, a retinoblastoma protein, acetylation of pRb is promoted. This signal cascade leads to disruption of the host's cell cycle control and enables cellular transformation by HPV.[14]

The regulatory mechanism of FIgM, a protein found in *Salmonella typhimurium*, operates as a result of its flexible structure.[2] *S. typhimurium* flagellum is constructed from a transmembrane basal body consisting of 35 different proteins.[2] The flexible structure of FIgM enables it to be exported through a small diameter channel within the basal body.[2,15] Any mutation resulting in a decrease in diameter causes FIgM levels to increase. Consequently, FIgM binds transcription factor sigma 28 (σ^{28}) leading to inhibition of mRNA synthesis and ultimately, assembly of flagellum.[15]

Due to their biological significance, it is important to elucidate IDP structure in order to better understand its role in cellular function and human disease.

1.2 Characterization of intrinsically disordered protein

In 1894, Emil Fischer proposed the first structural description of protein, known as the lock and key model, based on his observations of enzymes found in beer yeast.[16] He concluded hydrolysis occurs only when an enzyme and substrate fit together like a lock and key.[16] Forty years later, denaturation studies performed by Mirsky and Pauling supported conclusions made by Fischer.[17] When a protein's structure became denatured it was deemed no longer functional.[17] In the late 1950s, the first atomic resolution structures determined for myoglobin and lysozyme further supported the idea that all proteins have a stable tertiary structure.[18,19] However, in 1981 X-ray crystallography identified proteins with regions of missing electron density despite those regions being critical for function.[20,21] While it can be assumed those regions were likely disordered, such conclusions were not appropriate because user error, inability to solve the phase problem, or crystal defects could not be ruled out as factors.[22] Finally, nuclear magnetic resonance (NMR) revealed the functional, yet disordered tail of the histone H5 protein.[23] Since 2001, nearly 1,000 articles have been published pertaining to IDPs; a 500% increase in publications from 1995-2000.[3]

Structurally, IDPs lack a hydrophobic core and are highly flexible, displaying a wide range of conformational freedom.[24,25] Owing to their dynamic nature, traditional methods for characterization, such as x-ray crystallography and NMR, provide low structural resolution.[24] As a result, additional methods such as far-UV and near-UV circular dichroism (CD) spectroscopy,[26] fluorimetry,[27,28] and proteolysis studies[2] are being utilized to characterize disorder. Far-UV in combination with near-UV CD can detect disorder based on spectrum signatures of secondary structures.[2] Fluorescence studies

provide information on the presence of disorder by revealing the lack of a tightly packed, hydrophobic core.[27,28] The expanded nature of IDPs mean cleavage sites are more accessible for proteases, therefore proteolytic degradation studies are useful for identifying the presence of disorder.[29] In addition, characterization based on hydrodynamic parameters such as hydrodynamic radius, R_h, are being used to gain structural insight on intrinsic disorder. R_h has been widely used for quantifying protein structures for over 50 years[30,31] and can be measured using simple techniques based on light scattering,[25,32,33] size exclusion chromatography (SEC),[34,35] and sedimentation.[23,34]

1.3 Polyproline II contributes to intrinsically disordered structure

The polyproline II helix (P_{II}) is a non-classical secondary structure characterized by an extended left-handed helical conformation.[33,36] Many studies have demonstrated it is a preferred conformation in disordered proteins.[36,37] It consists of 3 residues per turn with a 3.1 Å rise and backbone dihedral angles of (-75°, +145°).[33] Due to its extended conformation and lack of intrachain hydrogen bonding, amino acid side chains are exposed to solvent.[36,38] While an ideal PP_{II} helix would be comprised of all prolines in the trans position, recent studies indicate all amino acids have a propensity to promote the PP_{II} helical structure.[33,38,39]

Creamer et al. were first to design a scale for predicting PP_{II} propensity using CD spectroscopy and a proline-based host peptide system, Ac-Pro_3-X-Pro_3-Gly-Tyr-NH_2, with X being any amino acid other than Gly, Tyr, and Trp. They measured PP_{II} helical content using the height of the maximum, $[\theta]_{max}$, from CD spectra.[33] As a result, Pro was identified as having the highest PP_{II} propensity (67%) with Val at the low end of the spectrum (49%).[33]

Kallenbach et al. used CD and J-coupling values determined from NMR to measure PP_{II} propensities for their host-peptide system, Ac-Gly-Gly-X-Gly-Gly-NH_2, excluding Gly and Pro.[38] They ranked PP_{II} propensity for each amino acid with Ala having the highest (81.8%) and His having the lowest (42.8%).[38]

In addition to the work performed by Creamer and Kallenbach groups, Hilser et al. developed a scale to predict PP_{II} propensity adding calorimetry as a component.[39] They calculated Pro as having the highest PP_{II} propensity (100%) while Gly had the lowest

(13%).[39]

By leveraging these propensity scales with experimental and molecular evaluations of R_h we hope to accurately predict structural characteristics of IDPs based on primary sequence. Using the intrinsically disordered N-terminal region of the human tumor suppressor protein p53 as an experimental model, preliminary studies have indicated a significant level of structural sensitivity to temperature[40] and properties that predict strong main chain bias for dihedral angles associated with secondary structure.[32,37]

If biological activity is closely aligned with structural features in IDPs, similar to the structure-function relationships observed in folded proteins,[41] then cold adaptation of IDP sequences would be expected as a compensatory strategy for the temperature-sensitive structural properties that have been documented in our model IDP[37,40,42] as well as other IDP systems.[43–45] To investigate this hypothesis, the structural properties of p53 homologs from species that live at cold temperatures were studied, specifically targeting the intrinsically disordered N-terminal region that regulates p53 activity.[28,46] This region contains an acidic transactivation domain (1-61) and a proline-rich domain (62-93) and contains seven phosphorylation sites.[47] By studying homologs from cold-adapted species, the effects of sequence variations can be observed using sequences that have been naturally selected to preserve activity. A homolog from beluga whale, *Delphinapterus leucas*, which lives in Artic conditions but maintains a constant body temperature similar to humans was used as a negative control. In addition, a homolog from European salmon, *Coregonus lavaretus,* was studied since this fish is cold blooded and lives at a cold temperature. The body temperature of beluga whales is approximately 37°C. In contrast, European salmon live at ~10°C.

1.4 Goal of research

The purpose of this research was to investigate differences in structural properties among human, whale, and salmon p53 N-terminal homologs. To accomplish this, R_h was determined using size exclusion chromatography (SEC) and dynamic light scattering (DLS) techniques. Temperature effects on R_h were measured using DLS. Temperature effects on secondary structure content were evaluated using circular dichroism spectroscopy (CD). The present study has allowed us to collect preliminary structural data that will provide the necessary groundwork for further investigations, such as how these structural changes affect DNA binding and phosphorylation of amino acids within disordered regions. By investigating these structural properties, more can be understood about the influence of sequence changes and temperature and the structure-function relationship of IDPs.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Chemicals and reagents used throughout this research were ACS grade or higher. In addition, water used in sample and buffer preparations was subjected to filtration and deionization by a Milli-Q Integral 3 pure water system from EMD Millipore (Darmstadt, Germany). Glassware, pipette tips, and other basic lab equipment were sterilized with a HICLAVE HV-50 autoclave vessel (Hirayama, Tokyo, Japan).

Incubation of agar plates at 37°C was carried out using a VWR forced air microbiological incubator (Radnor, PA). Incubation of broth cultures were carried out using the MaxQ 5000 incubated floor shaker from Thermo Scientific (Waltham, MA). Centrifugation steps were performed using Thermo Scientific's Sorvall LYNX centrifuge (Waltham, MA).

2.2 Sequences

2.2.1 Sequences of p53 homologs

Amino acid sequences were taken from wild-type consensus of the N-terminal region of p53 from human, beluga whale, and salmon. The gene sequences contained amino acid positions 1-93 for human (see Figure 2.1), 1-86 for whale (see Figure 2.2), and 1-83 for salmon (see Figure 2.3). Amino acid sequences were obtained from GenBank: p53(1-93) - BAC16799.1, p53(1-86) - AAL83290.1, and p53(1-83) - ACH73252.1.

```
1  MEEPQSDPSV  EPPLSQETFS  DLWKLLPENN  VLSPLPSQAM
41 DDLMLSPDDI  EQWFTEDPGP  DEAPRMPEAA  PPVAPAPAAP
81 TPAAPAPAPS  WPL
```

Figure 2.1. Amino acid sequence of human p53(1-93)
The N-terminal region consists of an acidic transcriptional activation domain (1-61; blue) and a proline-rich domain (62-93; red). Overall, p53(1-93) has two positively charged amino acids and seventeen negatively charged amino acids resulting in a net charge of -15 at physiological conditions. Additionally, 22 of 93 amino acids are proline (23.2%).

```
1  MEESQAELGV  EPPLSQETFS  DLWKLLPENN  LLSSELSPAV
41 DDLLLSPEDV  ANWLDERPDE  APQMPEPPAP  AAPTPAAPAP
81 ATSWPL
```

Figure 2.2. Amino acid sequence of whale p53(1-86)
p53(1-86) has two positively charged amino acids and seventeen negatively charged amino acids resulting in a net charge of -15 at physiological conditions. Additionally, 16 out of 86 amino acids are proline (18.6%).

```
1  MADLVENVSL   PLSQESFEDL   WKMNLNLMEV   QPPVTEAWVE
41 YDNFMMEAPL   QGEFDQSLFE   VSAPQPSIST   LDTGSPPTST
81 VPT
```

Figure 2.3. Amino acid sequence of salmon p53(1-83)
p53(1-83) has one positively charged amino acid and fourteen negatively charged amino acids resulting in a net charge of -13 at physiological conditions. Additionally, 9 out of 83 amino acids are proline (10.8%).

2.2.2 Sequence alignment of p53 homologs

Sequence alignment was performed using Clustal Omega to compare amino acid composition between human p53(1-93) and cold-adapted p53 homologs: whale(1-86) and salmon(1-83). Sequences were entered as FASTA sequences obtained from GenBank as mentioned in 2.2.

```
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP   60
MEESQAELGVEPPLSQETFSDLWKLLPENNLLSSELSPAVDDLLLSPEDVANWLDER--P   58
*** *:: .**********************:**   * *:***:***:*: :*: *   *

DEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPL                               93
DEAPQMPE-----PPAPAAPTPAAPAPATSWPL                               86
****:***      ************** ****
```

Figure 2.4. Sequence alignment between human p53(1-93) and whale p53(1-86)
The first line of each set (blue) is the amino acid sequence from human p53(1-93) and the second line (orange) is from whale p53(1-86). An asterisk (*) indicates positions that have a single, fully conserved residue. A colon (:) indicates conservation of amino acids with strongly similar properties. A period (.) indicates conservation of amino acids with weakly similar properties. Similarity of amino acid properties was based on scoring from Gonnet PAM 250 matrix. A score > 0.5 indicate strong similarity and a score ≤ 0.5 indicates weakly similar properties.

```
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP   60
--MADLVENVSLPLSQESFEDLWKMNLNLMEV-------------QPPVTEAWVEYDNFM   45
    :   .*. *****:*.****:  :   :              .*   * *.  *


DEAPRMPEAAPRVAPAPAAPTPAAPAPAPSWPL                               93
MEAPLQGEFDQSL-FEVSAPQPSISTLDTGSPPTSTVPT                         83
 ***   *    :    :** *:  :   . *
```

Figure 2.5. Sequence alignment between human p53(1-93) and salmon p53(1-83)
The first line of each set (blue) is the amino acid sequence from human p53(1-93) and the second line (green) is from salmon p53(1-83). An asterisk (*) indicates positions that have a single, fully conserved residue. A colon (:) indicates conservation of amino acids with strongly similar properties. A period (.) indicates conservation of amino acids with weakly similar properties. Similarity of amino acid properties was based on scoring from Gonnet PAM 250 matrix. A score > 0.5 indicate strong similarity and a score ≤ 0.5 indicates weakly similar properties.

2.3 Expression and purification of recombinant p53

The following methods describe protein expression, purification, and characterization protocols for p53(1-93). Each step was performed identically for whale p53(1-86) and salmon p53(1-83).

2.3.1 Cloning and transformation

Gene sequences were cloned onto a pJ404 plasmid expression vector under the control of a T5 promoter for isopropyl β-D-1-thiogalactopyranoside (IPTG) induced expression (DNA 2.0, Newark, CA). The plasmids also included a gene for ampicillin resistance (*amp*^{r}) and a 6x Histidine-tag (His-tag). Histidine has a high affinity for the nickel media utilized in nickel affinity chromatography.[48] Transformation was accomplished using *Escherichia coli* (*E. coli*) BL21 (DE3) pLysS competent cells (Novagen, Madison, WI).

2.3.2 Expression of recombinant p53 using Escherichia coli

Transformed cells were plated onto LB agar containing 100 μg/mL of ampicillin (Teknova, Hollister, CA) and incubated overnight at 37°C. The following day, a single colony was transferred to 10 mL of LB broth containing 100 μg/mL of ampicillin and incubated overnight at 30°C with orbital rotation at 170 RPM. The inoculated broth was sub-cultured into 1 L of LB broth containing 100 μg/mL of ampicillin at a final OD_{600} of 0.01. The subculture was incubated at 37°C with orbital rotation at 170 RPM until an OD_{600} of 0.6-0.7 was reached. IPTG was added at a final concentration of 0.5 mM to induce expression. Expression proceeded for 4 hours at 37°C with orbital rotation at 170

RPM. Finally, cells were collected by centrifugation at 30,200 x *g* for 15 minutes at 4°C. The supernatant was discarded and pellets were frozen overnight at -80°C.

2.3.3 Nickel affinity chromatography

Cell lysis was carried out by resuspending frozen pellets into 20 mL of lysis buffer (6 M guanidine hydrochloride, 10 mM Tris-HCl, 100 mM sodium phosphate, pH 8.0). Cells were then sonicated on ice using a Branson Sonifier S-450A (Emerson Electric, Danbury, CT) set to 80% duty and an output control of 5 for 4 cycles of 90 seconds on and 90 seconds off. After sonication, cells were centrifuged at 33,746 x *g* for 1 hour at 4°C. Following centrifugation, the supernatant was loaded into a 1.5 x 15 cm glass Econo-column (Bio-Rad, Hercules, CA) containing His-Select Nickel Affinity Gel (Sigma-Aldrich, St. Louis, MO) that had been previously equilibrated with 30 mL of lysis buffer. In order to observe the elution of protein, the column was connected to a Biologic LP system equipped with an ultraviolet (UV) absorbance monitor (Bio-Rad, Hercules, CA). Absorbance was measured at λ= 280 nm. After the addition of the centrifuged sample, an initial peak was observed (see Figures 2.6-2.8). This peak is known as "flow through" and contains proteins with no affinity for the nickel media. Once the peak from flow through returned to baseline an initial wash step (6 M guanidine hydrochloride, 10 mM Tris-HCl, 100 mM sodium phosphate, pH 8.0) was used to remove non-specific and weakly bound proteins. The wash buffer was added until the absorbance returned to baseline. Approximately, 30 mL of a low imidazole buffer (10 mM Tris-HCl, 100 mM sodium phosphate, 10 mM imidazole, pH 8.0) was added to remove proteins with a slight affinity for nickel. Finally, elution of target protein was achieved with ~30 mL of high

imidazole buffer (10 mM Tris-HCl, 100 mM sodium phosphate, 350 mM imidazole, pH 4.3). To remove imidazole, the eluate was dialyzed overnight at 4°C against 4 L of a Tris-buffered saline (20 mM Tris-HCl, 100 mM NaCl, pH 8.0) using Spectra/Por tubing with a 12-14 kDa molecular weight cut off (Spectrum Labs, Rancho Dominguez, CA).

2.3.4 Nickel affinity chromatograms

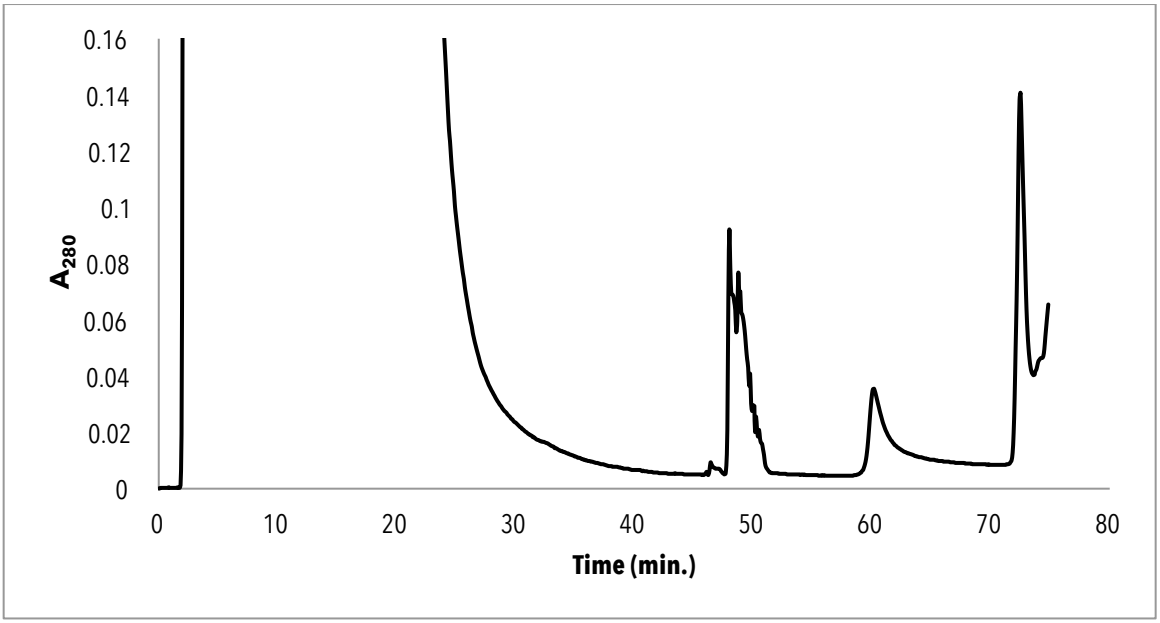

Figure 2.6. Chromatogram from nickel affinity purification of human p53(1-93) This chromatogram illustrates change in absorbance over time for the purification of recombinant human p53(1-93) from cell lysate. The first peak contained proteins lacking affinity for nickel. The second peak (~48 min) identified removal of non-specific and weakly bound proteins after the addition of an initial wash step. Next, a low imidazole step was used to remove proteins with a slight affinity for nickel (~61 min). The target protein contains a 6x His-tag allowing it to bind strongly with the nickel media. This required a high imidazole step for protein elution (~73 min). The last peak around ~75 min. was imidazole eluting from the column.

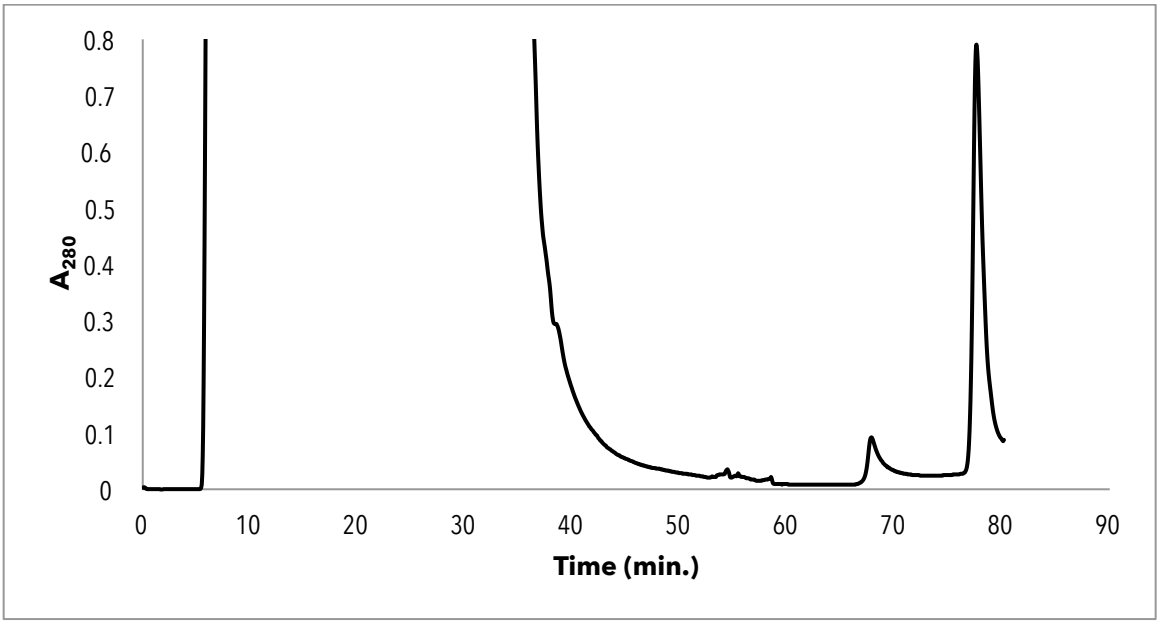

Figure 2.7. Chromatogram from nickel affinity purification of whale p53(1-86) This chromatogram illustrates change in absorbance versus time for the purification of recombinant whale p53(1-86) from cell lysate. The first peak contains proteins lacking affinity for nickel. The second peak (~55 min) identified removal of non-specific and weakly bound proteins after the addition of an initial wash step. Next, a low imidazole step was used to remove proteins with a slight affinity for nickel (~69 min). The target protein contains a 6x His-tag allowing it to bind strongly with the nickel media. This required a high imidazole step for protein elution (~78 min).

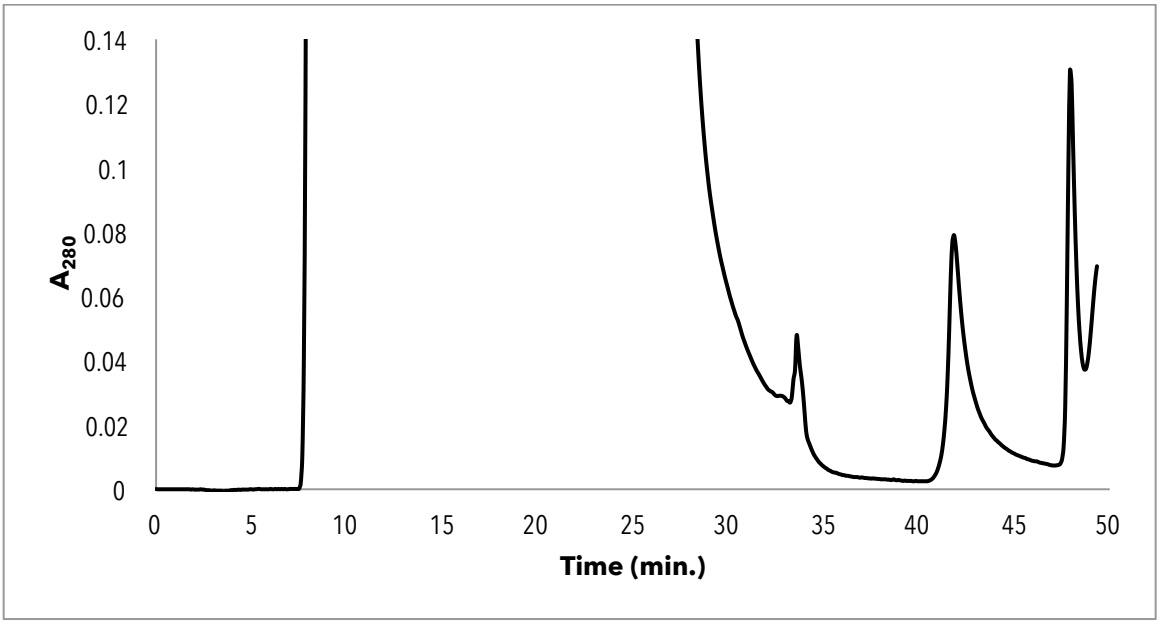

Figure 2.8. Chromatogram from nickel affinity purification of salmon p53(1-83) This chromatogram illustrates change in absorbance versus time for the purification of recombinant salmon p53(1-83) from cell lysate. The first peak contained proteins lacking affinity for nickel. The second peak (~34 min) identified removal of non-specific and weakly bound proteins after the addition of an initial wash step. Next, a low imidazole step was used to remove proteins with a slight affinity for nickel (~43 min). The target protein contains a 6x His-tag allowing it to bind strongly with the nickel media. This required a high imidazole step for protein elution (~48 min). The last peak around ~75 min. was imidazole eluting from the column.

2.3.5 Anion exchange chromatography

After overnight dialysis, the sample was subjected to thrombin digestion to remove the 6x His-tag. Fifty Units of thrombin (Sigma-Aldrich, St. Louis, MO) were added to the dialyzed sample and digestion was carried out in the dark for 4 hours at 25°C. The digested sample was then loaded into a 1.5 x 15 cm glass Econo-column (Bio-Rad, Hercules, CA) containing diethylaminoethyl (DE52) anion exchange media (Whatman, Maidstone, UK) and connected to a Biologic LP system with UV absorbance monitoring capability (Bio-Rad, Hercules, CA). Absorbance was measured at λ= 280 nm. Prior to sample loading, the DE52 media was degassed for 30 minutes and equilibrated with 40 mL of a saline/acetate buffer (20 mM sodium acetate, 25 mM NaCl, pH 4.8). To remove contaminating proteins, the column was washed with ~40 mL of the same saline/acetate buffer containing increasing concentrations of sodium chloride: 50 mM, 100 mM, and 150 mM (see Figures 2.9-2.11). Each concentration had a pH of 4.8. Elution of target protein was achieved using a high salt buffer (20 mM sodium acetate, 250 mM sodium chloride, pH 4.8). The eluate was dialyzed overnight at 4°C against 4 L of a phosphate-buffered saline (10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0) using Spectra/Por tubing mentioned previously.

Following dialysis, protein concentration was determined by measuring absorbance in a DU730 UV/Vis spectrophotometer (Beckman Coulter, Brea, CA) at λ=280 nm and solving for concentration, c, using the Beer-Lambert Law (see equation 1). According to ProtParam, the N-terminal region of p53 from human and whale both have an extinction coefficient of 16,500 $M^{-1}\,cm^{-1}$ and p53(1-83) from salmon has an extinction coefficient of 12,490 $M^{-1}\,cm^{-1}$.

$$A = \varepsilon cl \quad (1)$$

where A is absorbance at λ=280 nm, ε is extinction coefficient, c is concentration (mol/L), and l is path-length (cm).

2.3.6 Anion exchange chromatograms

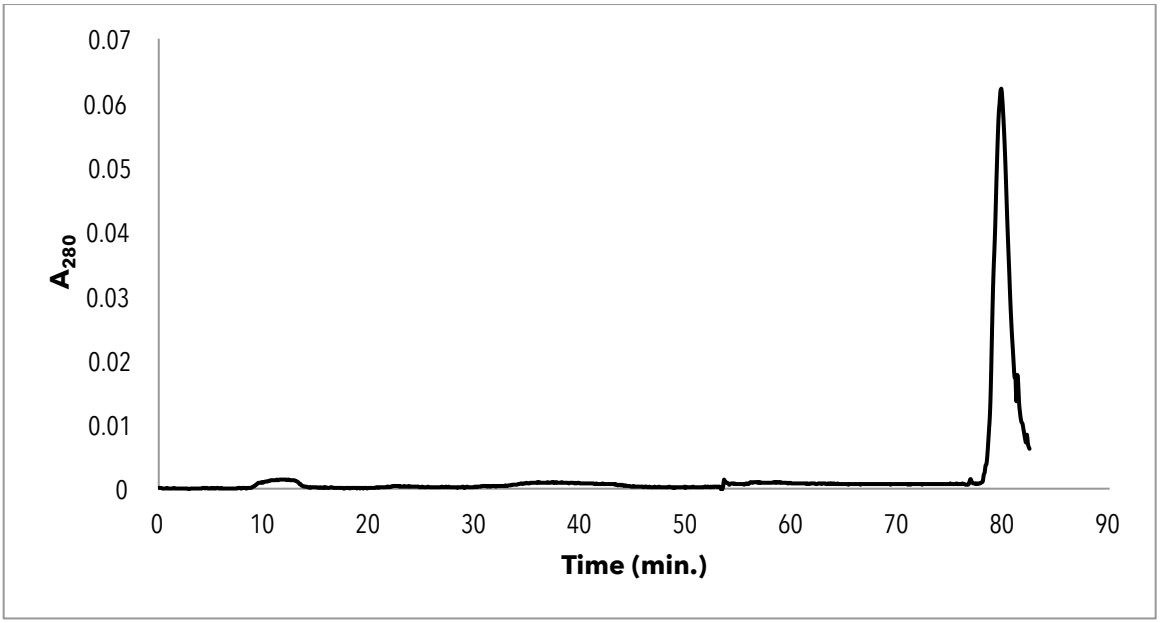

Figure 2.9. Chromatogram from anion exchange purification of human p53(1-93) This chromatogram illustrates change in absorbance versus time for human p53(1-93) after the addition of increasing concentrations of a sodium acetate/chloride buffer. After nickel affinity chromatography and thrombin digestion of the 6x His-tag, the sample was loaded onto a DE52 column. Proteins unable to bind to a positively charged column were observed as flow through (~10 min.). Wash buffers with low to medium salt concentrations (50-150 mM) were used to elute weakly bound proteins (~22, 40, 55 min., respectively). Finally, a high salt buffer was used to elute the target protein, p53(1-93) (~80 min).

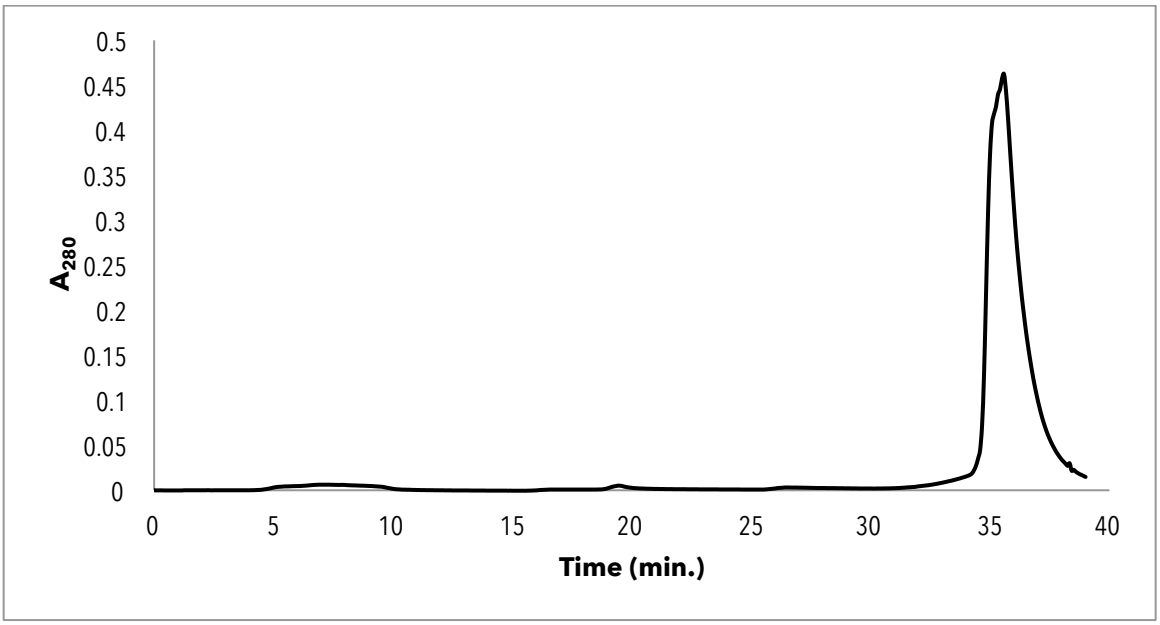

Figure 2.10. Chromatogram from anion exchange purification of whale p53(1-86) This chromatogram illustrates change in absorbance versus time for whale p53(1-86) after the addition of increasing concentrations of a sodium acetate/chloride buffer. After nickel affinity chromatography and thrombin digestion of the 6x His-tag, the sample was loaded onto a DE52 column. Proteins unable to bind to a positively charged column were observed as flow through (~7 min.). Wash buffers with low to medium salt concentrations (50-150 mM) were used to elute weakly bound proteins (~15, 20, 25 min., respectively). Finally, a high salt buffer was used to elute the target protein, p53(1-86) (~36 min).

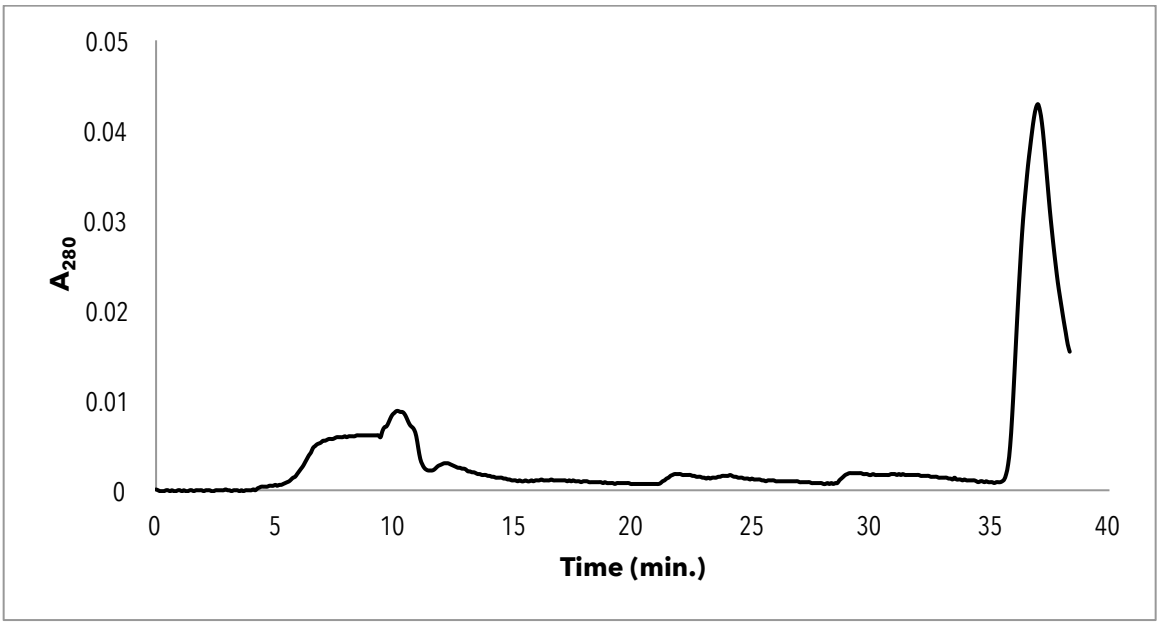

Figure 2.11. Chromatogram from anion exchange purification of salmon p53(1-83) This chromatogram illustrates change in absorbance versus time for salmon p53(1-83) after the addition of increasing concentrations of a sodium acetate/chloride buffer. After nickel affinity chromatography and thrombin digestion of the 6x His-tag, the sample was loaded onto a DE52 column. Proteins unable to bind to a positively charged column were observed as flow through (~7 min.). Wash buffers with low to medium salt concentrations (50-150 mM) were used to elute weakly bound proteins (~12, 23, 29 min., respectively). Finally, a high salt buffer was used to elute the target protein, p53(1-83) (~37 min).

2.3.7 Assessing purity of protein using gel electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to assess purity of collected protein. Samples were mixed at a ratio of 1:1 with a 2x Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS, 0.01% bromophenol blue) (Bio-Rad, Hercules, CA) containing 5% β-mercaptoethanol (J.T. Baker, Center Valley, PA) and then heated at 95°C for 5 minutes. Samples were loaded into a 4-20% precast Tris-HCl polyacrylamide gel (Bio-Rad, Hercules, CA) and electrophoresed within a Criterion cell (Bio-Rad, Hercules, CA) with a 1X Tris/glycine running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.3). Electrophoresis was carried out for 50 minutes at 200 V (see Figures 2.12-2.13).

2.3.8 Silver nitrate staining of polyacrylamide gels

To visualize protein bands, gels were silver stained using a method previously described by Chevallet et al.[49] Briefly, gels were fixed in 30% ethanol, 10% acetic acid for 30 minutes. Next, gels were rinsed twice in 20% ethanol for 10 minutes each. Third, gels were sensitized in 0.02% (w/v) sodium thiosulfate for one minute then rinsed in water two times for 1 minute each. Next, gels were impregnated in 12 mM silver nitrate for 20 minutes then rinsed with water for 10 seconds. Gels were then submerged in a basic developer solution (3% potassium carbonate (w/v), 3 mM formalin, 79 μM sodium thiosulfate) for approximately 5-10 minutes or until desired intensity of bands was achieved. Lastly, gels were rinsed in a stop solution (330 mM Tris-base, 2% (w/v) acetic acid) for 30 minutes followed by two rinses with water for 30 minutes each. Each step

was performed at room temperature using gentle agitation on an orbital bench-top shaker (Thermo Scientific, Waltham, MA).

Stained gels were imaged with a Bio-Rad Molecular Imager ChemiDoc XRS+ imaging system.

2.3.9 Gel images

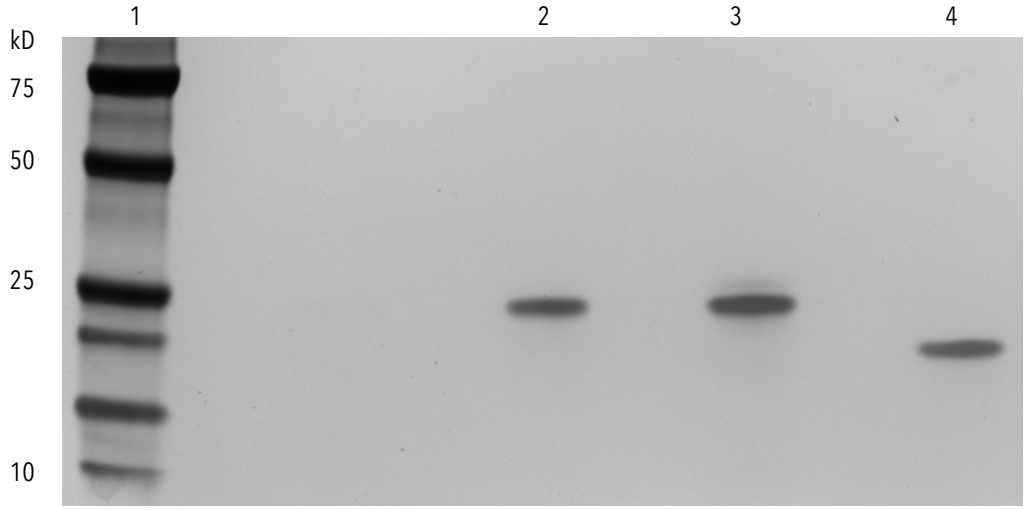

Figure 2.12. Silver stained SDS-PAGE gel for human p53(1-93) and whale p53(1-86)
Protein purity was assessed by gel electrophoresis and silver stained as mentioned in section 2.3.8. Lane 1 is the Precision Plus Protein ladder from Bio-Rad. 75, 50, 25, and 10 kDa bands are labeled for reference. Lanes 2 and 3 are single bands of human p53(1-93) after nickel and anion exchange chromatography from two separate purifications. Lane 4 is whale p53(1-86) after nickel and anion exchange chromatography. Human p53(1-93) has a molecular weight of 10.1 kDa and an apparent molecular weight of ~22 kDa. A single band around ~20-25 kDa indicates pure protein. Whale p53(1-86) has a molecular weight of 9.3 kDa and an apparent molecular weight ~ 20 kDa.

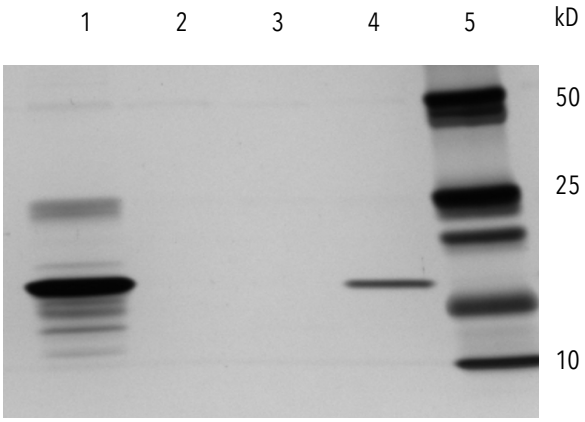

Figure 2.13. Silver stained SDS-PAGE gel for salmon p53(1-83)
Protein purity was assessed by gel electrophoresis and silver stained as mentioned in Section 2.3.8. Lane 5 is the Precision Plus Protein ladder from Bio-Rad. 50, 25, and 10 kDa bands are labeled for reference. Lane 1 is eluted protein from nickel affinity chromatography. Lane 4 is salmon p53(1-83) after anion exchange chromatography. Salmon p53(1-83) has a molecular weight of ~9.3 kDa, but an apparent molecular weight of ~17 kDa.

2.4 Detection of secondary structure content by circular dichroism spectroscopy

Once purity of protein was determined by SDS-PAGE, CD spectra were recorded in order to determine secondary structure content, specifically relative propensities of PP_{II} for p53 homologs: human, whale, and salmon.

First, protein was loaded into a 1 mm path-length quartz cuvette at a final concentration of 0.1-0.2 mg/mL. The cuvette was inserted into a Jasco J-710 spectrophotometer equipped with a PFD-425S Peltier unit (Easton, MD). Since far-UV light interacts with oxygen to form ozone (O_3), the spectrophotometer was flushed with nitrogen gas for 15 minutes to provide an oxygen-free environment.[50] After 15 minutes, the xenon lamp was turned on and the optical housing was continually purged with nitrogen gas at a flow rate of 5 L/min. Temperature scans from 5-85°C with 10° intervals were conducted. Samples were equilibrated at each temperature for 10 minutes. Spectra were collected with a resolution of 0.5 nm, a scan rate of 20 nm/min, were the average of 8 scans, and were baseline corrected by blanking with 10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0. CD measurements were converted into normalized mean molar ellipicity (θ_{MRE}) using the following equation,

$$\theta_{MRE} = \theta_d \frac{M}{c \times l \times n} = \left(\frac{deg \times cm^2}{dmol \times residue}\right) \quad (2)$$

where M is the molecular weight of the N-terminus of p53, c is the protein concentration, l is the path length of the cuvette, and n is the number of residues.

2.5 Determination of and temperature effects on hydrodynamic radius using dynamic light scattering

DLS is a well-established technique used to determine size and size distribution of molecules in solution.[51] This method is also applied to macromolecular size determinations (i.e., protein) by converting the diffusion coefficient, D, to an apparent hydrodynamic radius, R_h, using the Stokes-Einstein equation,[37,40]

$$R_h = \mathrm{kT}/(6\pi\eta\mathrm{D}) \quad (3)$$

where k is the Boltzman constant, T is absolute temperature, and η is solvent viscosity.

Samples to be measured were filtered through a 0.2 μm PVDF syringe filter and loaded into a 1 cm path-length quartz cuvette at a final concentration of 0.4-0.5 mg/mL. The cuvette was inserted into a Zetasizer Nano ZS with Peltier temperature control (Malvern Instruments, Malvern, UK). Temperature scans from 5°-75°C with 10° intervals were performed. Samples were equilibrated at each temperature for 15 minutes. R_h was calculated from an average of 12 or more runs per temperature scan and temperature scans cycled through at least 3 times. Solvent viscosity was calculated using a solvent builder software program provided by Malvern.

2.6 Determination of hydrodynamic radius using size exclusion chromatography

SEC was used to determine the R_h for N-terminal p53 homologs by comparing partition coefficients, K_D, to values from folded proteins with well-defined, globular structures. Staphylococcal nuclease, bovine erythrocyte carbonic anhydrase, chicken egg albumin, and horse heart myoglobin were chosen as protein standards since they have published crystallographic structures and cover a wide R_h range (see Table 1).

2.6.1 Preparation of protein standards

Staphylococcal nuclease was purified using methods described elsewhere.[42,52] The remaining proteins were purchased from Sigma-Aldrich (St. Louis, MO) and further purified by ion exchange chromatography.[53]

2.6.2 Preparation of column media

SEC experiments were performed using 3.75 g Sephadex G-75 media (GE Healthcare, Pittsburgh, PA) equilibrated in 10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0 and degassed for 30 minutes. The SEC media was loaded into a 1.5 x 30 cm glass Econo-column from Bio-Rad and packed by gravity. The packed column was attached to a Biologic LP system equipped with an UV absorbance monitor (Bio-Rad, Hercules, CA) (see Figure 2.13). A constant flow rate was maintained throughout the procedure.

2.6.3 Preparation of protein sample

Prior to being loaded onto the column, each protein sample was mixed with indicator dyes, blue dextran and 2,4-dinitrophenyl-L-aspartate (DNP-sspartate). Each sample contained 0.2–0.5 mg/mL protein in 10 mM sodium phosphate, 100 mM sodium

chloride, pH 7 with 0.3 mg/mL blue dextran and 0.03 mg/mL DNP-aspartate. All steps were carried out at room temperature (~20°C).

2.6.4 Calculating partition coefficients

Partition coefficients, K_D, were determined using the equation below,

$$K_D = (V_e - V_o) / (V_t - V_o) \quad (4)$$

where V_e is elution volume, V_o is void volume, and V_t is total volume. See schematic in Figure 2.8 for example.

Table 1. Properties of protein standards

Protein	mw (Da)	R_h (Å)
Chicken egg albumin	44,287	35.76
Bovine carbonic anhydrase	29,844	27.34
Staphylococcal nuclease	16,900	21.20
Horse myoglobin	16,950	21.83

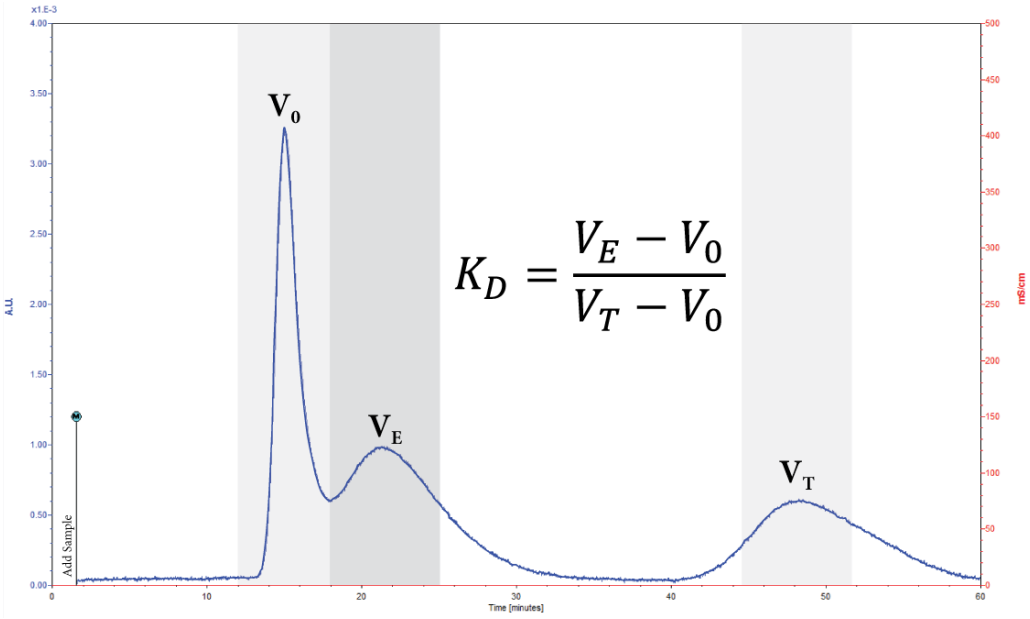

Figure 2.14. Schematic for determination of K_D using a chromatogram from SEC[54]
The partition coefficient, K_D, is determined using the equation above. The first peak is used to determine the void volume, V_o. The second peak is used to determine the elution volume, V_e. The third peak is used to determine the total volume, V_t. SEC is performed using a constant flow rate so time is substituted for volume.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Introduction

This research seeks to gain insight into temperature effects on structural properties of intrinsically disordered proteins using R_h as a reporter of size. Recent studies indicate IDPs have a surprising level of structural sensitivity to temperature.[40,42,44]

Kjaergaard et al. investigated temperature effects on 3 IDPs using CD spectroscopy, small angle X-ray scattering (SAXS), and NMR chemical shift analysis. Their CD data suggests increased temperature causes a heat-induced unfolding of PP_{II} structure.[44] Additionally, a decrease in radius of gyration, R_g, was observed using SAXS indicating IDP ensembles become more compact at elevated temperatures.[44] Compaction had initially been attributed to formation of α-helices,[55–57] but Kjaergaard et al. provided evidence against this as their results indicated a loss of helicity at elevated temperatures.[44]

Langridge et al. investigated temperature effects on the N-terminal region of human p53 protein using CD spectroscopy, DLS, and SDS-PAGE. Their results suggest temperature modulates a structural equilibrium between a preferred PP_{II} structure at low temperatures and a heat-induced denatured state at high temperatures.[40,42]

To further investigate temperature effects, we quantified structural characteristics of p53 homologs from human, whale, and salmon. By studying homologs from cold adapted species, the effects of sequence variations can be observed using sequences that have been naturally selected to preserve activity. To accomplish this, R_h was determined using SEC and dynamic light scattering (DLS) techniques. Temperature effects on R_h

were measured using DLS. Temperature effects on secondary structure content were evaluated using CD.

3.2 Temperature effects on secondary structure content

CD spectroscopy is a technique used to measure the difference in absorption of left-handed and right-handed circularly polarized light from chiral chromophores across a range of wavelengths. Chiral centers can be found in 19 out of 20 naturally occurring amino acids making CD a valuable tool for studying biological molecules, particularly proteins.[58] The secondary structure of proteins can be differentiated by spectrum signatures specific for alpha helices, beta-pleated sheets, PP_{II} helices, and random coils.[59]

A main structural component of IDPs is the presence of PP_{II} content, which can be identified by a peak between 220-225 nm. To observe temperature effects on PP_{II} structure, CD spectra for p53 homologs were recorded from 5°-85°C with 10° intervals. These results are provided in Figure 3.1. A local maximum was observed at 221 nm for each homolog verifying the presence of PP_{II} structure.[60–62] A local minimum at 200 nm and an isochromatic point at 210 nm were also observed.

When comparing spectra in Figure 3.1, human p53(1-93) has the most pronounced peak around 221 nm and the largest range of MRE decreasing from 0 to -25,000 deg cm^2 $dmol^{-1}$ res^{-1}. The spectra for whale appear similar to human having the same decrease in MRE, but a slightly less prominent peak at 221 nm. The spectra for salmon appear to flatten compared to human and whale.

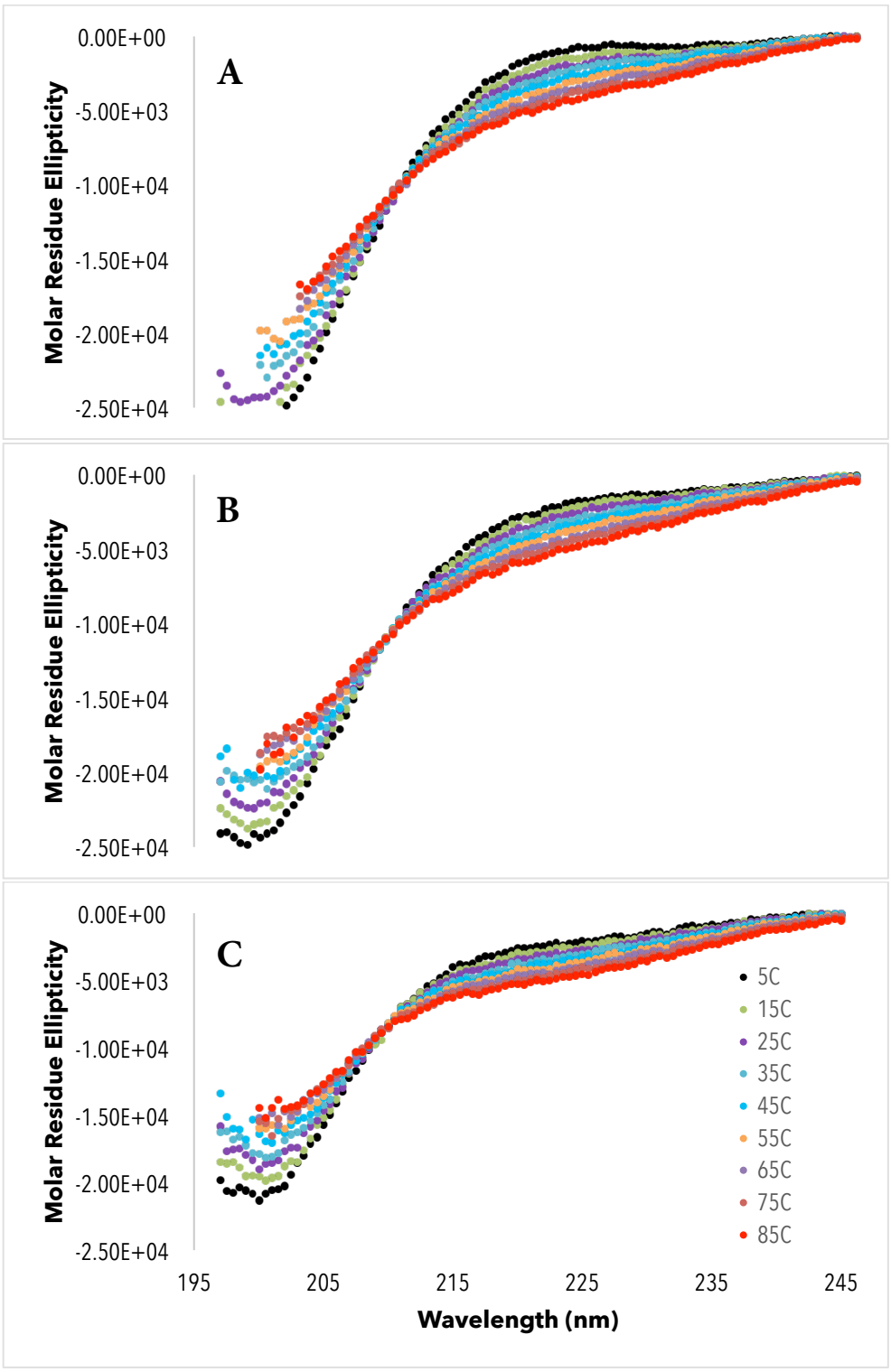

Figure 3.1. Temperature dependent CD spectra for N-terminal p53 homologs
CD spectra measured for A) human p53(1-93) B) whale p53(1-93) C) salmon p53(1-83). Measurements were taken at 197-240 nm from 5-85°C. Results are reported in molar residue ellipticity (MRE) units of deg cm^2 $dmol^{-1}$ res^{-1}.

Figure 3.2 provides a more discernible comparison of temperature effects on PP_{II}. In this figure molar residue ellipticity (MRE) is plotted relative to the 85°C spectrum. As temperature increased, the local maximum decreased for each homolog. Stated more generally, as temperature is increasing the PP_{II} structure is disappearing, or becoming thermally "unfolded."

Proline has been shown to have a high propensity for PP_{II} structure so we hypothesized this would be apparent when comparing peak height. Each homolog has varying fractional proline content: human p53(1-93) contains 22 prolines (23%), whale p53(1-86) contains 16 (18%), and salmon p53(1-83) contains 9 (11%).[33,38] Peak heights at 5° are: ~3,500 for human p53(1-93), ~3,000 for whale p53(1-86), and ~2,500 for salmon p53(1-83), which roughly correlate with proline content (see Figure 3.2).

Additionally, if you plot the difference of average MRE values (from 220-222 nm) relative to the 85°C spectrum, a linear temperature dependence is present (see Figure 3.3). This suggests each homolog lacks tertiary contacts that would otherwise cause an abrupt, cooperative unfolding event as seen in proteins with stable tertiary structures.[40]

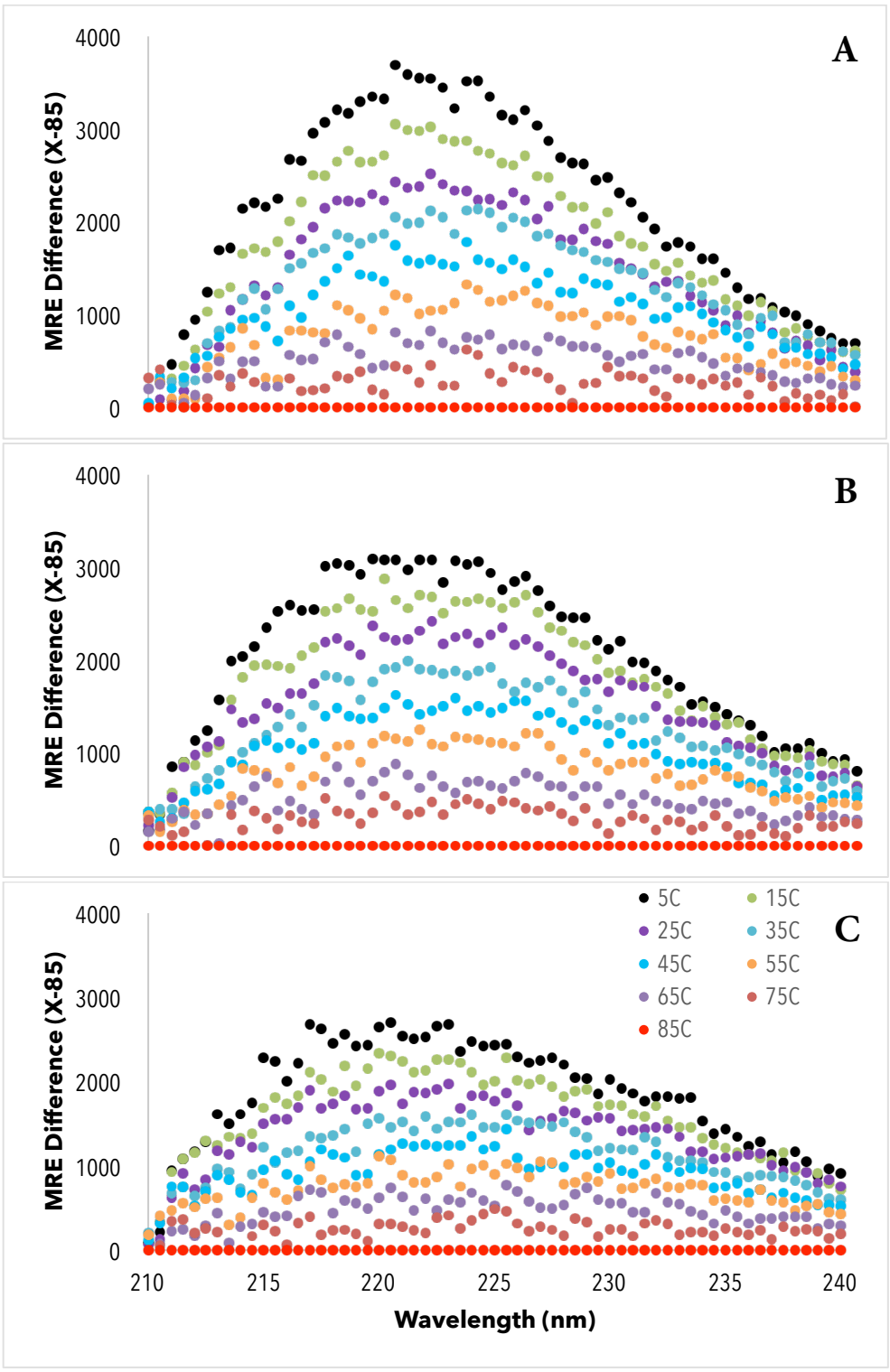

Figure 3.2. Temperature dependent CD spectra of local maxima for p53 homologs
Temperature-dependent CD spectra of local maxima for A) human p53(1-93) B) whale p53(1-86) C) salmon p53(1-83). The difference in molar residue ellipticity (MRE) at each temperature relative to the 85°C spectrum is illustrated. Results are reported in deg cm^2 $dmol^{-1}$ res^{-1}.

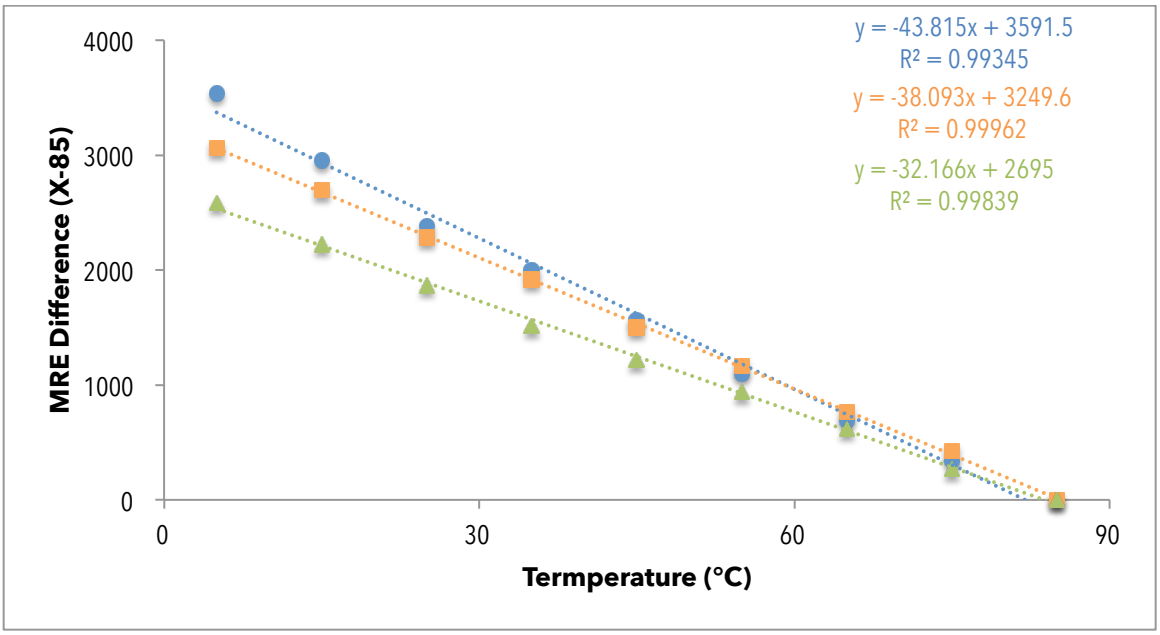

Figure 3.3. Summary of CD results for p53 homologs

Temperature-dependent CD spectra of local maxima at 221 nm calculated by taking the average MRE value from 220-222 nm, then subtracting the 85°C value from each reading. Blue circles represent human p53(1-93). Orange squares represent whale p53(1-86). Green triangles represent salmon p53(1-83). Results are reported in deg cm^2 $dmol^{-1}$ res^{-1}.

3.3 Determination of and temperature effects on hydrodynamic radius using dynamic light scattering

DLS is a spectroscopic technique that uses noninvasive backscatter to determine the size of molecules based on how they behave in solution, also known as Brownian motion.[63] To quantify the structural size of p53 homologs, the diffusion coefficient, D, was measured using DLS and converted to R_h using the Stokes-Einstein equation (see Equation 3). Results for each homolog are found in Figure 3.4.

Immediately apparent for human and whale homologs is a temperature-dependent trend where R_h decreases as temperature is increased. The R_h of p53(1-93) decreased from ~35 Å at 5°C to ~30 Å at 55°C, which is a 14% decrease in size. The R_h of whale decreased from ~31 Å at 5°C to ~ 25 Å at 55°C, an overall decrease of 19% in size. This temperature-dependent trend matches what was previously published for human p53(1-93).[37,40,42] Whale p53(1-86) was expected to behave similarly to human p53(1-93) since both species regulate their body temperature and have similar proline content.

Additionally, the decrease in R_h occurred gradually over a broad temperature range. This gradual compaction also matches results previously obtained for human p53(1-93) and is indicative of proteins absent of tertiary contacts.[40] In contrast, folded proteins with stable tertiary structures display cooperative thermal unfolding resulting in an abrupt change in R_h.[40,42]

The R_h of salmon decreased from ~30 Å at 5°C to ~24 Å at 55°C, which is a decrease in size of 20%. However, salmon does not exhibit the same temperature effects on R_h observed in human and whale, as there was no statistically significant trend in R_h versus temperature.

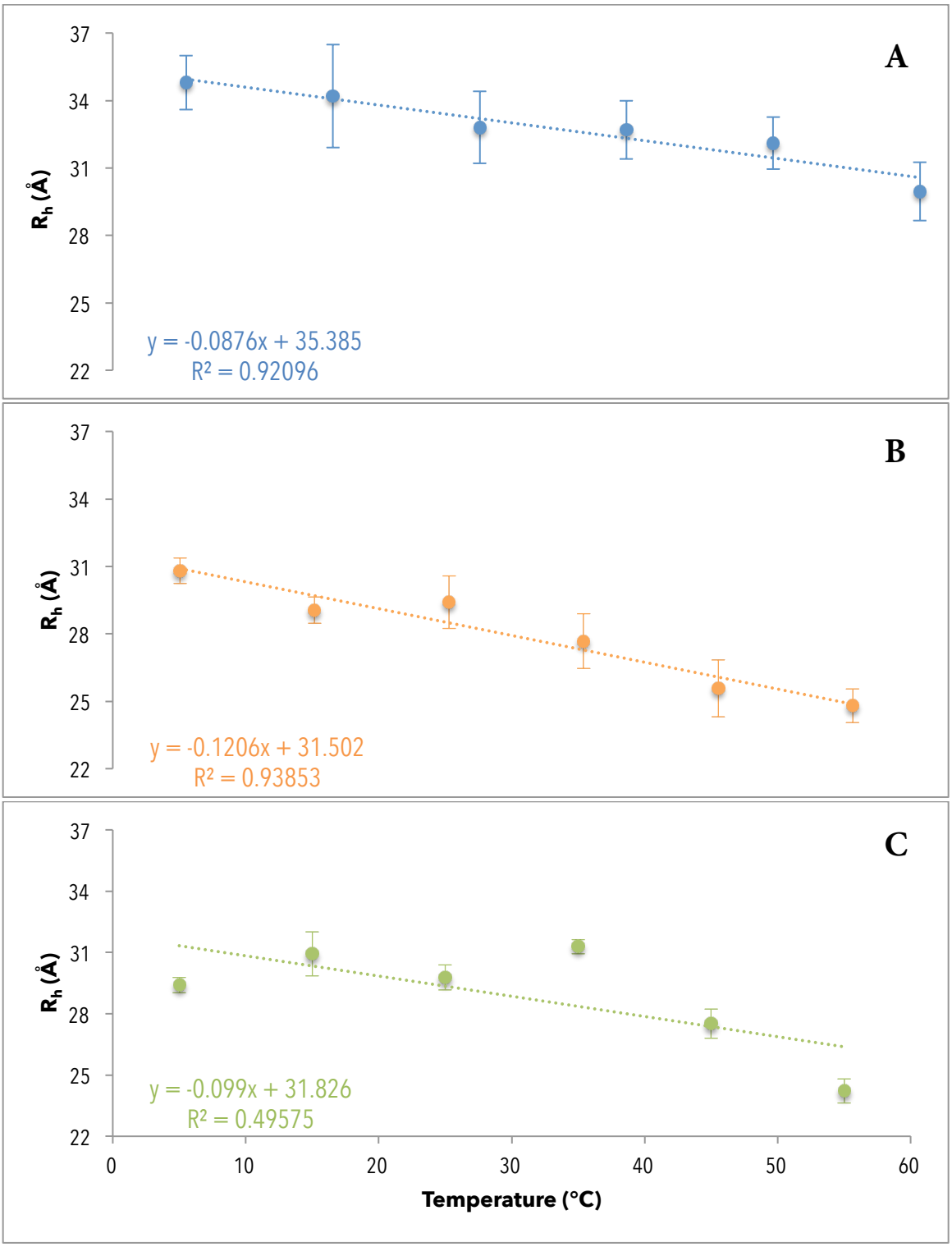

Figure 3.4. Temperature dependence of hydrodynamic radius for p53 homologs
DLS measurements were performed using ~0.4 mg/mL protein in 10 mM sodium phosphate, 100 mM sodium chloride, pH 7.0. Values are given as an average of at least 12 scans per temperature. A) Human p53(1-93) B) whale p53(1-86) C) salmon p53(1-83).

3.4 Determination of hydrodynamic radius using size exclusion chromatography

To confirm R_h determined by DLS, R_h was measured using SEC for p53 homologs and a set of protein standards. SEC is a chromatography method used to separate proteins based on size and/or molecular weight. A partition coefficient, K_D, can be calculated by comparing a protein's mobility through column media against two reference molecules, blue dextran and DNP-aspartate. Blue dextran has a large molecular weight (~2,000 kDa) and is used to determine void volume, V_o. DNP-aspartate has a small molecular weight (~300 Da) and is used to determine total volume, V_t.

We have previously shown a linear correlation can be made between K_D and R_h for proteins with known molecular weights and resolved crystal structures.[37] R_h is calculated from crystal structures as one-half the maximum distance between any two C_α atoms. Using this correlation, R_h for p53 can be extrapolated from the regression line generated by the protein standards.

In Tables 2-3 plus Figures 3.5 A and 3.6 A, elution volumes were correlated to molecular weight using equation 4. K_Ds for a set of protein standards were plotted against their R_h previously determined from their resolved crystallographic structures.[64–67] K_Ds are not comparable between different column preparations so protein standards were run each time SEC was performed. Since SEC separates proteins based on size (i.e. R_h), K_D for protein standards trend with their R_h. This correlation was used to extrapolate R_h for p53 homologs (see Figures 3.5 B-3.6 B). At 25°C, R_h calculated for human p53(1-93) was 31.22 Å. R_h determined for whale p53(1-86) and salmon p53(1-83) were estimated to be

29.72 Å and 28.28 Å, respectively. See Table 4 for comparison of SEC and DLS measured R_h.

During SEC, each homolog demonstrated slow column mobility resulting in an apparent molecular weight much larger than its actual molecular weight (see Figures 3.5 A and 3.6A). This was also observed during gel electrophoresis. Due to the expanded structure of IDPs their apparent molecular weight is often much larger than their actual molecular weight. For instance, in Figure 2.12 from section 2.3.9, when running human and whale p53 on a gel, their apparent molecular weights were ~25 and ~20 kDa; however, their actual molecular weights are 10.1 and 9.3 kDa, respectively. This is why it is necessary to use the correlation between R_h and K_D from protein standards to extrapolate R_h for p53 homologs.

Table 2. Summary of SEC results for human p53(1-93) and salmon p53(1-83)

Protein	$K_{D,mean}$	$K_{D,stdev}$	mw (Da)	Log mw	R_h (Å)
Albumin	0.111	0.011	44,287	4.646	35.76
CA	0.238	0.003	29,844	4.475	27.34
SNase	0.362	0.012	16,900	4.228	21.20
Myoglobin	0.393	0.002	16,950	4.229	21.83
p53(1-93)	0.185	0.002	10,108	4.005	31.22
p53(1-83)	0.242	0.015	9,393	3.973	28.28

CA is carbonic anhydrase, SNase is Staphylococcal nuclease, and mw is molecular weight. Table 2 is plotted in Figure 3.5. N=4.

Table 3. Summary of SEC results for whale p53(1-86)

Protein	$K_{D,mean}$	$K_{D,stdev}$	mw (Da)	Log mw	R_h (Å)
Albumin	0.089	0.009	44,287	4.646	35.76
CA	0.273	0.007	29,844	4.475	27.34
SNase	0.385	0.022	16,900	4.228	21.20
Myoglobin	0.406	0.004	16,950	4.229	21.83
p53(1-86)	0.219	0.003	9,398	3.973	29.72

CA is carbonic anhydrase, SNase is Staphylococcal nuclease, and mw is molecular weight. Table 3 is plotted in Figure 3.6. N=4.

Table 4. R_h values measured by SEC and DLS.

Protein	$K_{D,mean}$	$K_{D,stdev}$	$R_{h,SEC}$	$R_{h,DLS}$
Human p53(1-93)	0.089	0.009	31.22 ± 0.052	32.80 ± 1.202
Whale p53(1-86)	0.273	0.007	30.15 ± 0.0015	29.40 ± 1.166
Salmon p53(1-83)	0.385	0.022	28.28 ± 0.773	29.78 ± 0.617

R_h values were measured by SEC and DLS at 25°C and reported as Å. Standard error was determined for $R_{h,SEC}$ and $R_{h,DLS}$. N=4 and N=12, respectively.

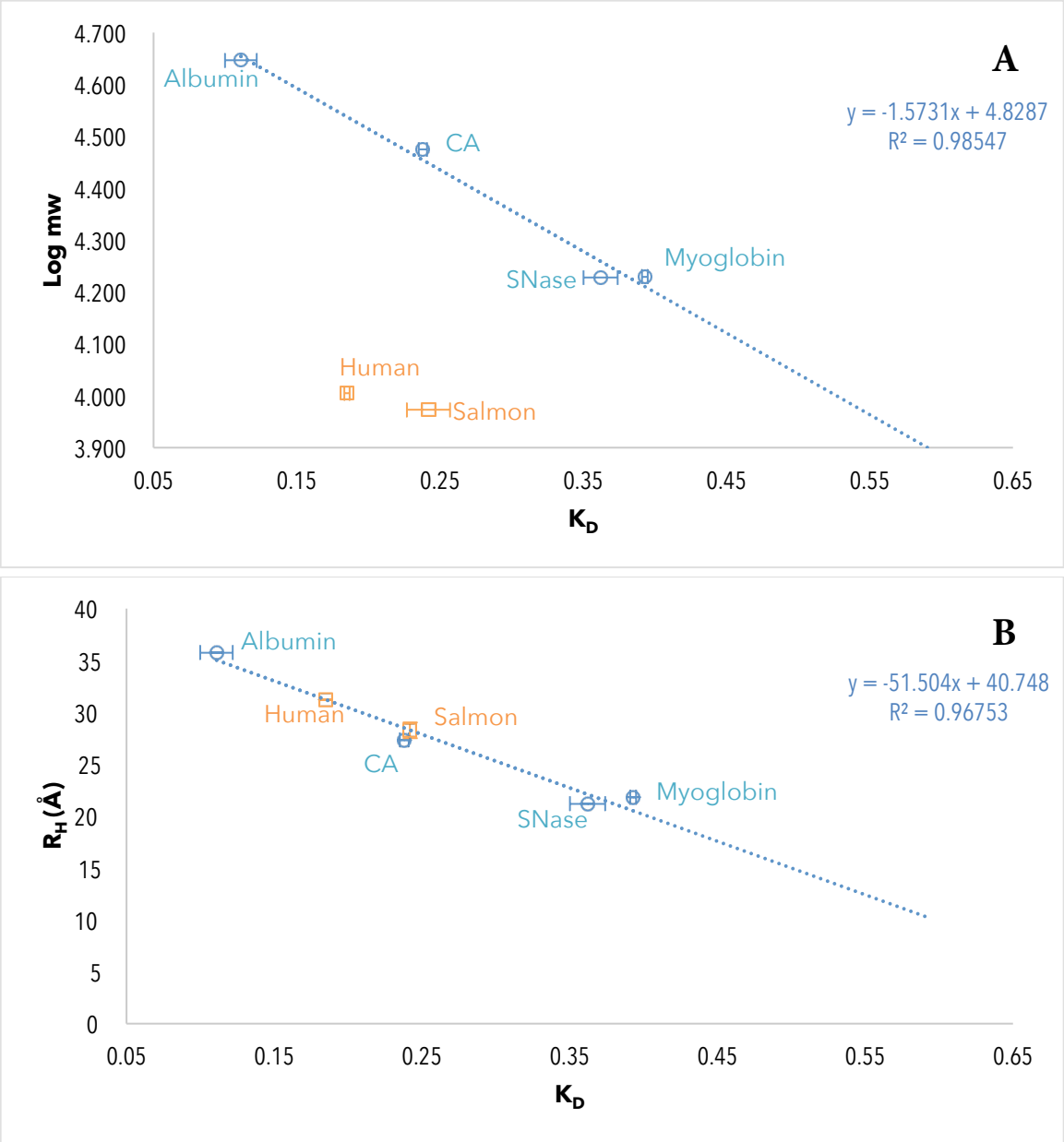

Figure 3.5. K_D measured by SEC for human p53(1-93), salmon p53(1-83), and protein standards

K_D was measured for protein standards (blue circles), human p53(1-93), and salmon p53(1-83) (orange squares). Proteins were run at room temperature and reported values are the average of four trials. Error bars represent standard error with N=4. A) p53(1-93) and p53(1-83) have molecular weights (mw) of 10.1 and 9.3 kDa, respectively. However their K_D values suggest a larger apparent mw (~31 and 27.5 kDa, respectively). B) Blue circles are R_h estimated from crystallographic structures of proteins standards. The dashed line is a linear fit of R_h versus K_D for protein standards, which was used to estimate R_h for p53(1-93) and p53(1-83).

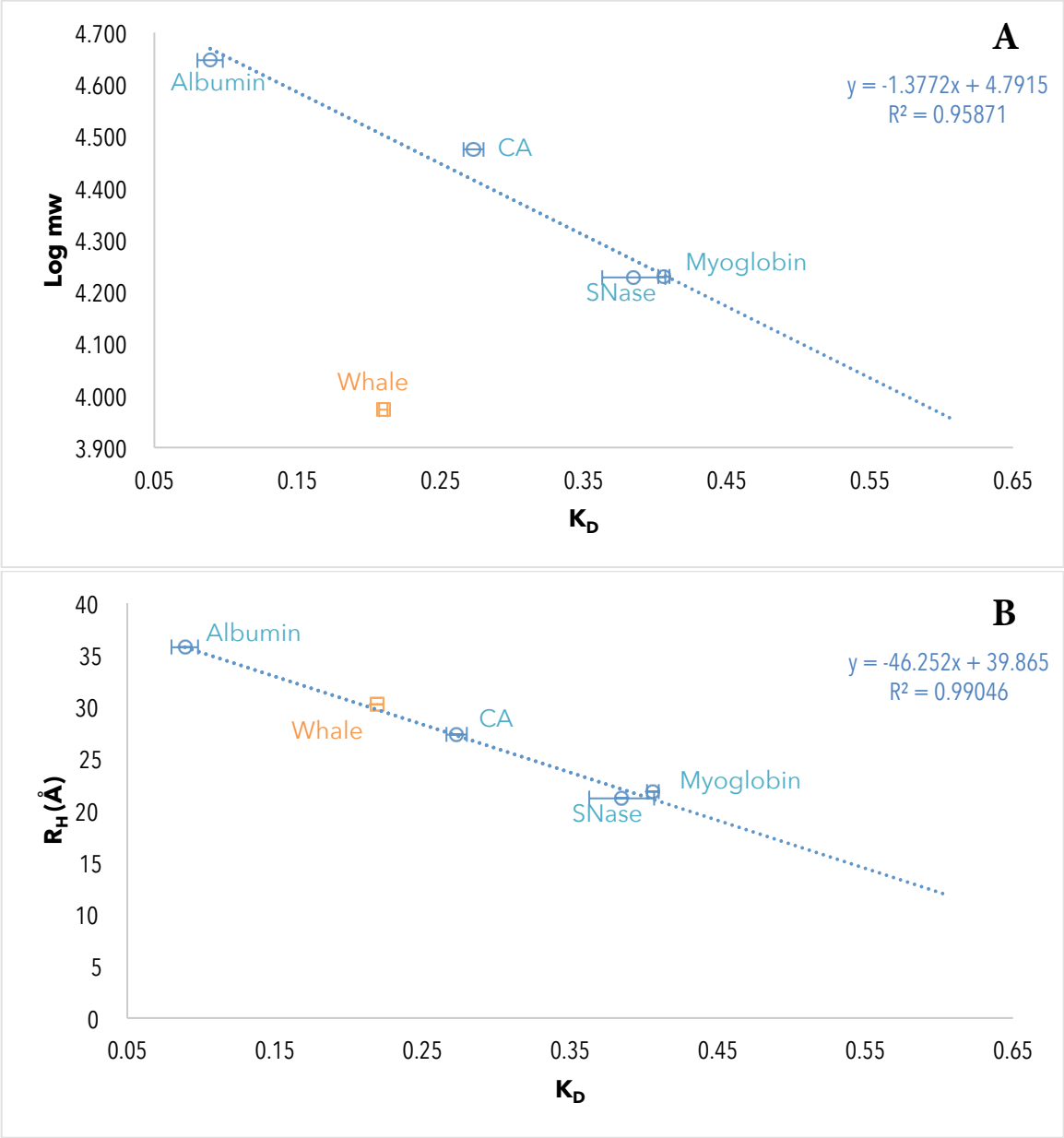

Figure 3.6. K_D measured by SEC for whale p53(1-86) and protein standards
K_D was measured for protein standards (blue circles) and whale p53(1-86) (orange square). Proteins were run at room temperature and reported values are the average of four trials. Error bars represent standard error with N=4. A) p53(1-86) has a molecular weight (mw) of ~9.3 kDa. However its K_D value suggests a larger apparent mw (~30 kDa). B) Blue circles are R_h estimated from crystallographic structures of proteins standards. The dashed line is a linear fit of R_h versus K_D for protein standards, which was used to estimate R_h for p53(1-86).

3.5 Conclusions

These results indicate that p53 homologs from cold-adapted species exhibit structural properties similar to results previously published for human p53(1-93).[40,42] Each homolog contains PP_{II} content, PP_{II} content trends with fractional proline content, and PP_{II} structure decreases with elevated temperature.

Human p53(1-93) and whale p53(1-86) exhibit similar temperature dependence as R_h has a strong, negative, linear correlation with temperature for both species. However, and perhaps most interestingly, this same level of structural temperature dependence was not seen in European salmon (Figure 3.4). This disagrees with our initial hypothesis that a temperature-dependent R_h is a defining characteristic of IDPs. Since conducting my experiments on cold-adapted species, other members of our lab have since observed a lack of structural temperature dependence in other IDPs (unpublished data from Lance R. English). The common factor with salmon p53(1-83) and these new IDPs that are being characterized is a significant decrease in fractional proline content. This leads us to our new hypothesis: temperature-dependent R_h is a defining characteristic of proteins with high fractional proline content.

A possible explanation for this phenomenon lies in the fact that European salmon cannot regulate its own body temperature.[68] European salmon are viable and able to grow at a temperature range of 6.0°C to 29.5°C.[68] Previous studies have shown that structural changes in disordered regions caused by post-translational modifications allow a protein to be recognized by other proteins, but cannot be bound by those same proteins when not modified.[69] This implies that certain structural characteristics are required for function, even in disordered regions. The lack of PP_{II} content in salmon p53(1-83) gives

it a relatively flat R_h (Figure 3.7) across its temperature viability range. If salmon p53(1-83) had the same PP_{II} content as its whale or human counterparts, the hydrodynamic properties would undergo massive changes in response to changing water temperature. Perhaps evolution has selected for a primary sequence in European salmon p53 that allows for a consistent size over the widest temperature range.

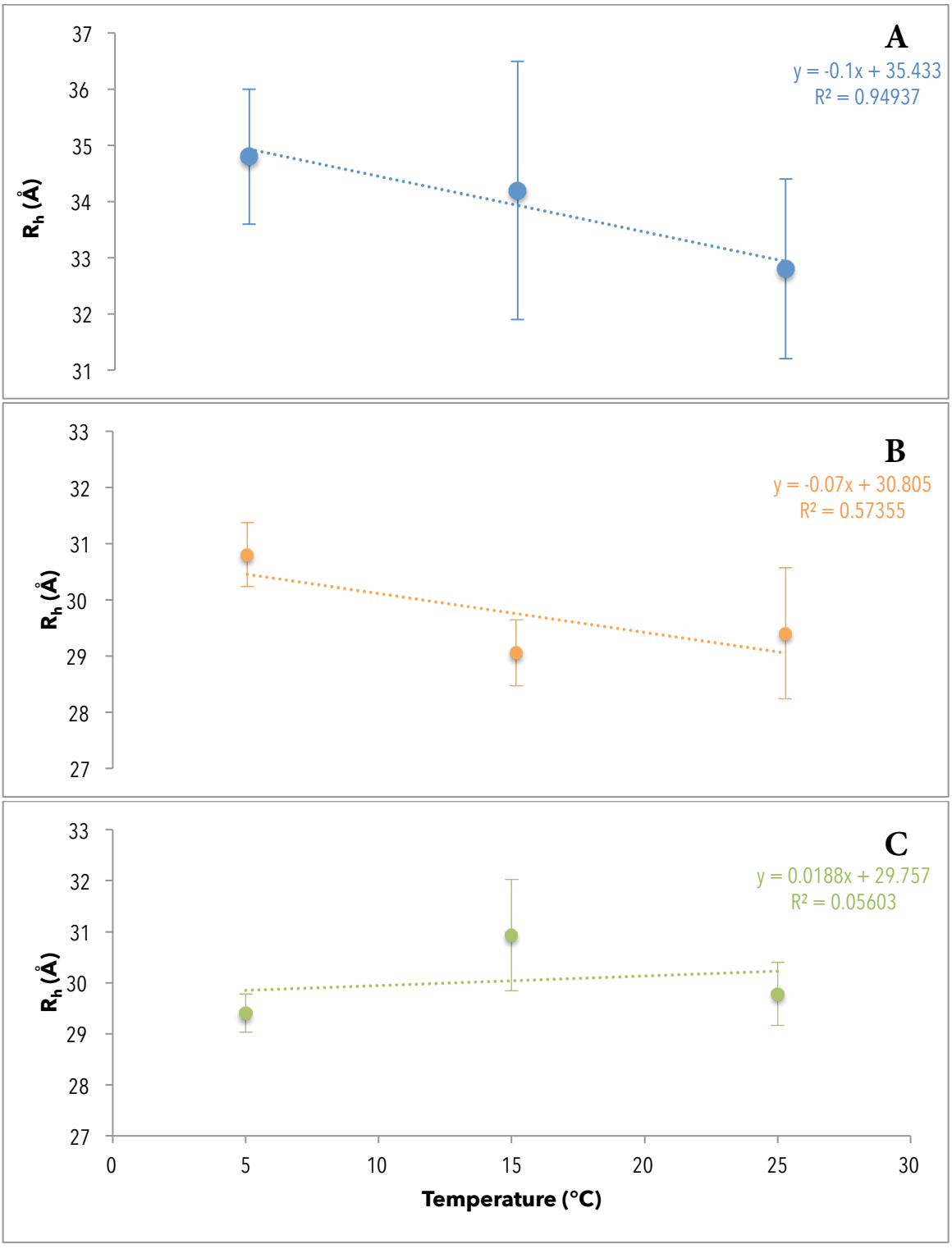

Figure 3.7. R_h measured by DLS across temperatures 5-25°C
This figure illustrates DLS measured R_h across 5-25°C. A) Human p53(1-93) B) whale p53(1-86) C) salmon p53(1-83). Error bars represent standard error, N=9.

This hypothesis raises the question regarding the usefulness of prolines. Why have the whale and human homologs evolved to have high proline content? Considering both organisms regulate their own body temperature[70,71] and have narrow viability ranges,[72] the temperature-dependent R_h probably does not offer any advantages. One explanation is that the disordered N-terminal region of p53, across all species, is required to be roughly 60 Å in diameter (R_h = 30 Å) spread across ~90 residues. As temperature increases, a more compact macrostate is thermodynamically favored. Thus, the additional PP_{II} content might be required to overcome this to keep an expanded conformation.

These preliminary investigations into structural characteristics of wild-type sequences from the N-terminal regions of p53 homologs lay the groundwork for future studies. By quantifying the structural characteristics for cold-adapted species using variants of wild type sequences we will advance our understanding of contributions from sequence, charge, and hydrophobic residues on temperature-dependence of IDP structure.

REFERENCES

1. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293,** 321–331 (1999).

2. Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19,** 26–59 (2001).

3. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **1804,** 1231–1264 (2010).

4. Van der Lee, R. *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **114,** 6589–6631 (2014).

5. Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18,** 756–764 (2008).

6. Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z. & Dunker, A. K. Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *J. Mol. Biol.* **323,** 573–584 (2002).

7. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27,** 527–533 (2002).

8. Namba, K. Roles of partly unfolded conformations in macromolecular self-assembly. *Genes Cells* **6,** 1–12 (2001).

9. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6,** 197–208 (2005).

10. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu. Rev. Biophys.* **37,** 215–246 (2008).

11. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32,** 1037–1049 (2004).

12. McLaughlin-Drubin, M. E. & Münger, K. The human papillomavirus E7 oncoprotein. *Virology* **384,** 335–344 (2009).

13. Uversky, V. N., Roman, A., Oldfield, C. J. & Dunker, A. K. Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *J. Proteome Res.* **5,** 1829–1842 (2006).

14. Jansma, A. L. *et al.* The high-risk HPV16 E7 oncoprotein mediates interaction between the transcriptional coactivator CBP and the retinoblastoma protein pRb. *J. Mol. Biol.* **426,** 4030–4048 (2014).

15. Dedmon, M. M., Patel, C. N., Young, G. B. & Pielak, G. J. FlgM gains structure in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12681–12684 (2002).

16. Fischer, E. Einflus der configuration auf die wirkung der enzyme. *Ber Dt Chem Ges* **27,** 2985–2993 (1894).

17. Mirsky, A. E. & Pauling, L. On the structure of native, denatured and coagulated proteins. *Proc Natl Acad Sci U S A* **13,** 739–766 (1936).

18. Kendrew, J. C. *et al.* A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181,** 662–666 (1958).

19. Blake, C. C. *et al.* Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* **206,** 757–761 (1965).

20. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. Protein disk of tobacco mosaic virus at 2.8 |[angst]| resolution showing the interactions within and between subunits. *Nature* **276,** 362–368 (1978).

21. Bode, W. The transition of bovine trypsinogen to a trypsinlike state upon strong ligand binding. II. The binding of the pancreatic trypsin inhibitor and of isoleucine-valine sequentially related peptides to trypsinogen and to p-guanidinobenzoate-trypsinogen. *J. Mol. Biol.* **127,** 357–74 (1979).

22. Romero, P. *et al.* Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput. Pac. Symp. Biocomput.* 437–448 (1998).

23. Aviles, F. J., Chapman, G. E., Kneale, G. G., Crane-Robinson, C. & Bradbury, E. M. The Conformation of Histone H5. *Eur. J. Biochem.* **88,** 363–371 (1978).

24. Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37,** 509–516 (2012).

25. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **19,** 23–30 (2009).

26. Ohgushi, M. & Wada, A. 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164,** 21–24 (1983).

27. Ferreon, A. C. M., Moran, C. R., Gambin, Y. & Deniz, A. A. Single-molecule fluorescence studies of intrinsically disordered proteins. *Methods Enzymol.* **472,** 179–204 (2010).

28. Bell, S., Klein, C., Müller, L., Hansen, S. & Buchner, J. p53 Contains Large Unstructured Regions in its Native State. *J. Mol. Biol.* **322,** 917–927 (2002).

29. Hubbard, S. J., Eisenmenger, F. & Thornton, J. M. Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci. Publ. Protein Soc.* **3,** 757–768 (1994).

30. Mandelkern, L., Krigbaum, W. R., Scheraga, H. A. & Flory, P. J. Sedimentation behavior of flexible chain molecules: Polyisobutylene. *J. Chem. Phys.* **20,** 1392–1397 (1952).

31. Tanford, C., Kawahara, K. & Lapanje, S. Proteins as Random Coils. I. Intrinsic Viscosities and Sedimentation Coefficients in Concentrated Guanidine Hydrochloride. *J. Am. Chem. Soc.* **89,** 729–736 (1967).

32. Tomasso, M. E., Tarver, M. J., Devarajan, D. & Whitten, S. T. Hydrodynamic Radii of Intrinsically Disordered Proteins Determined from Experimental Polyproline II Propensities. *PLOS Comput Biol* **12,** e1004686 (2016).

33. Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls, J. E. & Creamer, T. P. Host-guest scale of left-handed polyproline II helix formation. *Proteins Struct. Funct. Bioinforma.* **53,** 68–75 (2003).

34. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are 'natively unfolded' proteins unstructured under physiologic conditions? *Proteins* **41,** 415–427 (2000).

35. Schweers, O., Schönbrunn-Hanebeck, E., Marx, A. & Mandelkow, E. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J. Biol. Chem.* **269,** 24290–24297 (1994).

36. Theillet, F.-X. *et al.* The alphabet of intrinsic disorder. *Intrinsically Disord. Proteins* **1,** e24360 (2013).

37. Perez, R. B., Tischer, A., Auton, M. & Whitten, S. T. Alanine and proline content modulate global sensitivity to discrete perturbations in disordered proteins. *Proteins Struct. Funct. Bioinforma.* **82,** 3373–3384 (2014).

38. Shi, Z. *et al.* Polyproline II propensities from GGXGG peptides reveal an anticorrelation with β-sheet scales. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 17964–17968 (2005).

39. Austin Elam, W., Schrank, T. P., Campagnolo, A. J. & Hilser, V. J. Evolutionary conservation of the polyproline II conformation surrounding intrinsically disordered phosphorylation sites. *Protein Sci. Publ. Protein Soc.* **22,** 405–417 (2013).

40. Langridge, T. D., Tarver, M. J. & Whitten, S. T. Temperature effects on the hydrodynamic radius of the intrinsically disordered N-terminal region of the p53 protein. *Proteins Struct. Funct. Bioinforma.* **82,** 668–678 (2014).

41. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradović, Z. Intrinsic disorder and protein function. *Biochemistry (Mosc.)* **41,** 6573–6582 (2002).

42. Schaub, L. J., Campbell, J. C. & Whitten, S. T. Thermal unfolding of the N-terminal region of p53 monitored by circular dichroism spectroscopy. *Protein Sci.* **21,** 1682–1688 (2012).

43. Molloy, R. G. *et al.* Aquifex aeolicus FlgM protein exhibits a temperature-dependent disordered nature. *Biochim. Biophys. Acta* **1804,** 1457–1466 (2010).

44. Kjaergaard, M. *et al.* Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? *Protein Sci. Publ. Protein Soc.* **19,** 1555–1564 (2010).

45. Wuttke, R. *et al.* Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 5213–5218 (2014).

46. Wells, M. *et al.* Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 5762–5767 (2008).

47. Levine, A. J. p53, the Cellular Gatekeeper for Growth and Division. *Cell* **88,** 323–331 (1997).

48. Bornhorst, J. A. & Falke, J. J. [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.* **326,** 245–254 (2000).

49. Chevallet, M., Luche, S. & Rabilloud, T. Silver staining of proteins in polyacrylamide gels. *Nat. Protoc.* **1,** 1852–1858 (2006).

50. Nitrogen Purging for Circular Dichroism Spectrometers | Applied Photophysics. Available at: https://www.photophysics.com/lab/nitrogen-purging-circular-dichroism-spectrometers. (Accessed: 28th June 2016)

51. Lorber, B., Fischer, F., Bailly, M., Roy, H. & Kern, D. Protein analysis by dynamic light scattering: methods and techniques for students. *Biochem. Mol. Biol. Educ. Bimon. Publ. Int. Union Biochem. Mol. Biol.* **40,** 372–382 (2012).

52. Whitten, S. T. & García-Moreno E, B. pH Dependence of Stability of Staphylococcal Nuclease: Evidence of Substantial Electrostatic Interactions in the Denatured State. *Biochemistry (Mosc.)* **39,** 14292–14304 (2000).

53. English, L. & Tilton, E. The Complete Idiot's Guide to p53 Purification. (2014).

54. English, L. *Schematic for determination of KD using a chromatogram from SEC.* (2016).

55. Jarvet, J. *et al.* A left-handed 31 helical conformation in the Alzheimer Aβ(12–28) peptide. *FEBS Lett.* **555,** 371–374 (2003).

56. Uversky, V. N., Li, J. & Fink, A. L. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.* **276,** 10737–10744 (2001).

57. Dawson, R. *et al.* The N-terminal domain of p53 is natively unfolded. *J. Mol. Biol.* **332,** 1131–1141 (2003).

58. Voet, D. & Voet, J. in *Biochemistry 4e* 67–80 (John Wiley & Sons, 2010).

59. Kelly, S. M. & Price, N. C. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* **1,** 349–384 (2000).

60. Woody, R. W. Circular dichroism and conformation of unordered polypeptides. *Adv Biophys Chem* **2,** 37–79 (1992).

61. Rucker, A. L. & Creamer, T. P. Polyproline II helical structure in protein unfolded states: Lysine peptides revisited. *Protein Sci. Publ. Protein Soc.* **11,** 980–985 (2002).

62. Whittington, S. J., Chellgren, B. W., Hermann, V. M. & Creamer, T. P. Urea promotes polyproline II helix formation: implications for protein denatured states. *Biochemistry (Mosc.)* **44,** 6269–6275 (2005).

63. The Feynman Lectures on Physics Vol. I Ch. 41: The Brownian Movement. Available at: http://www.feynmanlectures.caltech.edu/I_41.html. (Accessed: 8th July 2016)

64. Stein, P. E. *et al.* Crystal structure of ovalbumin as a model for the reactive centre of serpins. *Nature* **347,** 99–102 (1990).

65. Saito, R., Sato, T., Ikai, A. & Tanaka, N. Structure of bovine carbonic anhydrase II at 1.95 A resolution. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 792–795 (2004).

66. Hynes, T. R. & Fox, R. O. The crystal structure of staphylococcal nuclease refined at 1.7 A resolution. *Proteins* **10,** 92–105 (1991).

67. Zahran, Z. N., Chooback, L., Copeland, D. M., West, A. H. & Richter-Addo, G. B. Crystal structures of manganese- and cobalt-substituted myoglobin in complex with NO and nitrite reveal unusual ligand conformations. *J. Inorg. Biochem.* **102,** 216–233 (2008).

68. Jonsson, B. & Jonsson, N. *Ecology of Atlantic Salmon and Brown Trout*. (Springer Netherlands, 2011).

69. Johnson, L. N. & Barford, D. The effects of phosphorylation on the structure and function of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **22,** 199–232 (1993).

70. Katsumata, E., Furuta, C., Katsumata, H., Watanabe, G. & Taya, K. Basal body temperature method for detecting ovarian cycle in the Captive Beluga (Delphinapterus leucas). *J. Reprod. Dev.* **52,** 59–63 (2006).

71. Benzinger, T. H. Heat regulation: homeostasis of central temperature in man. *Physiol. Rev.* **49,** 671–759 (1969).

72. Sund-Levander, M., Forsberg, C. & Wahren, L. K. Normal oral, rectal, tympanic and axillary body temperature in adult men and women: a systematic literature review. *Scand. J. Caring Sci.* **16,** 122–128 (2002).