PHOSPHOSERINE INCORPORATION INTO THE INTRINSICALLY DISORDERED N-TERMINAL DOMAIN OF THE P53 TUMOR SUPPRESSOR PROTEIN

by

George L. Parra, 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 2019

Committee Members:

Steven T. Whitten, Chair

Karen A. Lewis

Sean M. Kerwin

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair use

Duplication Permission

DEDICATION

This effort is dedicated to Claudia Campos, whose love and encouragement made this work possible. Also, to caffeine, beer, and carb heavy snacks, for providing mental energy and inspiration needed for long nights of writing. Finally, to myself for sticking with something, even if it did take some time. Good job buddy. Keep it up.

ACKNOWLEDGEMENTS

I am grateful to my sisters Ashley, Reanna, and Aliza and mother and father, Norma and George Sr. They have provided me nothing but encouragement, and moral and emotional support that I needed at each step of my life. I am also grateful to my other family members and friends who have supported me along the way.

A very special gratitude goes out those at the South Texas Doctoral Bridge Program for helping and providing the funding for tuition; namely Drs. Nicquet Blake, Babatude Oyajobi, Rachell Booth, Ronald Walter, and Raquel Salinas.

With a special mention to Dr. Steven T. Whitten and those who worked in this lab with me. Thank you Dr. Whitten for the opportunity to work in your lab, allowing me the space and resources to learn, and helping me along this journey.

I am also grateful for Dr. Karen Lewis and those in her lab for providing any additional support I needed in my writing and experimental design.

And finally, last but by no means least, also to my close friends, Lance English, Erin Tilton, Gabriela Herrera, Daniel Newton, Brandie Taylor, and Melissa Carrizales… it was great learning from and sharing this experience with you.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

Protein phosphorylation is used biologically as a mechanism to control many critical life processes. Since phosphorylation usually occurs within protein regions that are intrinsically disordered, quantitative and structural descriptions of phosphorylation effects in intrinsically disordered proteins (IDPs) are needed to understand the molecular basis of key aspects of development, aging, and disease. A few studies have investigated the structural effects of phosphorylation in IDPs but, overall, this subject remains mostly uncharacterized. The goal of this thesis project is to adapt literature-reported recombinant techniques for obtaining phosphorylated protein for use in our structural studies of the intrinsically disordered N-terminal region of the p53 tumor suppressor protein, p53(1-93). Healthy p53 activity is regulated by phosphorylation at sites in its disordered N-terminal region, whereas aberrant p53 activity has been linked to numerous human cancers. To obtain phosphorylated p53(1-93), two strategies were employed. The first used a commercially available kinase known as DNA-dependent protein kinase, that is thought to phosphorylate positions 15 and 37. The second is an orthogonal translation system consisting of two independent plasmids and genetically recoded *Escherichia coli* (*E. coli*). One plasmid is used for expression of p53(1-93) and the other for a phosphoserine (Sep) tRNA-synthetase system to permit phosphoserine incorporation into p53(1-93). Sep tRNA utilizes the anti-codon corresponding to TAG and, accordingly, recognizes the TAG stop codon. Two *E. coli* strains that were used have genes for serine phosphatase deleted. One strain has had all TAG stop codons recoded as TAA and release factors

corresponding to TAG silenced. Genes for expressing RNA synthetase and Elongation Factor specific to Sep tRNA are included in the Sep tRNA-synthetase plasmid. The TAG codon will be substituted into the p53(1-93) gene at positions corresponding to serine 15, 33, and 46 to mimic phosphorylation at those sites.

1. INTRODUCTION

Molecular control of cellular functions can be achieved through post-translational modifications (PTMs) of key enzymes and transcription factors in regulatory pathways. Over 200 kinds of post-translational modifications (PTMs) of proteins have been identified in cells[1–3], including covalent attachments of small molecules, such as glycosylation, methylation, and phosphorylation[4], to attachment of whole protein groups like ubiquitin, SUMO-1 and NEDD8. The proteolytic cleavage of the polypeptide chain for the activation of zymogens or removal of the protein via degradation is also considered a PTM[5]. PTMs can induce changes in structure, localization, and activity of the affected protein. A good example of a protein that is tightly controlled by PTMs, in that they affect its regulation, activity and localization is the tumor suppressor protein, p53. A few well studied PTMs of p53 are noted in Figure 1.1 and Table 1.1.

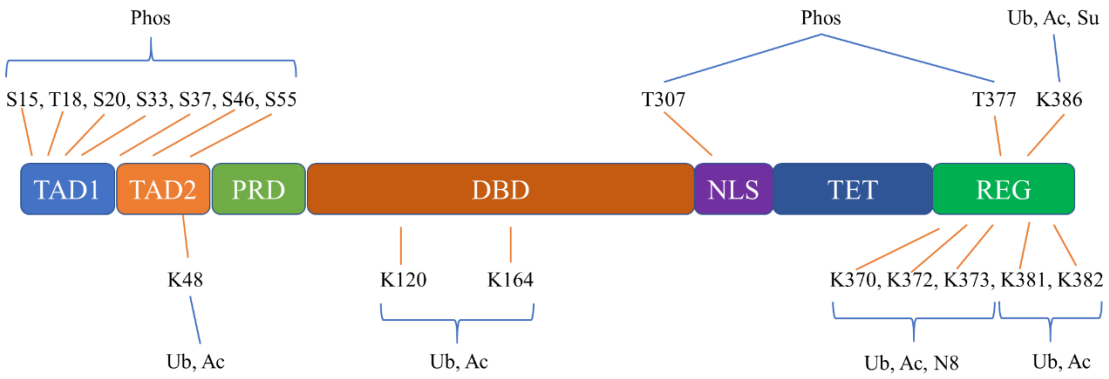

Figure 1.1. Post-translational modification sites of p53. A few well studied post-translationally modified amino acids of p53. Phos, phosphorylation; Ac, acetylation; Ub, ubiquitination; N8, neddylation; Su, SUMOylation.

Table 1.1 Modifiers of p53 activity. A few well studied modifiers of p53 activity through post-translational modification

Modification	Sites	Modifiers
Phosphorylation	S15, T18, S20, S33, S37, S46, S55, T307, T377	ATM, ATR, DNA-PK, CDK5, LRRK2
Acetylation/ Ubiquitination	K48, K120, K164, K381, K382, K370, K372, K373, K386	p300, TIP60,
Neddylation	K370, K372, K373	MDM2
SUMOylaiton	K386	SUMO-1

The p53 gene is oncogenic[6], meaning that its gene produce is expressed at high levels in tumors. The p53 protein is a transcription factor that acts mainly to induce cell cycle arrest or apoptosis in response to DNA damage and replication stress, at least for the unaltered wild type form[7]. It is made up of a core DNA binding domain that shows low thermodynamic stability at human body temperatures, flanked by two disordered regions at both carboxyl- and amino-termini. A depiction of this can be seen later in Figure 1.2. One PTM that is important for maintaining low p53 levels in healthy cells is ubiquitination. Ubiquitination involves the modification of a protein by adding the universally expressed regulatory protein ubiquitin by a ubiquitin ligase. Ubiquitination plays a role in cellular processes such as proteasomal degradation, DNA repair, cell cycle regulation, and gene expression. In healthy cells, p53 levels are maintained at low levels by a few ubiquitin ligase proteins that act to signal p53 degradation by adding the small

regulatory protein to the C-terminal region of p53. MDM2 is a protein ligase that binds to and blocks activation of the p53 N-terminal transactivation domain and when their interaction is disrupted, p53 can carry out its transcriptional activity[8]. Mono-ubiquitination of the C-terminal domain of p53 at lysine residues by MDM2 signals p300/CREB binding protein transcriptional coactivators to begin poly-ubiquitination of p53. Poly-ubiquitination tagged p53 then becomes a substrate for degradation by proteolysis in proteasomes.

When DNA damage occurs, p53 is stabilized by multiple PTMs which suppresses MDM2-mediated degradation allowing p53 to upregulate proteins that are associated with the cell cycle and pro-apoptotic activity. In healthy cells, this results in cell cycle arrest and mitigation of genomic aberrations by not allowing damaged cells to proliferate. One study illustrates how PTMs can affect both stability and activity of p53 toward this end. The N-terminal region of p53 contains seven phosphorylation sites which have been shown to upregulate p53 related transcription and activity[9]. In their study Lee and associates show that Cyclin-dependent kinase 5 (Cdk5) phosphorylation at serine residues 15, 33, and 46 in the N-terminal region of p53. Cdk5 is part of the cyclin dependent kinase family that was discovered for their role in regulating the cell cycle. In neuronal cells, this phosphorylation event stabilizes p53 and disrupts MDM2-mediated ubiquitination. This disruption allows for enhancement of phosphorylation dependent binding of p300 to act as an acetyl transferase[10]. Their data indicated that p300 acetylated p53s C-terminal region to enhance transcriptional activity and mitochondria mediated apoptosis of cells. Their study also revealed that Cdk5 activity increased p53

accumulation in the nucleus upon DNA damage. Their hypothesis was that the Cdk5 mediated disruption of MDM2-p53 interaction stopped MDM2 mediated nuclear export.

1.1 Protein Phosphorylation

Of the posttranslational modifications, phosphorylation is one of the most studied PTMs. Phosphorylation is a reversible modification that is controlled by enzymes known as kinases and phosphatases. Kinases most commonly add phosphate groups to serine, threonine, and tyrosine amino acids and fall into two groups; serine/threonine kinase, or tyrosine kinases. Phosphorylation is used to catalyze activation or deactivation of the target protein. Conversely, phosphatases reverse phosphorylation-induced activity by removing the phosphate group from a protein target. Phosphorylation is important in many eukaryotic cellular processes such as regulating the activity of transcription factors via protein-protein and DNA binding interactions, protein stability, cellular localization, modification of chromatin structures, and targeting for proteolytic degradation [11,12].

A good example of how phosphorylation impacts localization can be seen in p53. p53 activity is known to be modulated heavily by phosphorylation but, in one study by Ho, et al. p53 phosphorylation at threonine sites 307 and 377 by the leucine rich region kinase 2 (LRRK2) protein is a signal for localization. LRRK2 was named for the amount of leucine content that makes up the primary sequence of the protein. Proteins with leucine rich regions have been shown to play important roles in protein-protein interactions, such as signal transduction. The threonine sites of p53 that are phosphorylated by LRRK2 happen to be adjacent to the nuclear localization signal (NLS) regions of p53 that are depicted in Figures 1.1. After investigating the role of LRRK2 on p53 it was shown that phosphorylation at these sites results in the trans-localization of

p53 to the nucleus of the cells, which in turn leads to increased expression of p21, a cyclin-dependent kinase inhibitor that arrests the cell cycle.

The phosphorylation of sites in the disordered N-terminal region of p53 by multiple kinases activates p53 activity resulting in the expression of a number of pro-apoptotic and pro-survival proteins[13]. The N-terminal region of p53 has been shown to be intrinsically disorder and is made up of a transactivation domain (TAD) followed by a proline rich region. Intrinsically disordered protein (IDPs) are proteins that do not form stable structures in biological conditions. Mutations of phosphorylation sites in the N-terninal region of p53 can alter its levels and activity[14]. The transcriptional activity of p53 relies on phosphorylation at each of the seven serine residues and two threonine residues in the N-terminal domain[15]. Phosphorylation of different sites is needed for upregulation or downregulation of different p53 targets[13].

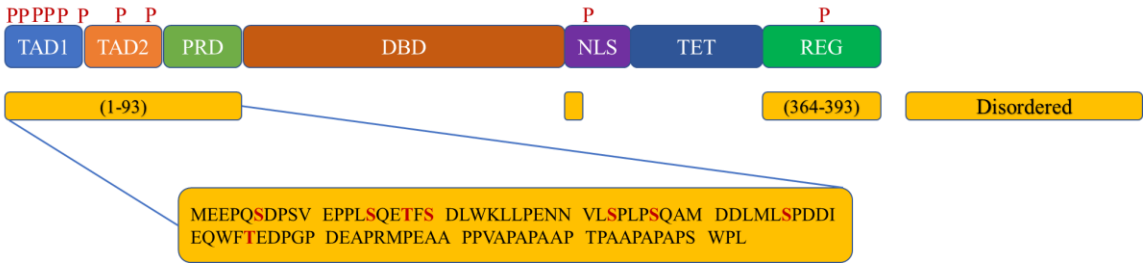

Figure 1.2. A depiction of p53 primary structure. P53 primary structure showing disordered N - and C - terminal regions. Sequence show *in vivo* phosphorylation sites that regulate p53 activity phosphorylation sites on toward the N-terminal region help signal nuclear transport of p53.

1.2 Intrinsically Disordered Proteins

Proteins and protein regions that do not form stable tertiary structures in biological conditions are known as intrinsically disordered proteins (IDPs). The disorder of these proteins is sequence dependent, but can be affected by macromolecular binding partners, small molecule ligands, and PTMs[16]. These interactions can induce stability in IDPs or cause structured proteins to become disordered[17,18]. Furthermore, experimental observations have shown that IDPs structures are sensitive to environmental and structural perturbations such as amino acid substitutions, changes in net charge, primary sequence rearrangements, temperature changes, and changes in pH[19,20,21,22]. IDPs can be characterized by low amounts of bulky hydrophobic residues, high amount of charged residues[23], and low complexity primary sequences of short repeating segments with few residue types[24,25]. This combination is found to promote disorder by favoring interactivity with water as well as allowing intermolecular charge repulsion, keeping the protein from forming a stable globular structure[26].

It has been shown that more than 30% of eukaryotic proteins are IDPs or contain large ID regions. These flexible proteins and protein regions display intermolecular flexibility which allows access to multiple conformational states and interaction points. Consequently, the same protein can adopt multiple biologically active structural configurations[27], allowing "promiscuous" binding[28] and functional capabilities that support signaling pathways[29]. The functions of IDPs can be regulated at the domain level, modulating activity and/or inducing structure changes in only some portions of the protein without impacting other regions[30]. This allows IDPs to function as key regulators of many cellular processes, such as transcription, translation, signal transduction, nuclear

transport, and the formation of stress granules and other subcellular membrane-less organelles[31].

Nuclear transport through the nuclear envelope is one such mechanism that is thought to be mediated by intrinsically disordered proteins. The nuclear envelope is made up of an inner and outer lipid layer, which creates one continuous lipid bilayer that encases the nucleus. The nuclear pore complex is a protein complex that spans the nuclear envelope of eukaryotes with an inner channel that is about 40 nm wide. This complex allows small molecules passive entry into the nucleus, however molecules approaching sizes over 5 nm need to be transported through the complex by shuttling nuclear transport receptors (NTRs). The nuclear pore complex is made up of approximately 34 distinct proteins of which about half are intrinsically disordered[32]. These proteins are called FG proteins due to their primary sequence consisting of stretches of phenylalanine and glycine repeats. It has been shown that NTRs bind selectively but weakly to FG domains however the mechanism of transport is still unclear. Two schools of thought have existed on how the nature of the FG proteins affected nuclear transport[33]. One hypothesis posits that the behavior of FG proteins during transport through the nuclear pore are driven largely by entropic, repulsive, interactions[34]. The other hypothesis is that the behavior is driven by attractive forces between FG proteins[35]. Davis and associates out of Cold Springs Harbor studied FG protein behavior using both computational and experimental data to help elucidate interactions that drive transport. They used purified single FG proteins, and FG proteins grafted to a planar surface at physiological densities to obtain hydrodynamic radius data and molecular compaction data. These experimental models were compared to molecular

dynamic simulations with FG proteins modeled as bead polymers. Molecular simulations were done on FG protein assemblies alone and within modeled DNA origami scaffolds. They found that the FG assemblies exhibited fast dynamics when adopting different morphologies and resealing after perturbation. They concluded that FG protein behavior is not governed by repulsive or attractive forces but rather relies on modulation of the two forces between the protein chains[33]. This balance of attractive and repulsive forces would allow for both sufficient sealing of the transport barrier after a pore soluble protein has been transported as well as allowing FG proteins to exhibit fast dynamic and morphological shifts needed for transport across the nuclear membrane[33].

IDPs have also been shown to be important for formation of proteinaceous membrane-less organelles (PMLOs). PMLOs form in various sizes throughout out the cell consisting of various combinations of local proteins depending on their environment. The formation of PMLOs allow cells to regulate many molecular interactions through isolating molecules of specific cellular processes. PMLOs form when the concentration of molecules gives rise to a change in local density such that a liquid-liquid phase separation (LLPS) of micron-sized droplets occurs. These droplets display liquid like qualities such as being able to drip, fuse, and relax back to spherical shapes after fusing . Though these droplets have no membrane support, their stability is maintained by weak electrostatic protein-protein, protein-RNA, and protein-DNA interactions[36]. IDPs have been shown to make up a significant amount of many human PMLOs[37]. An example of a PMLO driven by IDP activity can be seen in processing bodies (P-bodies), which are RNA/protein aggregates. The main IDP that is involved in LLPS of P-bodies is the LAF-1, with the key region being the RGG domain. This domain contains repeats of arginine

and glycine that impart a high negative charge as well as a preference for open, fluctuating, conformations. This allow LAF-1 protein to overlap multiple molecules and drives LAF-1-LAF-1 interactions that are responsible for LLPS even at low protein concentrations[38].

One group of proteins that are largely composed of IPDPs are proteins involved with transcriptional activity. In particular, eukaryotic transcription factors are thought to be comprised of up to 83% - 94% disordered regions[39]. Many transcription factors have high numbers of disorder-promoting residues (Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro) while being depleted in order-promoting residues (Trp, Tyr, Phe, Ile, Leu, Val, Cys, and Asn). Transcription factors show more disorder in activation domains when compared to DNA binding domains. One such IDP is the N-terminal domain of the transcription factor p53, p53(1-93).

The N-terminal region of p53 (termed "p53(1-93)") is made up of two transactivation domains (TAD1; 1-40, TAD2; 40-60) followed by a proline rich region (PRR; 64-93)[40], as shown in figure 1.2. p53(1-93) contains 22 prolines (23.3%) and a net charge of -15, both of which are characteristics of IDPs as discussed earlier in this this chapter. Another common characteristic of IDPs that is found in the N-terminal region of p53 is PTM sites and binding sites, each of which uniquely regulate the activity of the protein[41,42]. Studying the structural characteristics of the intrinsically disordered N-terminal region of p53 and how its character is influenced by phosphorylation and binding events would be useful for determining the mechanisms utilized by IDPs that allow IDP-regulation of signaling pathways.

1.3 Protein Phosphorylation and Intrinsic Disorder

One PTM that preferentially targets IDPs for modification of regulatory and signaling activities is phosphorylation. The phosphorylation of IDPs alters their activities by enhancing or disrupting weak but specific interactions with binding partners[43] as well as stabilizing and destabilizing secondary structural motifs[17]. The strong correlation between phosphorylation sites and intrinsically disordered regions has been demonstrated in part by Iakoucheva and colleagues who, after analyzing 1,500 known phosphorylation sites, observed that regions adjacent to phosphorylation sites display amino acid sequences that share qualities of intrinsically disordered regions[44]. They observed that amino acid compositions were mainly composed of disorder promoting amino acids, which tend to be less hydrophobic, charged, and more flexible than their order promoting counterparts. This observation goes hand in hand with the multiple studies that show phosphorylation events cause structural and dynamic changes in disordered regions for regulation of activity[17,18]. The relationship between the added charges via phosphorylation and structure should therefore be quantitively evaluated.

The N-terminal region (1-93) of the transcription factor p53 is a known IDP region that is modulated by phosphorylation. This region exhibits disorder associated with high negative charge, disorder promoting residues (namely prolines), and carries seven phosphorylation sites that exhibit biological relevance[9]. p53(1-93) has been extensively characterized by multiple methods including small angle X-ray scattering[45,46], nuclear magnetic resonance relaxation[47], and residual dipolar coupling[46]. Its structure modulation by phosphorylation has also been modeled in computational simulations.. One study by Ithuralde and Turjanski used solved crystal structures of p53(1-39) bound

to the TAZ2 domain of the CREB activator protein to model interactions between phosphorylated p53(1-39) and TAZ2. Their studies showed that phosphorylation led to conformational changes of the bound p53(1-39) that stabilized and increased binding to the TAZ2 domain[48].

Another study carried out by Okuda and Nishimura explored the role that phosphorylation in the TAD plays in binding to TFIIH, a transcription factor whose activity is modulated by p53[49]. It was hypothesized that that the p53 TAD would adopt an amphipathic alpha helix upon binding, as has been seen in previous binding studies using a non-phosphoryted p53 TAD[50,51] however, they showed that the phosphorylated p53 TAD exhibit an elongated structure upon binding. It was stated that the unphosphorylated structure likely bound in an elongated manner to TFIIH as well, but that phosphorylation likely stabilized the elongated structure and strengthened the binding interaction[52].

Sequence motifs of phosphorylation sites on the N-terminal region have been identified as either being directly adjacent to charged amino acids glutamate or aspartate or part of a P-X-X-P motif in which the P is proline, one X either a serine or threonine, and the other X is typically a Lysine[53]. A few but not all the kinases that phosphorylate these motifs are shown in Table 1.1. These sites are highly disordered with mutations that lower local disorder propensities shown to decrease the likelihood of phosphorylation[53]. Furthermore, p53(1-93) has already been used as a model to establish the effects on structure from temperature changes[20], amino acid substitutions[21], and charge-charge interactions[54] on IDPs. Therefore, phosphorylated p53 would be a useful model for exploring how phosphorylation effects structure of IDPs. Obtaining enough

phosphorylated protein for structural studies remains a struggle due to the challenge in purifying site-specific phosphorylated proteins in large enough quantities[55,56].

1.4 Purification of Phosphorylated Proteins

One method to produce phosphorylated proteins for structural studies is using a commercial kinase that is known substrate for the protein of interest to phosphorylate *in vitro*[57]. This involves adding the kinase to a protein sample along with ATP to produce phosphorylated protein. For example, DNA-dependent protein kinase (DNA-PK) is a protein kinase that is known to phosphorylate p53 as well as other regulatory proteins on serine or threonine residues. As the name of the enzyme implies, the activity of DNA-PK is dependent of the presence double-stranded DNA, so commercially available sonicated calf thymus DNA is added to samples of recombinant protein to activate the enzyme and induce phosphorylation. DNA-PK generally phosphorylates sites that are followed in sequence by a glutamine (SQ or TQ motif) [58,59], however, it has also been shown to also phosphorylate serine and threonine residues at non-SQ or TQ positions that are beside a hydrophobic residue such as tyrosine, leucine, alanine, proline, or methionine [60,61,59].

DNA-PK has been shown to phosphorylated p53(1-93) at Serine 15 *in vivo*. Recombinant p53(1-93) contains two SQ motifs, located at positions 15 and 37, with serine 37 being less likely to be phosphorylated due to the lack a charged amino acid at site 36. There are also serine residues that are followed in sequence by a hydrophobic amino acid at positions 9, 33, 45, and 90. Only the serine at residue 90 is part of a phosphorylation motif (P-X-X-P) which DNA-PK does not phosphorylate *in vivo.* There are two threonine residues beside a phenylalanine at positions 18 and 55, with neither being a part of a common phosphorylation motif. Theoretically, only serine 15 would be

phosphorylated by DNA-PK. Phosphorylation of the model IDP p53(1-93) by DNA-PK *in vitro* may be useful in providing enough phosphorylated protein for structural studies.

The orthogonal translation system is another method that can be used for site-specific incorporation of phospho-amino acid residues onto proteins. This system utilizes an RNA synthetase/tRNA pair that is assigned a specific codon for and incorporation of phosphoserine during translation[62]. An RNA synthetase, otherwise known as a tRNA-ligase, catalyzes the esterification of a specific amino acid, or it's precursor to one of its compatible tRNAs to form an aminoacyl tRNA (aa-tRNA). The phosphoserine orthogonal translation system (SepOTS) takes advantage of an O-phosphoseryl-tRNA synthetase (SepRS) found in *Methanosarcina* and *Methanococcus* to aminoacylate phosphoserine onto a tRNA. In OTS methods it is common to assign the new amino acid to a one of the existing stop codons[63]. The SepOTS uses the stop codon UAG otherwise known as the amber stop codon.

The first iteration of a system utilizing $tRNA^{Cys}$ from *Methanocaldococcus jannaschii* and SepRS from *Methanococcus maripaludis* as an orthogonal translation pair for the insertion of Sep was reported by Park and colleagues[56]. The SepRS/$tRNA^{Cys}$ pair was modified to create a tRNA that would recognize the UAG stop codon for insertion. A Sep elongation factor (SepEF) was also included in the system to specifically mediate the entry of the Sep-tRNA into the free site of the ribosome. The original pair was used in BL21 *E. coli*. however, did not incorporate Sep efficiently. This was due to too much competition for the UAG stop codon from the release factor 1 (RF1), which allows the termination of protein translation upon recognizing the amber stop codon. C321.ΔA *E. coli* cells have since been create by deleting the gene for RF1 and recoding all TAG sites

to TAA allowing for the deletion of RF1 with little compromise to fitness[62]. This system may also be useful in production of phosphorylated p53(1-93) for downstream analysis.

1.5 Project Goal

This project aims to test if methods such as DNA-PK mediated phosphorylation or SepOTS could be used to produce recombinant phosphorylated p53(1-93) for studies on the effects of protein phosphorylation on the structural properties of IDPs. Methods such as trying to use Asp and Glu to mimic phosphoserine and phosphothreonine respectively have been successfully used for activity studies but the method is problematic for true structural studies[64]. For instance, the carboxyl groups do carry an extra negative charge but fail to fully represent the -2 charge of a phosphate group. Furthermore, the Asp/Glu residues are smaller and contain different geometry than phosphorylated serine or threonine. This has led to methods for the production of better mimics or the incorporation of phospho-amino acids.

1.5.1 DNA-PK Mediated Phosphorylation

To attempt to phosphorylate recombinantly-expressed p53(1-93) and two p53(1-93) variants, DNA-dependent protein kinase (DNA-PK) was used. Variants of p53(1-93) are Ala^- and Pro^-, in which all alanine or proline residues have been replaced with glycine. These mutants have been previously shown to express well. The activity of DNA-PK is dependent on the presence double-stranded DNA, so commercially available sonicated calf thymus DNA was added to samples of recombinant protein to activate the enzyme and induce phosphorylation.

1.5.2 Phosphorylation of Recombinant p53 at Genetically-defined Positions

The recombinant expression system shown in Figure 1.3 below, is an orthogonal translation system that utilizes two independent plasmids as well as the genomically recoded *Escherichia coli* strain[56]. Two *E. coli* strains have been chosen for site-specific insertion of phosphoserine on TAG sites. For our system, one plasmid will encode the intrinsically disordered N-terminal region of p53, p53(1-93), with TAG codons inserted at serine 15, 33, and 46 combined. These serine positions have been shown to be biologically relevant to p53 activity[10]. The other plasmid will code for tRNA for phosphoserine which contains an anti-codon corresponding to a TAG stop codon, as well as an RNA synthetase and Elongation Factor specific to the phosphoserine tRNA.

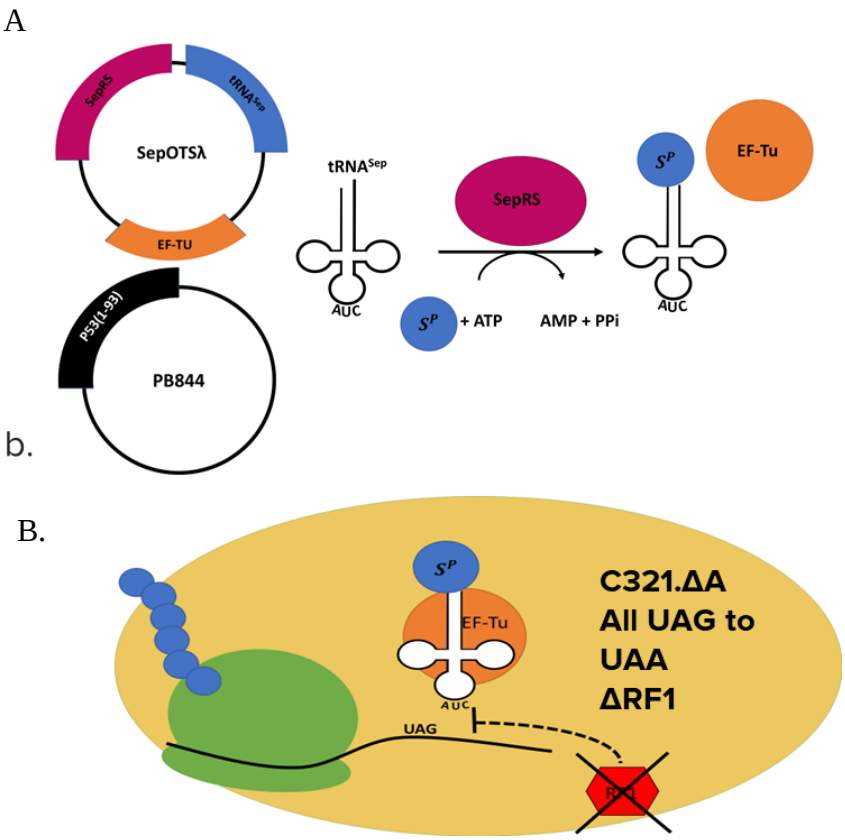

Figure 1.3. A schematic of the Phosphoserine Orthogonal Translation System (Sep-OTS). A. One plasmid codes for $tRNA^{Sep}$, SepRS, and EF-Tu. The second plasmid codes for p53(1-93) with specific serine codons changed to UAG (amber codon). $tRNA^{Sep}$ is aminoacylated with Sep by SepRS. **B.** EF-Tu then delivers Sep-$tRNA^{Sep}$ to the ribosome. Site-specific incorporation of Sep at UAG is directed via the CUA anticodon of $tRNA^{Sep}$.

The goal of this project is to use these two techniques, DNA-PK mediated phosphorylation and an orthogonal translation system for phosphoserine substitution into the protein, to achieve site-specific phosphorylation of recombinant p53(1-93) at purity levels and yield appropriate for structural studies to be conducted in the future.

2. MATERIAL AND METHODS

2.1 Materials and Equipment

All water used for reagents, buffers, and growth media was filtered and deionized using either a Milli-Q Integral 3 pure water system from EMD Millipore (Billerica, MA) or an Elga Purelab Classic water system from Veolia Water Solutions & Technologies (Paris, France). The chemicals used in this project were ACS grade or higher. Equipment such as glass wear, pipette tips, test tubes, and Eppendorf tubes, as well as growth media were autoclaved using a Hirayama HICLAVE HV-50 autoclave (Kasukabe-Shi, Japan). VWR disposable 16×100-mm borosilicate test tubes were used for culturing. Either a A&D GH-200 analytical balance (Tokyo, Japan) or an Ohaus PA84C Pioneer analytical balance (Parsippany, NJ) were used to weigh out chemicals used in buffers and reagents. The pH meters used for pH adjustment of reagents were a Beckman Coulter Φ 510 pH meter (Brea, CA) or Thermo Scientific Orion Star A111 pH meter (Waltham, MA). A Welch DryFast 2032 Ultra Diaphragm Pump (Niles, IL) was used for degassing solutions and vacuum filtration.

All inoculated agar plates were incubated in a VWR 120 V forced air microbiological incubator (Radnor, PA). All inoculated liquid broth cultures were incubated in a Thermo Fisher Scientific MaxQ 5000 floor-model shaker (Waltham, MA). All centrifugation steps were done using a Thermo Fisher Scientific Sorvall LYNX 6000 Superspeed centrifuge (Waltham, MA). A Branson Sonifier S-450A (Danbury, CT) was used to lyse bacterial cells. A Bio-Rad Biologic LP low pressure chromatography system (Hercules, CA) was used to perform nickel affinity chromatography, ion exchange chromatography, and size exclusion chromatography. 12-14 kDa molecular weight cut-

off dialysis tubing from Spectrum Labs, Spectra/Por (Rancho Dominguez, CA) was used for all dialysis steps. CD spectroscopy was performed using a Jasco J-710 spectropolarimeter with a Jasco spectropolarimeter power supply, and Jasco PFD-425S Peltier for temperature control (Easton, MD).

2.2 Protein Sequences

For the initial attempt at obtaining phosphorylated p53(1-93) using DNA-dependent protein kinase, the primary amino acid sequence of the protein was taken from human wild type p53 (WT) and includes a transactivation domain (TAD, residues 1-60) and a proline rich region (PRR, residues 61-93) (see Figure 2.1).

1 MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQAM DDLMLSPDDI
51 EQWFTEDPGP DEAPRMPEAA PPVAPAPAAP TPAAPAPAPS WPL

Figure 2.1. Sequence of WT p53(1-93). Intrinsically disordered N-terminal region of WT p53 includes the transactivation domain in red and the proline rich region in blue.

Two over-expressing mutants were generated in which either all alanine residues (Ala^-) or all proline residues (Pro^-) were substituted for glycine (see Figures 2.2 and 2.3).

1 MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQGM DDLMLSPDDI
51 QWFTEDPGP DEGPRMPEGG PPVGPGPGGP TPGGPGPGPS WPL

Figure 2.2 Sequence of Ala^{-} p53(1-93). Intrinsically disordered N-terminal region of WT p53 includes the transactivation domain in red and the proline rich region in blue. Alanine to Glycine substitutions are underlined.

1 MEEGQSDGSV EGGLSQETFS DLWKLLGENN VLSGLGSQAM DDLMLSGDDI
51 EQWFTEDGGG DEAGRMGEAA GGVAGAGAAG TGAAGAGAGS WGL

Figure 2.3. Sequence of Pro^{-} p53(1-93). Intrinsically disordered N-terminal region of WT p53 includes the transactivation domain in red and the proline rich region in blue. Proline to Glycine substitutions are underlined.

For attempts to purify recombinant phosphorylated WT p53(1-93), the primary amino acid sequence used was taken from the human wild type p53 as above with the exception that serine positions 15, 33, and 46 were substituted with phosphoserine (Figure 2.4).

1 MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQAM DDLMLSPDDI
51 EQWFTEDPGP DEAPRMPEAA PPVAPAPAAP TPAAPAPAPS WPL

Figure 2.4. Sequence of Sep p53(1-93). Intrinsically disordered N-terminal region of WT p53 includes the transactivation domain in red and the proline rich region in blue. Phosphoserine insertions are black and underlined.

2.3 Plasmid Expressions and Purifications

2.3.1 Transformation: DNA-PK Phosphorylation of p53(1-93)

The following transformation methods are for WT p53(1-93) and the two glycine substitution variants, Ala^- p53(1-93) and Pro^- p53(1-93).

The DNA that encodes the protein sequences for the WT p53(1-93) along with the Ala^- and Pro^- variants, as shown in Figures 2.1 – 2.3, were synthesized and cloned into a pj404 vector by ATUM (Menlo Park, CA). Each plasmid contains the DNA sequences mentioned above with a 6x-histidine tag at the N-terminus and a thrombin cut site for downstream removal of the 6x-His tag. Expression of the gene of interest was controlled by an isopropyl-β-D1-thiogalactopyranoside (IPTG) inducible T5 promotor. The plasmid also contained the lactose repressor protein gene (lacI), a high-copy origin of replication (pUC), and an ampicillin resistance gene (amp^r).

Plasmid DNA received from ATUM was diluted to a concentration of 10 ng/µL using autoclaved deionized water, aliquoted out in 5-µL increments, and stored at -80 ºC. Transformation was done by adding 2.5 µL of plasmid DNA to a 50-µL aliquot of *E. coli* BL21 (DE3) pLysS competent cells Novagen (EMD; Billerica, MA) that had been thawed on ice for 10 minutes. The sample was gently flicked 3 times to fully mix then incubated on ice for at least 30 minutes. The cells were then heat shocked in a 42ºC water bath for exactly 30 seconds before returning them to ice to incubate for at least 2 minutes. Finally, 1 mL of 2xYT broth media was added to the cells which were then set to incubate at 37ºC for 60 minutes at 225 RPM for outgrowth.

Once outgrowth was completed 100 µL of the cells was pipetted onto 3 separate Lysogeny Broth (LB) plates with 100 µg/mL of ampicillin to maximize recovery of

transformed cells. All plates were incubated at 37ºC for 15-24 hours. After incubation an inoculation loop was flame sterilized, quenched on the agar, used to select a single colony from one plate to transfer to a test tube of 5 mL 2xYT with ampicillin, and the inoculated sample was incubated at 30ºC for 15 to 20 hours with 225 RPM shaking. To create glycerol stocks of a newly transformed cell, 750 µL of overnight culture was thoroughly mixed with 250 µL of autoclaved 50% (v/v) glycerol in a 2-mL cryovial and stored at -80ºC.

2.3.2 Transformation: Phosphoserine Orthogonal Translation System

The following transformation methods are for the Sep p53(1-93) and SepOTSλ plasmids for the phosphoserine orthogonal translation system.

The DNA that encodes the protein sequence for Sep-p53(1-93) with Amber stop codons (TAG) corresponding to serine sites 15, 32, and 46 was synthesized and cloned into both pD864-WR or pD884-SR vectors by ATUM (Menlo Park, CA). Each plasmid contains the DNA sequences as shown in Figure 2.4 along with a 6x-histidine tag at the N-terminus and a thrombin cleavage site for downstream removal of the 6x-His tag. Expression of the gene of interest on both plasmids was controlled by a rhamnose inducible promotor (rhaBAD). The pD864-WR plasmid contains a high-copy origin of replication (pUC) and a weak ribosome binding site. The pD884-SR plasmid contains a low-copy origin of replication (p15a) and a strong ribosome binding site. Both plasmids carry an ampicillin resistance gene (amp^r). The Sep-p53(1-93) expression plasmids received from ATUM were diluted to a concentration of 10 ng/µL using autoclaved deionized water, aliquoted out in 5-µL increments, and stored at -80 ºC.

The phosphoserine orthogonal translation system plasmid (SepOTSλ), and the serine suppressor tRNA plasmid (SupD) were ordered from Addgene (Watertown, MA). Plasmids were provided in bacterial stabs of DH5α and Top10 *E. coli* cells respectively. These cell types are both descendants of the K-12 *E. coli* strain and generally used for plasmid storage and cloning. The SepOTSλ plasmid contain four repeats of the gene for an Amber stop codon suppressor tRNA specific to phosphoserine (4xtRNA-Sep-A37) and an elongation factor (SepRS9/EFSep21) specific to that amino-acylated-tRNA. The SupD plasmid contains 2 Amber stop codon suppressor tRNAs specific to serine. Expression of the genes of interest on both plasmids was controlled by an isopropyl-β-D1-thiogalactopyranoside (IPTG) inducible promotor. The plasmids also contain the lactose repressor protein gene (lacI), a high-copy origin of replication (pBR332), and a kanamycin resistance gene (kan^r). A QIAprep Spin Miniprep Kit from Qiagen (Hilden, Germany) was used to isolate each plasmid following the protocol provided in the kit. The extracted plasmids, SepOTSλ and SupD, were diluted to a concentration of 10 ng/µL using autoclaved deionized water, aliquoted out in 5-µL increments, and stored at -80 ºC.

The *E. coli* strains BL21ΔserB and C321.ΔA were ordered from Addgene to use as the expression vectors for the SepOTS method. C321.ΔA has 321 genome wide UAG to UAA conversions, as well as the deletion of the UAG associated release factor 1 (RF1) gene (Δ*prfA*)[65]. C321.ΔA also carries resistance to zeocin (Δ*mutS:zeo*, Δ*tolC*) and the deletion of the gene for phosphoserine phosphatase (*SerB-*/Δ*SerB*) to keep phosphoserine at optimum levels within the cell for incorporation. BL21ΔserB has the phosphoserine phosphatase gene deleted as well (*SerB-*/Δ*SerB*).

Transformation was done by adding 4 µL a Sep-p53(1-93) expression plasmid DNA to a 50-µL aliquot of competent *E. coli* C321.ΔA or BL21ΔserB that had been thawed on ice for 10 minutes. The sample was gently flicked 3 times to fully mix then incubated on ice for at least 30 minutes. The cells were then heat shocked in a 42ºC water bath for exactly 30 seconds before returning them to ice to incubate for at least 2 minutes. Finally, 1 mL of 2xYT broth media was added to the cells which were then set to incubate at 37ºC for 60 minutes at 225 RPM for outgrowth.

Once outgrowth was completed, 100 µL of the cells was pipetted onto 3 separate Lysogeny Broth (LB) plates with 100 µg/mL of ampicillin to maximize recovery of transformed cells. All plates were incubated at 37ºC for 15-24 hours. After incubation an inoculation loop was flame sterilized, quenched on the agar, used to select a single colony from one plate to transfer to a test tube of 5 mL 2xYT with ampicillin, and the inoculated sample was incubated at 30ºC for 15 to 20 hours with 225 RPM shaking. To create glycerol stocks of a newly transformed cell 750 µL of overnight culture was thoroughly mixed with 250 µL of autoclaved 50% (v/v) glycerol in a 2-mL cryovial and stored at -80ºC. A total of four transformation products were produced: two C321.ΔA transformation products containing one of the Sep-p53(1-93) expression plasmids, and two BL21ΔserB transformation products containing one of the Sep-p53(1-93) expression plasmids.

The four newly transformed vectors with Sep-p53(1-93) were made chemically competent again. Transformation was carried out a second time on each of the four previous transformation products by adding 4 µL of SepOTSλ or SupD plasmid DNA to a 50-µL aliquot of competent cells that had been thawed on ice for 10 minutes. The

sample was gently flicked 3 times to fully mix then incubated on ice for at least 30 minutes. The cells were then heat shocked in a 42ºC water bath for exactly 30 seconds before returning them to ice to incubate for at least 2 minutes. Finally, 1 mL of 2xYT broth media was added to the cells which were then set to incubate at 37ºC for 60 minutes at 225 RPM for outgrowth.

Once outgrowth was completed 100 µL of the cells was pipetted onto 3 separate Lysogeny Broth (LB) plates with 100 µg/mL of ampicillin and 50 µg/mL kanamycin to maximize recovery of transformed cells. All plates were incubated at 37ºC for 15-24 hours. After incubation an inoculation loop was flame sterilized, quenched on the agar, used to select a single colony from one plate to transfer to a test tube of 5 mL 2xYT with ampicillin and kanamycin, and the inoculated sample was incubated at 30ºC for 15 to 20 hours with 225 RPM shaking. To create glycerol stocks of a newly transformed cell 750 µL of overnight culture was thoroughly mixed with 250 µL of autoclaved 50% (v/v) glycerol in a 2-mL cryovial and stored at -80ºC. The transformation process was repeated yielding six transformation products. Two C321.ΔA transformation products containing one of the Sep-p53(1-93) expression plasmids along with the SepOTSλ were obtained. Two BL21ΔserB transformation products containing one of the Sep-p53(1-93) expression plasmids along with the SepOTSλ were also obtained. Finally, one C321.ΔA transformation product containing pBR322 ORI Sep-p53(1-93) and SupD, and one BL21ΔserB transformation product containing p15a ORI Sep-p53(1-93) and SupD. The same transformation steps were also used to create a BL21ΔserB transformation product with an Ala^{-} p53(1-93) plasmid with SepOTSλ.

2.3.3 Expression: WT p53(1-93) and substitution variants

The glycerol stock from the transformation steps were used to streak a LB plus ampicillin plate, which was incubated for 15-24 hours at 37ºC. A single colony was transferred to a 5 mL 2xYT plus ampicillin test tube using a flame sterilized inoculation loop. The sample was incubated at 30ºC with 225 RPM shaking for 15-20 hours. After incubation, the overnight culture was used to bring 1 L of sterile 2xYT plus ampicillin in a 4-L flask to an optical density (OD) of ~0.001 A.U. at a wavelength of 600 nm (OD_{600}). The culture was then incubated at 37ºC with 225 RPM shaking until the culture reached an OD_{600} between 0.6 and 0.7 A.U., which signifies mid-exponential growth phase.

When the OD_{600} is within the 0.5 – 0.6 range protein expression was induced by adding IPTG to the final culture to a final concentration of 1.0 mM. The culture was allowed to continue incubation at 37ºC with 225 RPM shaking for 5 hours. Cells were then transferred to 250-mL centrifuge tubes and centrifuged for 15 minutes at 30,2240 x g at 4ºC using a F14-6x25y rotor. After centrifugation, the supernatant was decanted off the newly formed cell pellets and discarded. The cell pellets were stored overnight at -80ºC.

2.3.4 Expression: Sep p53(1-93)

The glycerol stocks from the transformation step were used to streak a LB plus ampicillin, kanamycin, and bleomycin plate, which was incubated for 15-24 hours at 37ºC. A single colony was transferred to a 5 mL 2xYT media plus ampicillin, kanamycin, and bleomycin test tube using a flame sterilized inoculation loop. The sample was incubated at 30ºC with 225 RPM shaking for 15-20 hours. For protein purification experiments, after incubation was complete the overnight culture was used to inoculate 1

L of sterile 2xYT plus ampicillin, kanamycin, and bleomycin in 4-L Erlenmeyer flasks. For protein expression experiments, overnight culture was used to inoculate 10 mL or 100 mL of sterile 2xYT plus ampicillin, kanamycin, and bleomycin in a test tube or 500-mL Erlenmeyer flask. An initial optical density (OD) of ~0.15 A.U. at a wavelength of 600 nm (OD_{600}) was used for both purification and expression trials. The cultures were then incubated at 37 ºC with 225 RPM shaking until the culture reached an OD_{600} between 0.8 and 0.9 A.U. for protein expression.

For experimentation with glucose concentration effects on OD_{600} concentrations of 0.1 %, 0.5 %, 1.0 %, and 2.0 % w/v glucose was added to the 2xYT growth media that was used for the final expression step.

When the OD_{600} is within 0.8 to 0.9 OD_{600}, protein expression was induced by adding IPTG to the final culture to a final concentration of 1.0 mM, and Rhamnose to a final concentration between 0.2 % w/v. The culture was allowed to continue incubation at either 20 ºC, 30 ºC, or 37 ºC with 225 RPM shaking for 5 hours (37 ºC) to 20 hours (20 ºC, 30 ºC). Negative controls were also used at the 10 ml and 100 ml volumes in which the cultures were not exposed to IPTG or rhamnose. 1 mL samples were aliquoted from the 10 mL and 100 mL expressions, at 0 hours, 1 hours, 5 hours, and 20 hours into induction depending on temperature of induction. Aliquots were spun down at 14,000 x g in a Beckman Coulter Microfuge 16 centrifuge and supernatant discarded. Pellets were stored at -80 ºC for gel electrophoresis. Induced cells at 1 L and 4 L volumes were then transferred to 250-mL centrifuge tubes and centrifuged for 15 minutes at 30,2240 x g at 4ºC using a F14-6x25y rotor. After centrifugation, the supernatant was decanted off the newly formed cell pellets and discarded. The cell pellets were stored overnight at -80ºC.

For experiments on ethanol effects on protein production concentrations of 2.0 % v/v was added during the induction step and the rest of the process was carried out as stated above.

2.3.5 Nickel affinity chromatography

Cell pellets from the 1 L protein expressions were thawed and resuspended in lysis buffer (6 M guanidine, 10 mM Tris, 100 mM Na_2PO_4, pH 8.0) such that 30 mL of lysis buffer was used for 3 cell pellets at a time. The resulting solution was transferred to a 50-mL conical and sonicated on ice for 90 seconds on and 90 seconds off for 4 rounds at 80% duty, and 50% output. The lysate was centrifuged for 1 hour at 4ºC at 33,746 x g using the F14-14 x 50cy rotor.

During centrifugation, about 15 mL of HIS-select nickel affinity gel from Sigma-Aldrich (St. Louis, MO) was added to a 1.5 x 30 cm Bio-Rad Econo-column, which was then attached to the low-pressure liquid chromatography system. An Equilibration Buffer (6 M guanidine, 10 mM Tris, 100 mM Na_2PO_4, pH 8.0) was run through the column until the conductivity and absorbance at 280 nm (A_{280}) became constant and level.

Once the column was equilibrated and the centrifugation step complete, the supernatant of the sample was loaded into the nickel affinity column and washed with 45 mL of Wash Buffer #1 (6 M guanidine, 10 mM Tris, 100 mM Na_2PO_4, pH 8.0), followed by 40 mL of Wash Buffer #2 (10 mM Tris, 100 mM Na_2PO_4, pH 8.0). Next, 40 mL of Wash Buffer #3 (10 mM Tris, 100 mM Na_2PO_4, 10 mM imidazole, pH 8.0) was added to the column to elute any protein with weak interactions to the nickel media from the column. Finally the Elution Buffer (10 mM Tris, 100 mM Na_2PO_4, 350 mM imidazole, pH 4.3) was used to elute the protein of interest, which was collected in a 50-mL conical

from the moment the conductivity and absorbance began to rise up until the conductivity and absorbance began to level off. Dialysis of the eluate was carried out overnight in Tris-buffered solution (TBS; 20 mM Tris, 100 mM NaCl, pH 8.0) at 4ºC.

In the case of the p15a ORI p53(1-93) plasmid, tobacco etch virus (TEV) protease was used for the cleavage of the 6xHis tag. To remove the 6xHis tag, 100 U per 1 mg protein of 7xHis tagged TEV protease (Sigma-Aldrich, St. Louis, MO) was added to the dialyzed sample in a 50-mL conical wrapped in foil to keep out light. The sample was placed on a Thermo Fisher Scientific Large 3-D rotator and the digestion was run for 16 hours at room temperature at low rotation. Using the same type of HIS-select nickel affinity gel as the previous step, the nickel media was added to a 1.5 x 30 cm Bio-Rad Econo-column The column was then attached to the low-pressure liquid chromatography system. An Equilibration Buffer (10 mM Tris, 100 mM Na_2PO_4, pH 8.0) was run through the column until the conductivity and absorbance at 280 nm (A_{280}) became constant and level.

Once the column was equilibrated and the digestion step complete, the digested sample was loaded into the nickel affinity column and washed with 45 mL of Wash Buffer #1 (10 mM Tris, 100 mM Na_2PO_4, pH 8.0). The protein of interest was collected during the flow through in a 50-mL conical. Collection started from the moment the conductivity and absorbance began to rise up until the conductivity and absorbance began to level off. If no protein came off this step then 40 mL of Wash Buffer #2 (10 mM Tris, 100 mM Na_2PO_4, pH 8.0) was added. Any eluate that comes off the column is collected as before. Next, 40 mL of Wash Buffer #3 (10 mM Tris, 100 mM Na_2PO_4, 10 mM imidazole, pH 8.0) was added to the column to elute any protein that was not cleaved.

Finally, the Elution Buffer (10 mM Tris, 100 mM Na_2PO_4, 350 mM imidazole, pH 4.3) was used to elute more un-cleaved protein, which was collected in a 50-mL conical. Un-cleaved eluate was re-dialyzed in Tris-buffered saline (TBS; 20 mM Tris, 100 mM NaCl, pH 8.0) at 4ºC and prepped for re-digestion. Dialysis of the cleaved eluate was carried out overnight in 4 L of ddH_2O.

After overnight dialysis of the cleaved and purified sample, the protein was concentrated by splitting the sample evenly in 2-mL microcentrifuge tubes and placing them into a Thermo-Fisher Scientific Savant DNA 120 SpeedVac. The SpeedVac was run on medium without heat until samples reached a volume between 0.2 and 0.1 mL, after which the protein samples were combined and re-dialyzed overnight in phosphate-buffered solution (10 mM sodium phosphate, 100 mM NaCl, pH 7.0) at 4 ºC. After dialysis the sample was aliquoted evenly in 1.5 mL microcentrifuge tubes and stored in the -80 ºC freezer.

2.3.6 Ion Exchange Chromatography

To remove the Histidine tag, fifty units of thrombin (Sigma-Aldrich, St. Louis, MO) were added to the dialyzed sample in a 50-mL conical wrapped in foil to keep out light. The sample was placed on a Thermo Fisher Scientific Large 3-D rotator and the digestion was run for 4 hours at room temperature at low rotation. While thrombin digestion was carried out, 10 mL of diethylaminoethyl (DEAE) Sephacel media (GE Healthcare, Chicago, IL) was decanted and mixed with 30 mL of Column Media Buffer (20 mM sodium acetate, 25 mM NaCl, pH 3.5) in 50-mL beaker. The DEAE media was degassed for at least 30 minutes before being transferred to a 1.5 x 30 cm Bio-Rad Econo-column and attached to the low-pressure liquid chromatography system. The

media was equilibrated using Equilibration Buffer (20 mM sodium actate, 25 mM NaCl, pH 4.8) until the A_{280} and conductivity leveled off.

Once the media was equilibrated and the thrombin digestion carried out to the sample was added to the DEAE column and washed with 50 mL Wash Buffer 50 (20 mM sodium acetate, 50 mM NaCl, pH 4.8) to remove weakly binding protein from the media. More weakly binding proteins were washed off the column with Wash Buffer 150 (20 mM sodium acetate, 150 mM NaCl, pH 4.8) before the target protein was eluted off the column with Elution Buffer (20 mM sodium acetate, 400 mM NaCl, pH 4.8). The eluate was collected as soon as the A_{280} and conductivity began to rise and stopped when they leveled off. Finally, the collected protein sample was dialyzed overnight in a phosphate-buffered solution (10 mM sodium phosphate, 100 mM NaCl, pH 7.0) at 4ºC.

After overnight dialysis, the purified protein was concentrated by splitting the sample evenly in 2-mL microcentrifuge tubes and placing them into a Thermo-Fisher Scientific Savant DNA 120 SpeedVac. The SpeedVac was run on medium without heat until samples reached a volume between 0.2 and 0.1 mL, after which the protein samples were combined and re-dialyzed overnight in the same phosphate-buffered solution as above (10 mM sodium phosphate, 100 mM NaCl, pH 7.0). After dialysis the sample was aliquoted evenly in 1.5 mL microcentrifuge tubes and stored in the -80 ºC freezer.

2.3.7 DNA-PK Reaction and Ion Exchange Recovery

For the DNA-PK phosphorylation method part of each purified p53(1-93) variant (WT, Ala^-, or Pro^-) were subjected to phosphorylation reactions. Reactions consisted of 50mM HEPES (pH 7.5), 1mM DTT, 0.1mM EDTA, 0.2mM EGTA, 10mM MgCl2, 0.1M KCl, 1.14mM of one p53(1-93) variant at a time, 80µg/mL BSA, 0.2mM ATP, 400

units/mL and 10µg/mL linear double-stranded DNA. One unit of DNA-PK is the amount of enzyme needed for incorporation of 1 pmol of phosphate into a DNA-PK substrate in one minute at 30°C. Each reaction was carried out for 10 minutes at 30°C before being stopped by 20 µL of 30% acetic acid.

Before the reaction was done 10 mL of diethylaminoethyl (DEAE) Sephacel media (GE Healthcare, Chicago, IL) was decanted and mixed with 30 mL of Column Media Buffer (20 mM sodium acetate, 25 mM NaCl, pH 3.5) in 50-mL beaker. The DEAE media was degassed for at least 30 minutes before being transferred to a 1.5 x 30 cm Bio-Rad Econo-column and attached to the low-pressure liquid chromatography system. The media was equilibrated using Equilibration Buffer (20 mM sodium acetate, 25 mM NaCl, pH 4.8) until the A_{280} and conductivity leveled off.

Once the phosphorylation reaction was done being carried out the sample was added to the DEAE column and washed with 50 mL Wash Buffer 50 (20 mM sodium acetate, 50 mM NaCl, pH 4.8) to remove weakly binding protein and DNA from the media. The column was then washed with Wash Buffer 150 (20 mM sodium acetate, 150 mM NaCl, pH 4.8) before the target protein was eluted off the column with Elution Buffer (20 mM sodium acetate, 400 mM NaCl, pH 4.8). The eluate was collected as soon as the A_{280} and conductivity began to rise and stopped when they leveled off. Finally, the collected protein sample was dialyzed overnight in a phosphate-buffered solution (10 mM sodium phosphate, 100 mM NaCl, pH 7.0) at 4ºC.

After overnight dialysis, the purified protein was concentrated by splitting the sample evenly in 2-mL microcentrifuge tubes and placing them into a Thermo-Fisher Scientific Savant DNA 120 SpeedVac. The SpeedVac was run on medium without heat

until samples reached a volume between 0.2 and 0.1 mL, after which the protein samples were combined and re-dialyzed overnight in the same phosphate-buffered solution as above (10 mM sodium phosphate, 100 mM NaCl, pH 7.0). After dialysis the sample was aliquoted evenly in 1.5 mL microcentrifuge tubes and stored in the -80 ºC freezer.

2.3.8 Gel Electrophoresis of Purified Samples

The purity of the proteins expressed and purified was gauged via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). To perform SDS-PAGE, 20 µL of sample was mixed with 20 µL 2xSDS-reducing Laemmli Buffer (2% w/v SDS, 25% v/v glycerol, 5% v/v 2-mercaptoethanol, 0.01% w/v Bromophenol Blue and 0.0625 M Tris HCl, pH 6.8), in a 500µL-microcentrifuge tube which was then heated to 95ºC for 7 minutes with 300 RPM in an Eppendorf Thermomixer R (Hamburg, Germany). A pre-cast Bio-Rad 4-20% Criterion Tris-HCl gel was placed into a Bio-Rad Criterion Cell, and the top of the gel and gel rig were filled with Running Buffer (250 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) to points indicated on the equipment by the manufacturer. Once the sample was finished being heated, 10-20 µL of sample mixture was loaded into one well, 5 µL of Bio-Rad Precision Plus Protein Standard were added to one to two other wells and left-over wells were loaded with 10 µL 1xSDS-rudcing Laemmli Buffer. The cell with loaded samples was connected to a Bio-Rad PowerPac HV power supply and gel electrophoresis was carried out at 100 V for 20 minutes to ensure proper stacking of the protein sample, followed by ~40 minutes at 200 V at room temperature.

2.3.9 Hand Cast Gels for SDS-PAGE

Hand cast gels were used to run SDS-PAGE on aliquoted time course pellets to check for protein expression. An acrylamide concentration of 14% was used. A Bio-Rad Mini-PROTEAN

Gel Casting Stand System (Bio-Rad Laboratories, Hercules, CA) was set up with a 1.5 mm spacer plate and a small plate flush against gaskets of the apparatus. A 10-well comb was placed in the plates and a black line was marked 1 cm below the bottom of the comb and then placed aside. 20 mL of resolving gel solution (14 % w/v Acrylamide/bis, 1 % w/v SDS, 0.375 M Tris-HCl, pH 8.8) was degassed for 15 minutes without covering. After degassing, APS and TEMED were added to the resolving solution at final concentrations of 0.0005 % w/v and 0.001 % v/v, respectively and the solution was stirred to initiate polymerization. A glass pipette was used to pipette in the solution to the black line and the resolving solution was covered in 100% isopropyl alcohol while the gel polymerized for 45 minutes. While the resolving gel was polymerizing a 10 mL stacking gel solution (14 % w/v Acrylamide/bis, 1 % w/v SDS, 0.125 M Tris-HCl, pH 6.8) was prepared and degassed for 15 minutes. When the resolving gel was done polymerizing and the stacking solution was done degassing, APS and TEMED were added to the stacking solution to final concentrations of 0.0005 % w/v and 0.002 % v/v, respectively. The stacking solution was mixed, the isopropyl alcohol was then drained off the resolving gel and the stacking solution was pipetted on top using a glass pipette. The comb was placed in and the stacking gel allowed to polymerize for 30 minutes before use.

2.3.10 Gel Electrophoresis of Whole Cell Lysate Samples

The degree of protein expression from the 10 mL and 100 mL expression experiments was gauged via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Cell pellets from aliquoted samples were mixed with 100 µL 1xSDS-reducing Laemmli Buffer (1% w/v SDS, 25% v/v glycerol, 5% v/v 2-mercaptoethanol, 0.01% w/v Bromophenol Blue and 0.0625 M Tris HCl, pH 6.8), in a 500µL-microcentrifuge tube which was then heated to 95ºC for 7 minutes with 300 RPM in an

Eppendorf Thermomixer R (Hamburg, Germany). The cell lysate was then spun down at *14,000* x g in a Beckman Coulter Microfuge 16 centrifuge for 5 minutes and supernatant saved. Hand cast Mini-PROTEAN gels were placed in a Mini-PROTEAN Tetra gel rig, and the top of the gel and gel rig were filled with Running Buffer (250 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) to points indicated on the equipment by the manufacturer. Once the sample was ready, 10-20 µL of sample mixture was loaded into one well, 5 µL of Bio-Rad Precision Plus Protein Standard were added to one to two other wells and left-over wells were loaded with 10 µL 1xSDS-rudcing Laemmli Buffer. The cell with loaded samples was connected to a Bio-Rad PowerPac HV power supply and gel electrophoresis was carried out at 70 V for 20 minutes to ensure proper stacking of the protein sample, followed by ~40 minutes at 140 V at room temperature.

2.3.11 Coomassie Staining

After electrophoresis was carried out precast gels were removed from their casing and soaked with deionized water for 5 minutes before being transferred to Coomassie Blue stain (10% Acetic Acid, 40% Methanol, 0.05% Brilliant Blue) to soak for anywhere between 90 minutes to overnight. Coomassie stain was carefully poured back into its bottle and the gel was washed two to three times with water to remove excess stain. The gel was then soaked in Coomassie De-stain (10% acetic acid, 40% methanol) for 1 to 6 hours depending on how long it was stained for. When the gel background was clear the de-stain was discarded, and the gel washed in deionized water for 30 minutes. A Bio-Rad Molecular Imager ChemiDoc XRS+ imaging system was used to image the gel.

2.3.12 Silver Staining

After electrophoresis of hand cast gels, the gels were removed from the glass plates and rinsed in water a few times. Gels were soaked in fixing solution (30% ethanol, 10% acetic acid) for 30 minutes. After fixing the gels were rinsed twice in a 20% ethanol solution for 10 minutes each time. Gels were next subjected to sensitization via 0.02% w/v sodium thiosulfate for 1 minute, followed immediately by rinsing twice with water for 1 minute each time. The sensitized gels were impregnated in 12 mM silver nitrate for 20 minutes to 2 hours. After impregnation the gel was rinsed quickly (~10 seconds) with water and developed in a basic developing solution (2.3% w/v sodium carbonate, 9.125×10^{-5}% v/v formaldehyde, 1.25×10^{-5} % w/v sodium thiosulfate) for 3 to 10 minutes. Once bands developed the developer was disposed of and gels were rinsed in stop solution (2% acetic acid, 330 mM Tris) for 30 minutes up to 2 hours. The gels were rinsed in water twice for at least 30 minutes each before imaging with the Bio-Rad Molecular Imager ChemiDoc XRS+ imaging system.

2.4 Mass Spectrometry

Samples that were analyzed using electrospray ionization mass spectrometry (ESI-MS) were dialyzed in 4 L of ddH_2O overnight at 4 ºC. Dialyzed samples were then diluted to 10 – 15 mM using autoclaved ddH_2O prior to being submitted for mass spectrometry by Bo Song in Dr. Xiaopeng Li's lab. Molecular weights of recombinant protein (M_r) were calculated using

Equation 2.1 $$\boldsymbol{p = m/z}$$

Equation 2.2 $$\boldsymbol{p_1 = \frac{M_r + z_1}{z_1}}$$

Equation 2.3 $$\boldsymbol{p_2 = \frac{M_r + (z_1 - 1)}{z_1 - 1}}$$

where p_1 and p_2 are adjacent peaks and z_1 is the charge of peak one.

2.5 Circular Dichroism Spectroscopy

Circular Dichroism was used to investigate the secondary structure, namely PP_{II} propensity, of WT p53(1-93) and the variants as well is if their phosphorylated counterparts' secondary structure was affected by the addition of phosphate groups.

A sample of protein was loaded into a 1 mm pathlength Jasco quartz cuvette at a final concentration of 0.17 mg/mL in a phosphate buffered saline (10 mM Na_2PO_4, 100 mM NaCl, pH 7.0). A Jasco J-710 spectropolarimeter with a Jasco spectropolarimeter power supply, and a Jasco PFD-425S Peltier were used to acquire measurements from the 197-245 nm range in 0.5 nm increments at a rate of 20 nm/minute. Before taking measurements, nitrogen was used to flood the spectrophotometer for 15 minutes after which the xenon lamp was turned on. Eight scans were taken at each temperature from 5 ºC – 85 ºC in 10 ºC increments with the nitrogen gas set at a flow rate of 5 L/min. Raw CD data was reported as an average of the eight scans and converted to molar residue ellipticity (MRE) using Equation 2.1. The data was plotted using Microsoft Excel.

Equation 2.4: $$MRE = \frac{(CD\ millidegrees)*(0.1)*\left(mean\ residue\ weight\frac{g}{mol}\right)}{(pathlength\ cm)*(concentration\frac{mg}{mL})}$$

2.6 Size Exclusion Chromatography

Distribution coefficients (K_d), which are used to calculate hydrodynamic radii (R_h) of WT and variant p53(1-93) as well as RSNase scrambled sequence proteins, was determined via size exclusion chromatography (SEC). Sephadex G-100 gel filtration media from GE Healthcare (Pittsburgh, PA) was used for p53(1-93) experiments, while Sephadex G-75 gel filtration media, also from GE Healthcare, was used for scrambled sequence RSNase expreiments. All p53(1-93) samples were combined with indicator dyes, Blue Dextran and 2, 4-dinitrophenyl-L-asartate (DNP-aspartate), for the determination of the void and total volumes for the column for each reading. Blue Dextran and DNP-aspartate were run separately from the scrabled RSNase samples as the Blue Dextran and DNP-aspartate were shown to interact with the RSNase proteins. Four liters of phosphate buffered saline solution (100 mM NaCl, 10 mM Na_2PO_4, pH 7.0) was filtered using a GE Healthcare Whatman 1 qualitative filter paper using a Welch Dryfast Ultra vacuum pump connected to a filter flask. This solution was degassed in the flask for 10 minutes after filtration. The PBS was transferred to a 4-L beaker.

For SEC experiments 8 g of Sephadex G-75 media was equilibrated in about 150 mL of phosphate buffered saline solution (100 mM NaCl, 10 mM Na_2PO_4, pH 7.0) overnight at room temperature. Phosphate buffered solution was decanted from the top of the mixture until a 75:25 media:solvent ration was by volume was reached. The media was degassed for at least 3 hours. The degassed media was poured into a 1.5 x 30 cm Bio-Rad Econo-column which was attached to the Biologic low-pressure liquid chromatography system. The top of the column was connected to Tygon tubing that was anchored to the bottom of the 4-L beaker of degassed phosphate buffered solution. The

media was allowed to pack by gravity overnight while a constant flowrate was maintained. The same flowrate was maintained through the whole experiment. Four protein standards were chosen for comparing their partition coefficients (K_D) to the K_Ds the p53(1-93) proteins. Due to the wide range among R_h values, the protein standards chosen were *Staphylococcal* nuclease, Horse heart myoglobin, Bovine carbonic anhydrase, and Chicken egg albumin.

Protein standards, except for Chicken egg albumin, and p53(1-93) proteins were measured in the following way. 90 µL of BD-DNP-Asp (3 mg/mL Blue Dextran D5751, 0.75 mg/mL DNP-Aspartate) was mixed with 10 µL of protein sample right before measuring. This mixture was pipetted to the top of the media bed and allowed to settle into the bed for a few seconds before pipetting in ~3 mL of phosphate buffered solution. The column was reconnected to Tygon tubing and opened for initiate flowthrough. The start time was noted for the determination K_D values. Chicken egg albumin was measured separately by running a 90 µL sample of BD-DNP-Asp followed by a sample of Chicken egg albumin, noting the start time of each loaded sample for the determination of K_D. Each sample was measured a minimum of three times.

3. PHOSPHORYLATION OF RECOMBINANT P53(1-93) USING DNA-DEPENDENT PROTEIN KINASE

3.1 Introduction

It is known that phosphorylation is widely used in biological systems as a regulatory control mechanism for protein function[66]. Most phosphorylation sites are located in regions of proteins that are intrinsically disordered[67]. The mechanisms whereby phosphorylation regulates the properties of disordered protein structures are, unfortunately, not understood. The purpose of this project is to develop a system for producing recombinant protein that is intrinsically disordered and can be phosphorylated at amounts and purity levels that are sufficient for structural analysis. The intrinsically disordered N-terminal region of p53 tumor suppressor protein, p53(1-93), was chosen as the experimental model because it is indeed intrinsically disordered[46] and can be recombinantly expressed and recovered from cellular lysate at amounts and purity levels that allow for structural characterization[19,21,68]. Recombinant p53(1-93) also contains serine and threonine positions that map to sites in the unaltered wild type protein that are phosphorylated *in vivo* to regulate p53 activity[10,69].

To attempt to phosphorylate recombinantly-expressed p53(1-93), DNA-dependent protein kinase (DNA-PK) was used. DNA-PK is a commercially available protein kinase that is known to phosphorylate p53 and other regulatory proteins on serine or threonine residues that are followed in sequence by a glutamine (SQ or TQ motif) [58,59]. Recombinant p53(1-93) contains two SQ motifs, located at positions 15 and 37. In vitro, DNA-PK has been shown to also phosphorylate serine and threonine residues at non-SQ or TQ positions that are followed by a hydrophobic residue such as tyrosine, leucine,

alanine, proline, or methionine [60,61,59]. Recombinant p53(1-93) contains four serine residues that are followed in sequence by a hydrophobic amino acid at positions 9, 3, 45, and 50, as well as one threonine followed by a phenylalanine at position 18. As the name of the enzyme implies, the activity of DNA-PK is dependent of the presence double-stranded DNA, so commercially available sonicated calf thymus DNA is added to samples of recombinant protein to activate the enzyme and induce phosphorylation.

Besides using recombinant p53(1-93) derived from the wild type (WT) sequence, two substitution variants, Ala^- p53(1-93), and Pro^- p53(1-93), were expressed and isolated for our phosphorylation trials. The Ala^- p53(1-93) and Pro^- p53(1-93) variants each have all alanine or all proline residues substituted for glycine, respectively. For unknown reasons, substituting these residues for glycine causes greater expression levels and concomitant higher yield when compared to recombinant WT p53(1-93) expression and recovery. Differences in the expression levels can be seen in Figure 3.1 and Figure 3.2 shown later in this chapter. Higher yields were desired in order to obtain recombinant protein at levels that would allow for many phosphorylation experiments with DNA-PK.

3.2 Expression and Purification of Recombinant p53(1-93)

Expression of recombinant p53(1-93) in *Escherichia coli* cells was done using a plasmid-based *lac* expression system that utilizes isopropyl-β-1-thiogalactopyranoside (IPTG) for induction. The plasmid is a pUC derived plasmid vector with a high copy number of ~500-700 plasmids per cell. The induction of this plasmid uses a T7 phage RNA polymerase which is highly specific and does not interact with native cellular machinery. BL21 DE3 *E. coli* cells were used that have compatibility with *lac* induction systems by carrying the gene for T7 RNA polymerase under control of the *lac* promoter and a higher stability of expressed proteins from the deficiency of proteases.[70] IPTG is a molecular mimic of the lactose metabolite, allolactose, which binds to the lac repressor protein. This binding causes a conformational change in the lac repressor protein that promotes dissociation from the *lac* operon DNA and allows transcription to proceed. Unlike allolactose, IPTG is not a natural metabolite of *E. coli* and therefore is not broken down and remains at a constant cellular concentration to keep transcription running.

To facilitate the recovery of recombinant p53(1-93) from *E. coli* lysate, a 6x histidine tag (His tag) was inserted at the N-terminal end of p53(1-93) and each glycine substitution variant. His tags have high affinity for Ni^{2+}, allowing simple recovery by nickel affinity chromatography as a purification step. Nickel affinity chromatography is a form of immobilized metal ion chromatography (IMAC) that uses the strong affinity between Ni^{+2} and histidine as the basis for purification. The 6x His tag on the recombinant proteins binds to the immobilized Ni^{2+} media in a column that is hooked up to a chromatograph, which utilizes tryptophan absorbance at 280 nm to show when recombinant protein is eluted through the column. p53(1-93) contains three tryptophan

residues. Bacterial proteins that do not bind Ni^{2+} are observed to flow through the column immediately and cause an initial, large increase in absorbance as shown in Figures 3.1A and 3.2A. The chromatography beads are then washed with a low concentration NaCl buffer to elute any proteins that may bind non-specifically to the column. Finally, proteins with high Ni^{2+} affinity, such as those containing His tags, are eluted by a wash buffer that has high imidazole concentration. Imidazole is structurally similar to the histidine side chain and, accordingly, out competes the His tagged proteins for Ni^{2+} binding at the high concentrations. Elution is seen as a sharp increase in absorbance from the column. Following protein elution is a shoulder in the observed absorbance, indicating imidazole has saturated the media and is now flowing directly through the column (Figure 3.2.1A and 3.2.2A).

A consensus sequence (Leu-Val-Pro-Arg-Gly-Ser) for thrombin cleavage was included in the p53(1-93) gene, after the His tag and before the protein primary sequence. This was done to separate, via thrombin digestion, the recombinant protein from its affinity tag. Recombinant protein recovered from Ni^{2+} affinity chromatography was then dialyzed into a buffer solution that was appropriate for thrombin digestion (20 mM Tris-HCl, 100 mM NaCl, pH 8).

Taking advantage of the high acidity of p53(1-93) and the low solubility of thrombin at solution pH less than 5, anion exchange chromatography was utilized to further purify digested p53(1-93) by separating it from thrombin, the cleavage products, and any trace amounts of remaining bacterial proteins. Anion exchange chromatography takes advantage of the surface charge of proteins for separation. In this case, p53(1-93), which has a net charge of -15 from 17 glutamic and aspartic residues and only 2 lysine

and arginine, will strongly bind to positively charged chromatography beads. The p53(1-93) primary sequence has no histidine, tyrosine, or cysteine residues to contribute to its net charge. Q sepharose, which uses a quaternary ammonium functional group to supply positive charge, was selected as the anion exchange media. Chromatography was performed at pH 4.8 to precipitate thrombin and yet remain at a high enough solution pH to ensure that glutamic and aspartic side chains are deprotonated. To elute bound protein from the column, first a somewhat low ionic strength (50 mM NaCl) was used to remove weak ionic bonding contaminants, and then a final ionic strength of 250 mM was used to elute p53(1-93). The collection peak, containing digested recombinant p53(1-93), is shown in Figures 3.1B and 3.2B.

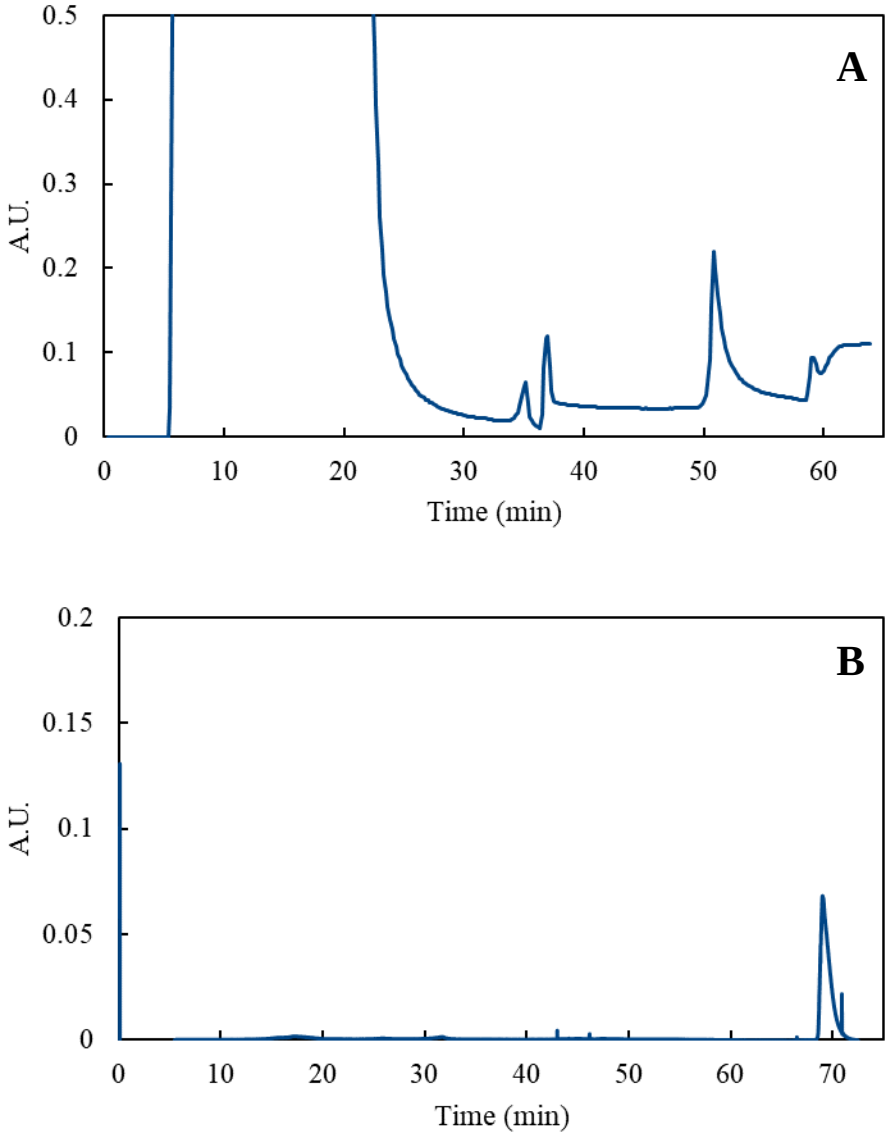

Figure 3.1 Chromatograms showing the purification of WT p53(1-93). The blue line represents absorbance at 280 nm. (A) Nickel affinity chromatography with contaminating proteins eluting at ~35 and 51 minutes, and p53(1-93) with a 6x His tag eluting from the column at 59 minutes. (B) Anion exchange chromatography, showing the elution of thrombin-digested p53(1-93) at ~70 minutes.

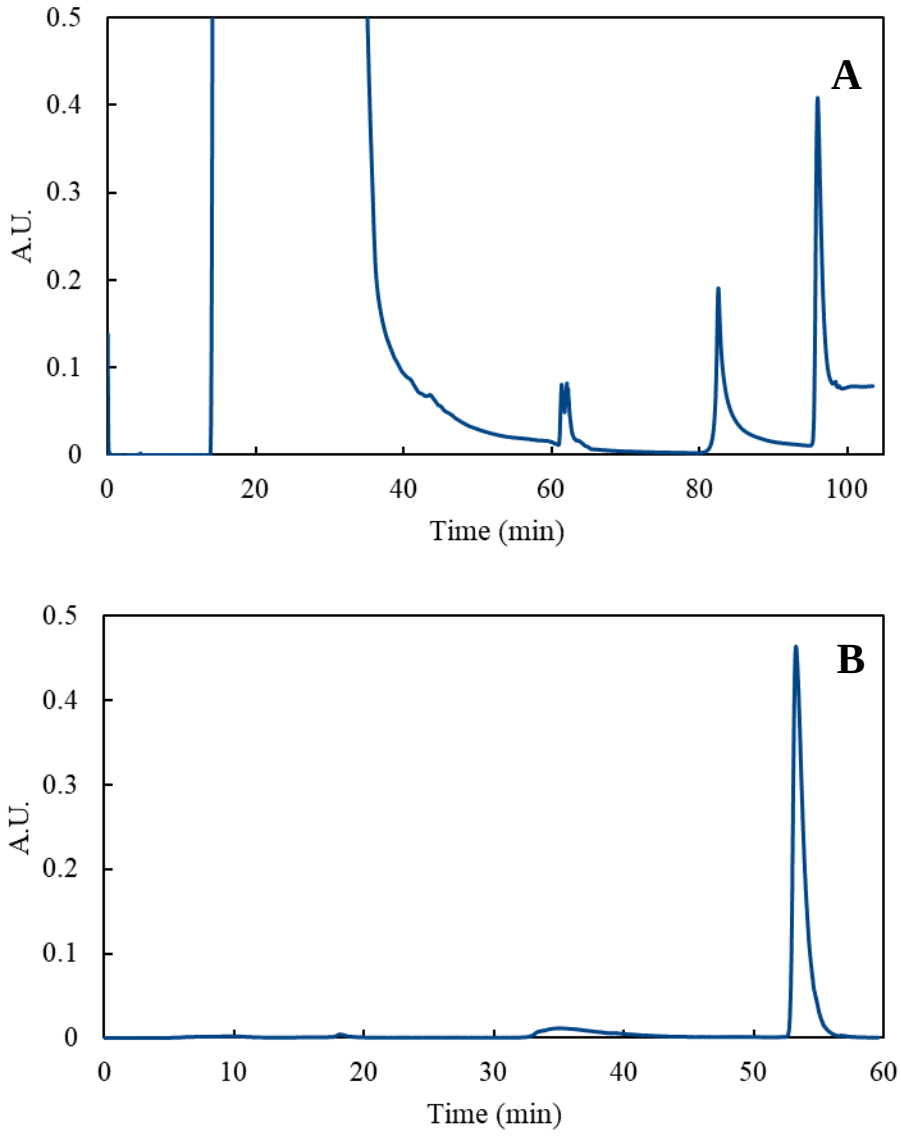

Figure 3.2. Chromatograms showing the purification of Ala$^-$ p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~60 and 85 minutes, and Ala$^-$ p53(1-93) with a 6x His tag eluting from the column at 95 minutes. (**B)** Anion exchange chromatography, showing the elution of thrombin-digested Ala$^-$ p53(1-93) at ~54 minutes.

3.3 Assessing concentration and purity of p53(1-93)

The p53(1-93) protein fragment contains three tryptophan residues that absorb UV light at a wavelength of 280 nm. The concentrations of purified recombinant p53(1-93) and the substitution mutants were quantified using absorbance spectroscopy at 280 nm. A UV spectrometer was blanked at 280 nm with 900 µL of buffer (100 mM NaCl, 10 mM Na_2HPO4, pH 7) against which purified p53(1-93) was dialyzed. A quartz cuvette with a pathlength of 1 cm was used. After the blank, 100 µL of recombinant protein were loaded into the cuvette for a x10 dilution factor, and an absorbance reading was taken. The Beer-Lambert equation, shown below, was then used to convert the measured absorbance into concentration in moles per liter units by

Equation 3.1 $$A = Cl\varepsilon \rightarrow C = \frac{A}{l\varepsilon}$$

where A is the absorbance measured at 280 nm, l is the pathlength of the cuvette in cm, ε is the wavelength-dependent molar extinction coefficient with units of M^{-1} cm^{-1}, and C is the concentration with units of molarity. The extinction coefficient was approximated as the weighted sum of the extinction coefficients at 280 nm for tryptophan (W), tyrosine (T), and cysteine (C) by

Equation 3.2 $$\varepsilon = (nW \times 5500) + (nY \times 1490) + (nC \times 125)$$

where n is the number of each residue type in the p53(1-93) sequence and the coefficients represent the extinction coefficient of each amino acid at 280 nm. Recombinant p53(1-93) has three W (i.e., tryptophan residues) and no T or C, giving an extinction coefficient of 16,500 $M^{-1}cm^{-1}$. Multiplying the molar concentration by the molecular weight gives the concentration in mg/mL units. WT p53(1-93), Ala^- p53(1-93), and Pro^- p53(1-93)

were shown to have final concentrations of 0.587 mg/mL, 0.730 mg/mL, and 2.070 mg/mL respectively. These results are given in Table 3.1.

Next, the purity of recovered recombinant protein was assessed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by Coomassie or silver staining for visualization. During SDS-PAGE, protein samples are mixed with SDS to cover the proteins with a uniform negative charge. The sample is then mixed with 2-mercaptoethanol and boiled for 5 minutes to denature protein and prevent it from aggregating. The samples were loaded into wells of a precast gel and an electric current is run through the gel. The uniformly negative proteins are electrostatically attracted to a positive electrode with smaller fragments moving through the gel faster than larger ones according to molecular weight. After a given time (~1 hr) to allow electrophoresis to proceed, the acrylamide gel is stained for visualization of recombinant and bacterial proteins.

Silver staining is a method that takes advantage of acidic side chains to visualize the protein in a gel. Samples are "fixed" into a gel by first removing SDS by a wash with 10% acetic acid and 30% ethanol solution. Then the gel is incubated in a silver nitrate solution where silver binds to acidic side chains in any proteins that are present. The bound silver ions are reduced to elemental silver with formaldehyde, causing deep brown to black bands to form where silver has been deposited. Figure 3.3 shows a developed silver stained gel in which purified WT p53(1-93) and Pro^- p53(1-93) can be seen in lanes 6 and 7 respectively.

Coomassie brilliant blue dye can also be used for the visualization of proteins in electrophoresis gels. After SDS-PAGE, gels are soaked in excess Coomassie solution

were the Coomassie binds to the protein through Van der Waals interactions with sulfonic acid groups and protein amine groups. Following de-staining of the gel in a 40% methanol and 10% acetic acid solution, protein bands become visible. Purified Ala^{-} p53(1-93) can be seen in lane 6 of Figure 3.4 and with relatively few impurities present.

Table 3.1 Final concentrations of recombinant p53(1-93) variants.

Protein	A.U.	Conc. (mol/L)	Conc. (mg/mL)	Total mg
WT p53(1-93)	0.956	$5.797 \cdot 10^{-5}$	0.587	1.173
Ala^- p53(1-93)	1.210	$7.333 \cdot 10^{-5}$	0.730	2.190
Pro^- p53(1-93) *	3.671	$2.224 \cdot 10^{-4}$	2.070	2.070

**Pro^- p53(1-93) samples were provided by Lance English*

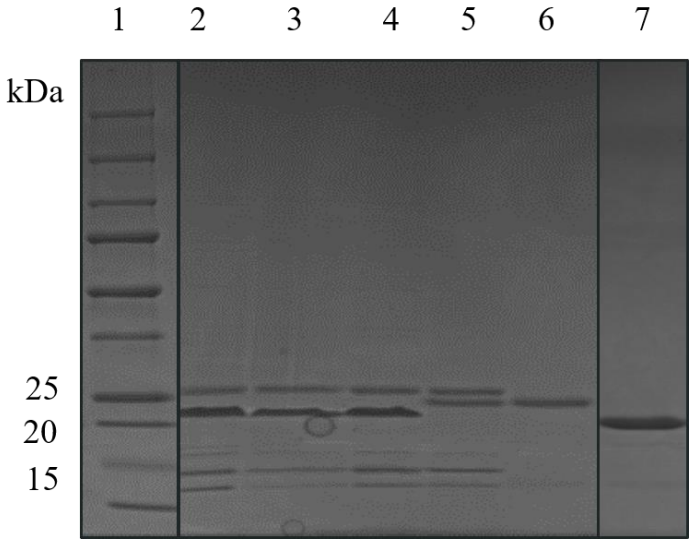

Figure 3.3. SDS-PAGE analysis of recombinant p53(1-93) and Pro^- p53(1-93) purifications. Silver-stained SDS-PAGE gel from the results of WT purification, and a purity check of the Pro^- sample. Lane 1: Protein Standards; Lane 2 - 4: Ni-NTA WT eluate samples; Lane 5: Thrombin digested WT eluate samples; Lane 6: Anion exchange WT eluate; Lane 7: Purified Pro^- was provided by Lance English.

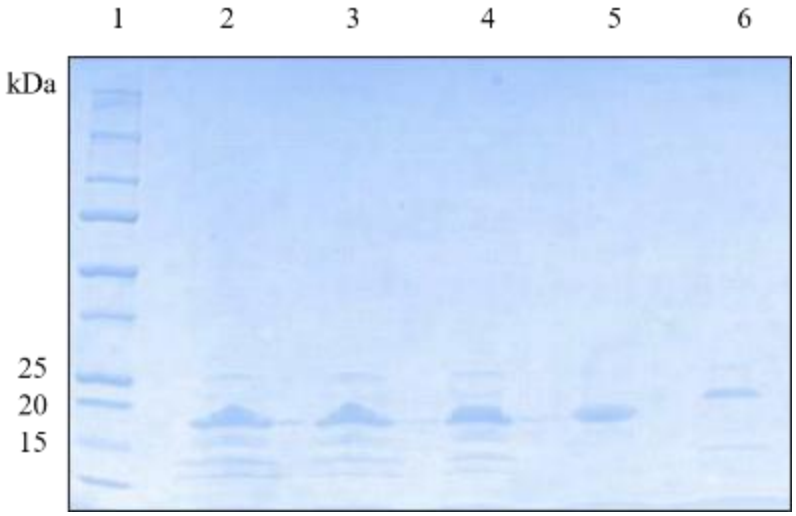

Figure 3.4. SDS-PAGE analysis of recombinant Ala$^-$ p53(1-93) purification.

Coomassie-stained SDS-PAGE gel showing results from Ala$^-$ p53(1-93) purification. Lane 1: Protein Standards; Lane 2 - 4: Ni-NTA Ala$^-$ p53(1-93) eluate samples; Lane 5: Thrombin digested Ala$^-$ p53(1-93) eluate samples; Lane 6: Anion exchange Ala$^-$ p53(1-93) eluate.

3.4 Phosphorylation of Recombinant p53(1-93)

Each recombinant p53(1-93) variant has two SQ motifs, one of which (serine 15) is known to be phosphorylated by DNA-PK *in vivo,* as well as five hydrophobic motifs that DNA-PK has been shown to potentially phosphorylate *in vitro*[58,59,71]. After purification via anion exchange chromatography, p53(1-93) and each glycine substitution variant was subjected to phosphorylation using DNA-PK and calf thymus DNA. Reactions consisted of 50mM HEPES (pH 7.5), 1mM DTT, 0.1mM EDTA, 0.2mM EGTA, 10mM MgCl2, 0.1M KCl, 1.14mM of one p53(1-93) variant at a time, 80µg/ml BSA, 0.2mM ATP and 10µg/ml linear double-stranded DNA. Each reaction was carried out for 10 minutes before being stopped by 20 µL of 30% acetic acid.

Taking advantage of the acidic nature of recombinant p53(1-93), proteins were again subjected to anion exchange chromatography, this time to separate p53(1-93) from DNA-PK. Chromatograms of the recombinant proteins that underwent phosphorylation can be seen in Figures 3.5 – 3.7. Gel images from SDS-PAGE were used to show that the purity of each sample was maintained after phosphorylation. Although there was no significant differences in the migration of p53(1-93) bands after phosphorylation for any of the purifications (Figures 3.8 and 3.9), it has been shown that the electrophoretic mobility of IDPs is often not indicative of molecular weight[19]. Purity of recombinant WT and Pro^- p53(1-93) after attempted phosphorylation is shown in Figure 3.8, and Ala^- p53(1-93) purity is shown in Figure 3.9.

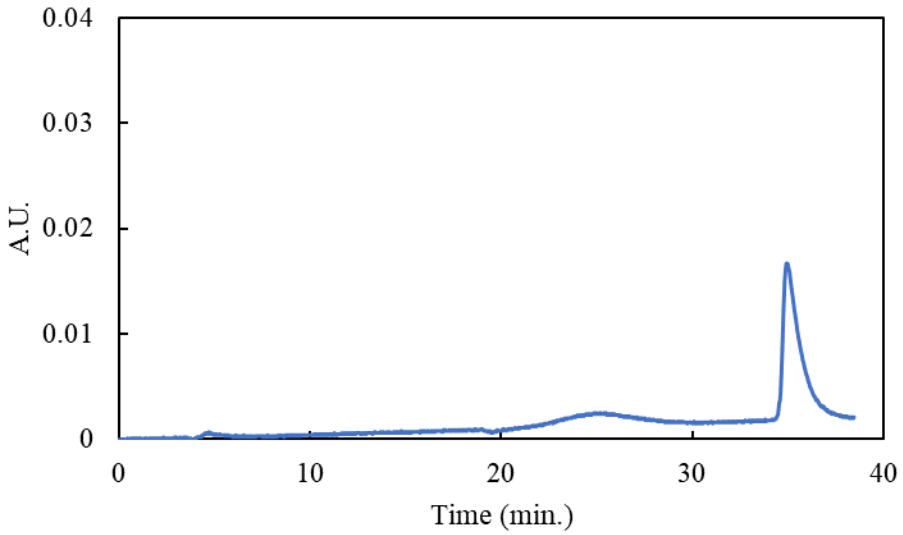

Figure 3.5. Anion exchange chromatography of phosphorylated WT p53(1-93).

Addition of elution buffer led to the elution of the target protein at around 36 minutes.

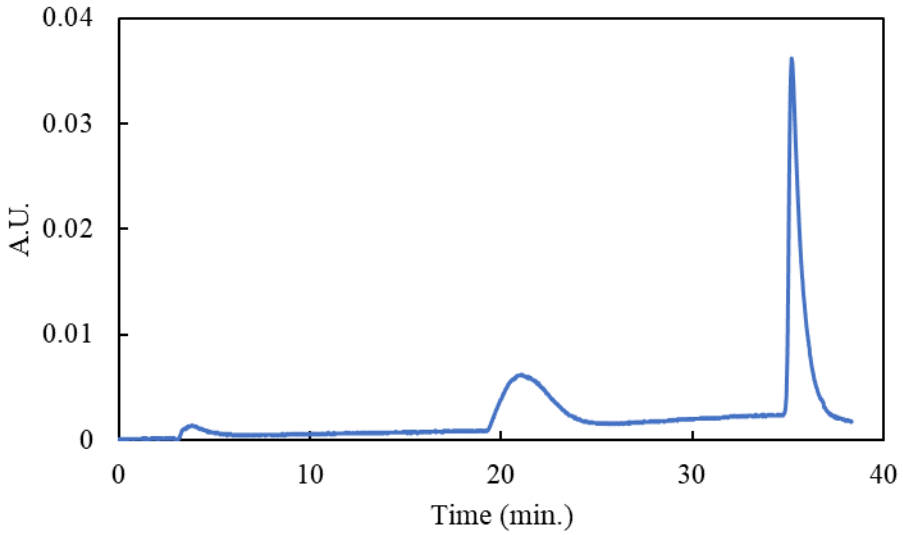

Figure 3.6. Anion exchange chromatography of phosphorylated Ala^{-} p53(1-93).

Addition of elution buffer led to the elution of the target protein at around 36 minutes.

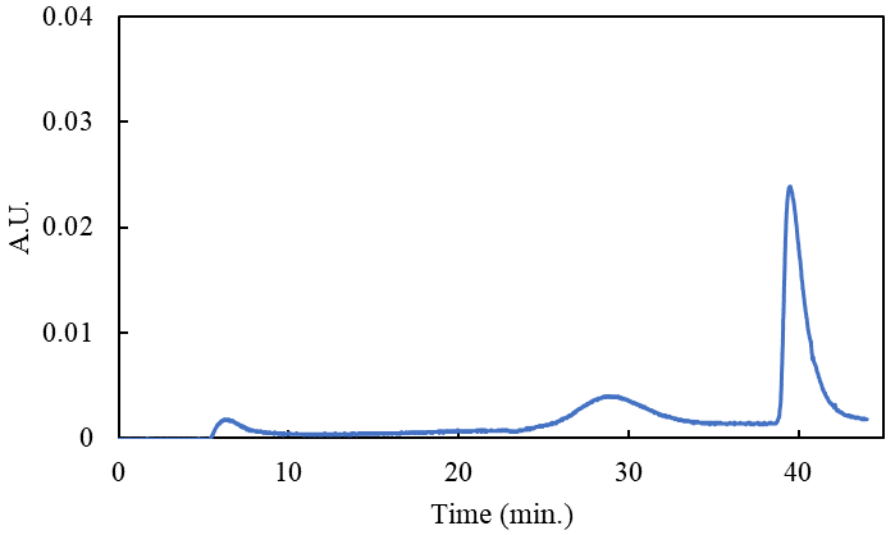

Figure 3.7. Anion exchange chromatography of phosphorylated Pro$^-$ p53(1-93). Addition of elution buffer led to the elution of the target protein at 40 minutes.

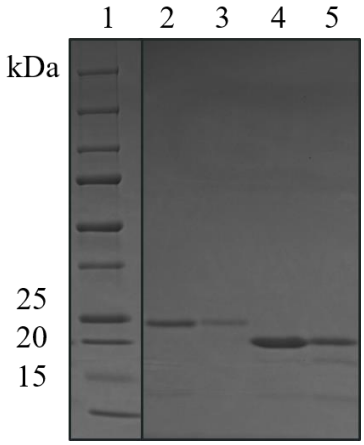

Figure 3.8. WT p53(1-93) and Pro$^-$ p53(1-93) SDS-PAGE of Phosphorylated and non-Phosphorylated forms. Silver-stained SDS-PAGE gel from WT phosphorylation, and Pro$^-$ phosphorylation. Lane 1: Protein Standards; Lane 2: Anion exchange WT eluate; Lane 3: Anion exchange phosphorylated WT eluate; Lane 4: Pro$^-$ sample; Lane 5: Anion exchange phosphorylated Pro$^-$ eluate.

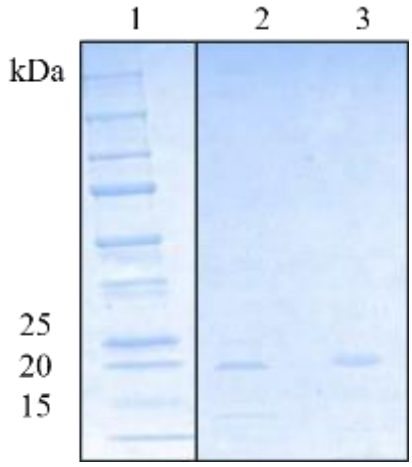

Figure 3.9. Ala$^-$ p53(1-93) SDS-PAGE of Phosphorylated and non-Phosphorylated forms. Coomassie-stained SDS-PAGE gel from Ala$^-$ phosphorylation. Lane 1: Protein Standards; Lane 2: Anion exchange Ala$^-$ eluate; Lane 3: Anion exchange phosphorylated Ala$^-$ eluate.

3.5 Structural Characterization by Size Exclusion Chromatography

Size exclusion chromatography (SEC) can be used to determine the molecular weights of proteins[72,73]. IDPs such as p53(1-93) often have elongated structures and their molecular weights can be grossly overestimated from SEC-based methods[73]. However, SEC can also be used to measure the hydrodynamic radius (R_h) of a protein, which is the radius of a hard sphere that diffuses at the same rate as that protein[73]. It is known that changes in net charge can increase the mean R_h of IDPs[74]. For that reason, SEC was chosen to investigate changes in mean R_h as an indicator of a successful phosphorylation event, since each phosphorylation event should change the net charge of p53(1-93) by -2.

During SEC, p53(1-93) and the variants were each eluted with indicator dyes, blue dextran and 2-4-dinitrophenyl-L-aspartic acid (DNP aspartate), that were used to determine the void and total column volumes. The column was packed with porous G-75 sepharose beads. Blue dextran elutes first because of its very large size (>2000 kDa) to give the void volume, V_o, followed by p53(1-93), giving its elution volume, V_e, and lastly DNP aspartate, with the smallest size (~300 Da), gives the total volume, V_t. These volumes are used to determine the equilibrium distribution coefficient (K_d) by

Equation 3.3

$$K_d = \frac{V_e - V_O}{V_t - V_O}$$

Distribution coefficients are mostly independent of the column dimensions and allow comparisons to be made from measurements that used different columns. The K_d values can be used to estimate the mean R_h of a protein when plotted along with globular protein controls that have known mean R_h and using the observed linear correlation[72,73]. This is shown in Figures 3.10 and 3.11.

We can typically resolve differences in mean R_h of ~0.5 Å using SEC, based on the standard deviation in experimental K_D[22], shown in Table 3.2 for the proteins in our study. SEC at pH 7 revealed that WT p53(1-93) had an estimated mean R_h from its K_D of 32.0 Å, with Ala^- at 28.8 Å, and Pro^- at 26.8 Å. The K_D of each protein was measured in triplicate or more. After phosphorylation, the SEC-estimated mean R_h values were 31.1 Å, 29.5 Å, and 27.2 Å, respectively, for these proteins. The results are summarized in Table 3.2 and Figure 3.3. Ala^- and Pro^- phosphorylation seemed to cause an increase in mean R_h, matching the expected outcome. WT p53(1-93), however, decreased in mean hydrodynamic size, which is not consistent with the idea that it had been phosphorylated.

Table 3.2. SEC parameters for recombinant p53(1-93) and globular proteins. Molecular weight for each protein is calculated from sequence, except for the phosphorylated versions (P).

Protein	$K_{D, mean}$	$K_{D, sd}$	M.W. (Da)	R_h (Å)
Chicken egg Albumin	0.173	0.002	42861.3	-
p53(1-93) WT	0.243	0.003	10123.3	32.0
(P) p53(1-93) WT	0.261	0.007	-	31.1
(P) p53(1-93) Ala^-	0.294	0.002	-	29.5
p53(1-93) Ala^-	0.309	0.002	9954.9	28.8
Bovine Carbonic Anhydrase	0.328	0.009	28982.5	-
(P) p53(1-93) Pro^-	0.341	0.001	-	27.2
p53(1-93) Pro^-	0.350	0.001	9241.8	26.9
Horse Heart Myoglobin	0.441	0.003	16951.4	-
Staphylococcal Nuclease	0.478	0.001	16811.2	-

Table 3.3. Comparison of SEC-measured mean hydrodynamic size for phosphorylated and non-phosphorylated p53(1-93)

Protein	$R_{h, non\text{-}Phosphorylated}$ (Å)	$R_{h, Phosphorylated}$ (Å)	ΔR_h (Å)
p53(1-93) WT	32.0	31.1	-0.9
p53(1-93) Ala^-	28.8	29.5	+0.7
p53(1-93) Pro^-	26.9	27.2	+0.3

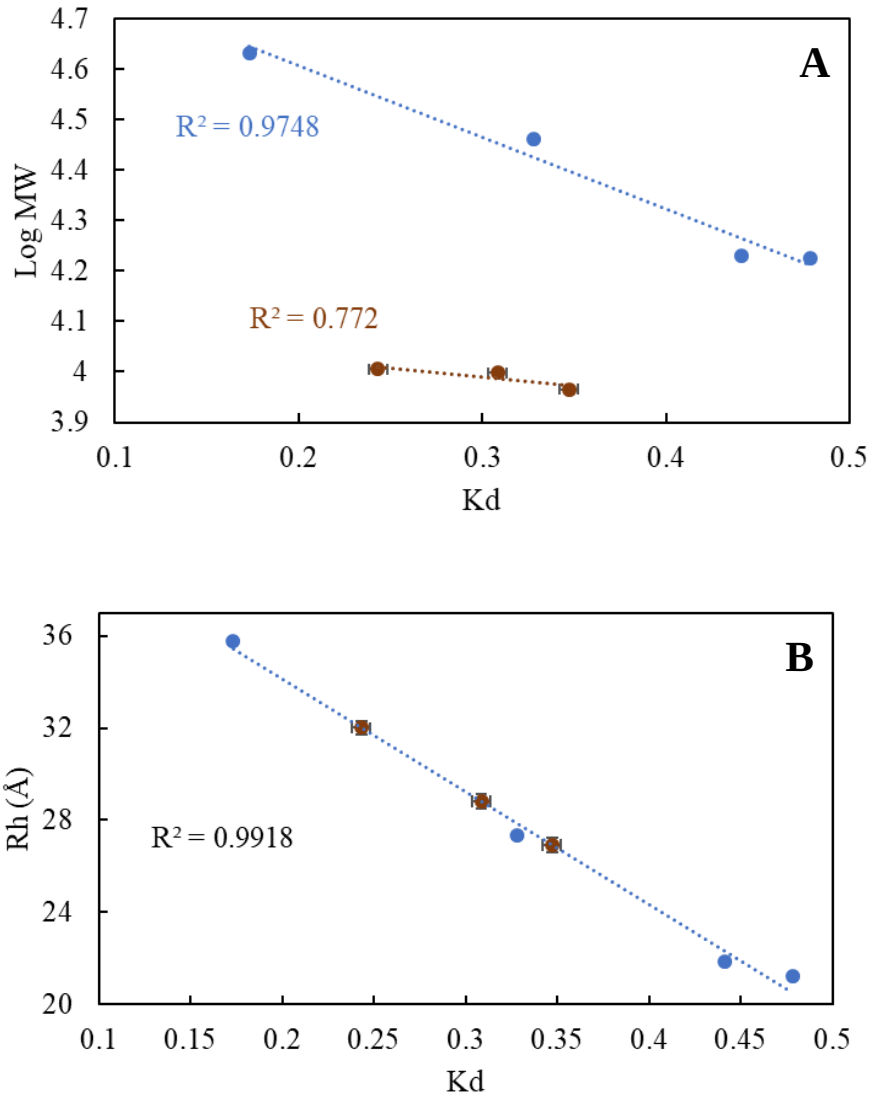

Figure 3.10. Comparison of K_D to molecular weight and apparent mean R_h for recombinant p53(1-93). Blue dots represent globular protein standards, while brown dots represent disordered recombinant p53(1-93) proteins. **(A)** Log MW as a function of k_d. Regression lines show that of globular and intrinsically disordered proteins are structurally different. **(B)** Mean R_h of disordered proteins from the linear extrapolation of the trend for globular proteins. Globular proteins used crystallographic structures to obtain R_h. For this set of globular proteins, we showed that crystal structure calculated R_h matches the mean R_h that is measured using dynamic light scattering techniques[75–78].

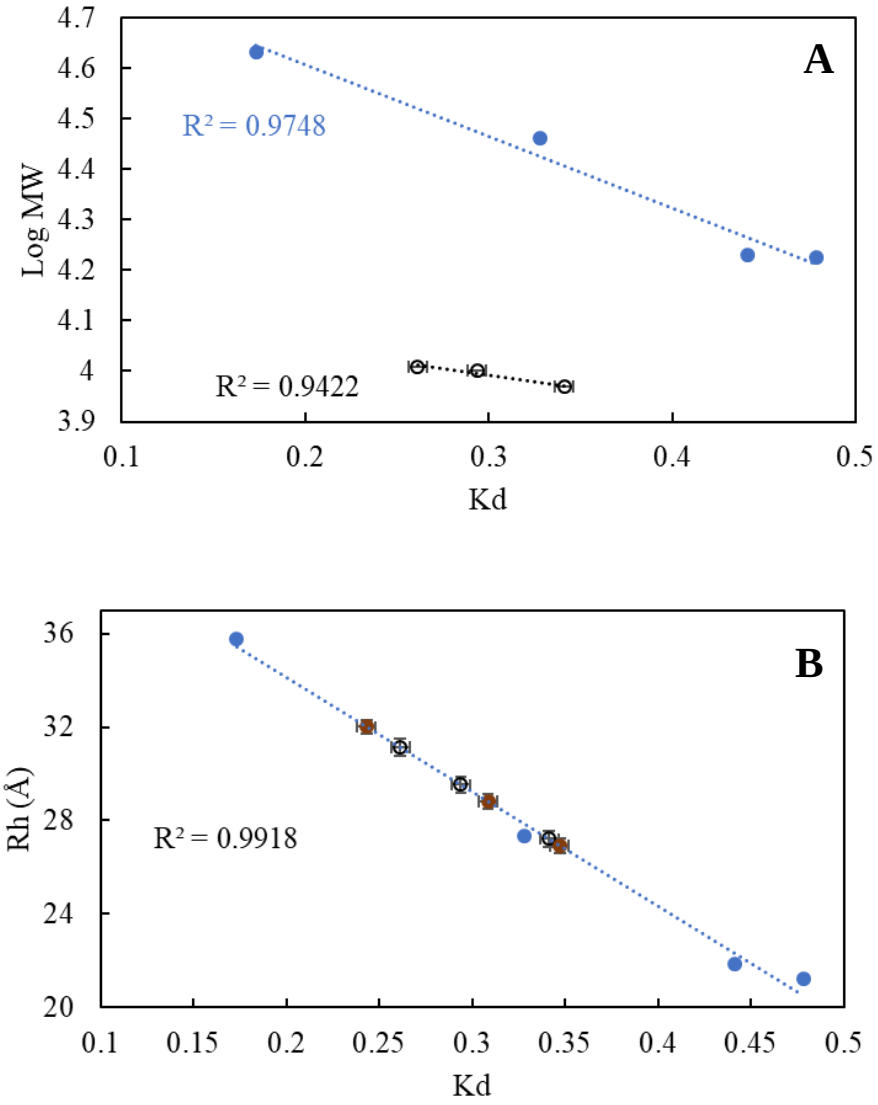

Figure 3.11. Comparison of K_D to molecular weight and apparent mean R_h for phosphorylated p53(1-93). Blue dots represent globular protein standards, while black circles represent phosphorylated recombinant p53(1-93) proteins, and brown dots represent non-phosphorylated recombinant p53(1-93). (A) Log MW as a function of kd. Regression lines show that of globular and phosphorylated IDPs remain structurally different. (B) Mean Rh of phosphorylated IDPs from the linear extrapolation of the trend for globular proteins. Globular proteins used crystallographic structures to obtain Rh. Non-phosphorylated Rh values shown by brown dots are to display average shift in Rh.

3.6 Structural Characterization by Circular Dichroism

Circular Dichroism (CD) is a spectroscopic technique in which circularly polarized light is used to determine the structural properties of a molecule containing chiral groups[79]. Amino acids in a polypeptide chain each have at least one chiral center, except for glycine, and absorb left and right circularly polarized light in different amounts to cause a rotation in the plane of polarization of the superposition of light that has been transmitted through the sample. This rotation in the plane of polarization is referred to as the Molar Residue Ellipticity (MRE) when normalized according to the molar concentration of protein in the sample and the number of amino acid residues in its primary sequence. Accordingly, the MRE value at any given wavelength represents the CD signal per peptide unit. Based on the structural conformation of the protein, the measured MRE shows characteristics in the CD spectrum that are specific for secondary structures that are present such as α-helices, β-sheets, β-turns, and PP_{II}[80]. Because of this, the secondary structure content of a protein can be estimated using its CD spectrum.

Intrinsically disordered proteins tend to report CD spectra that have MRE at 221 nm that are close to zero[81,82]. This is observed in the spectrum of recombinant p53(1-93), given in Figure 3.12, which also shows a local maximum at ~221 nm (panel B) that correlates with the presence of PPII structures[81,82]. The intensity of this local maximum is linearly dependent on temperature (panel C). As the PPII bias is driven by a favorable enthalpy[83,84], the effect of increased temperature will be to populate non-PPII states at the expense of PPII, causing the decrease in signal intensity. Also, since the PPII bias is locally driven[83,85] and noncooperative[19], the temperature effect should be broad and linear. The temperature dependence to the WT p53(1-93) CD spectrum shows an

isochromatic point between 209.5 nm and 210.5 nm, which is characteristic of the PPII-random coil transition[86]. The local minimum at 200 nm is another feature that is shared with disordered proteins[87].

The CD spectrum of phosphorylated WT p53(1-93), given in Figure 3.13, is vastly different from the spectrum of non-phosphorylated WT. The temperature dependent local maximum at 220, the local minimum at 200, and the isochromatic point between 209.5 and 210.5 are not present anymore. No other characteristic spectra are present, so the likely cause of the drastic change is impurities in the sample after attempted phosphorylation, a hypothesis that is supported by the results of electrospray ionization mass spectroscopy performed on phosphorylated WT p53(1-93).

Shown in Figure 3.14.A, the CD spectrum of Ala$^-$ p53(1-93) consists of a local maximum at around 221 nm that is linearly dependent on temperature and an isochromatic point between 209.5 nm and 210.5 nm, both of which indicate PP_{II} structure. There is also a local minimum at 200 nm in the Ala- p53(1-93) spectrum that is commonly observed in disordered proteins[87]. The local CD maximum at 221 nm decreased in intensity compared to the WT p53(1-93) spectrum, which can be accounted for by the substitutions of alanine for glycine in the Ala- variant, since alanine is known to have a much higher propensity for PP_{II} structure than glycine[68,88].

The CD spectrum of phosphorylated Ala$^-$ p53(1-93) is similar to the spectrum of non-phosphorylated Ala$^-$, with the only difference being the local maxima is shifted slightly upfield to 222 nm and the magnitude of this peak is decreased. This decrease in peak height may be interpreted as a redistribution of statistical coil throughout the protein

indicating the possibility of a structural change caused by phosphorylation, however more studies are needed[89]. The isochromatic point remains between 209.5 nm and 210.5 nm.

The CD spectrum of Pro$^-$ p53(1-93) consists of a local maximum peak at around 221 nm that is linearly dependent on temperature and an isochromatic point between 209.5 nm and 210.5 nm, both of which indicate PP_{II} structure. The local minimum at 200 nm in the Pro- p53(1-93) spectrum is similar to what was observed in the WT and Ala$^-$ p53(1-93) spectra, indicating the protein is disordered[87]. The local CD maximum at 221 nm, and the overall PP_{II} propensity decreased in intensity compared to WT and Ala$^-$ p53(1-93) spectra, which can be accounted for by the substitutions of proline for glycine in the Pro- variant, since proline is known to have a much higher propensity for PP_{II} structure than either alanine or glycine[68,88].

The CD spectrum of phosphorylated Pro$^-$ p53(1-93) is similar to the spectrum of phosphorylated Ala$^-$ p53(193) spectra. Once again, going from the non-phosphorylated to the phosphorylated spectrum, there is a decrease in intensity of the local CD maximum at 221 nm. As with Ala$^-$, this shift may be interpreted as a redistribution of statistical coil throughout the protein, indicating a structural change brought on by phosphorylation[89]. The isochromatic point remains between 209.5 nm and 210.5 nm indicating the PP_{II}-coil transition remains, albeit with a lower PPII content when compared to the non-phosphorylated protein.

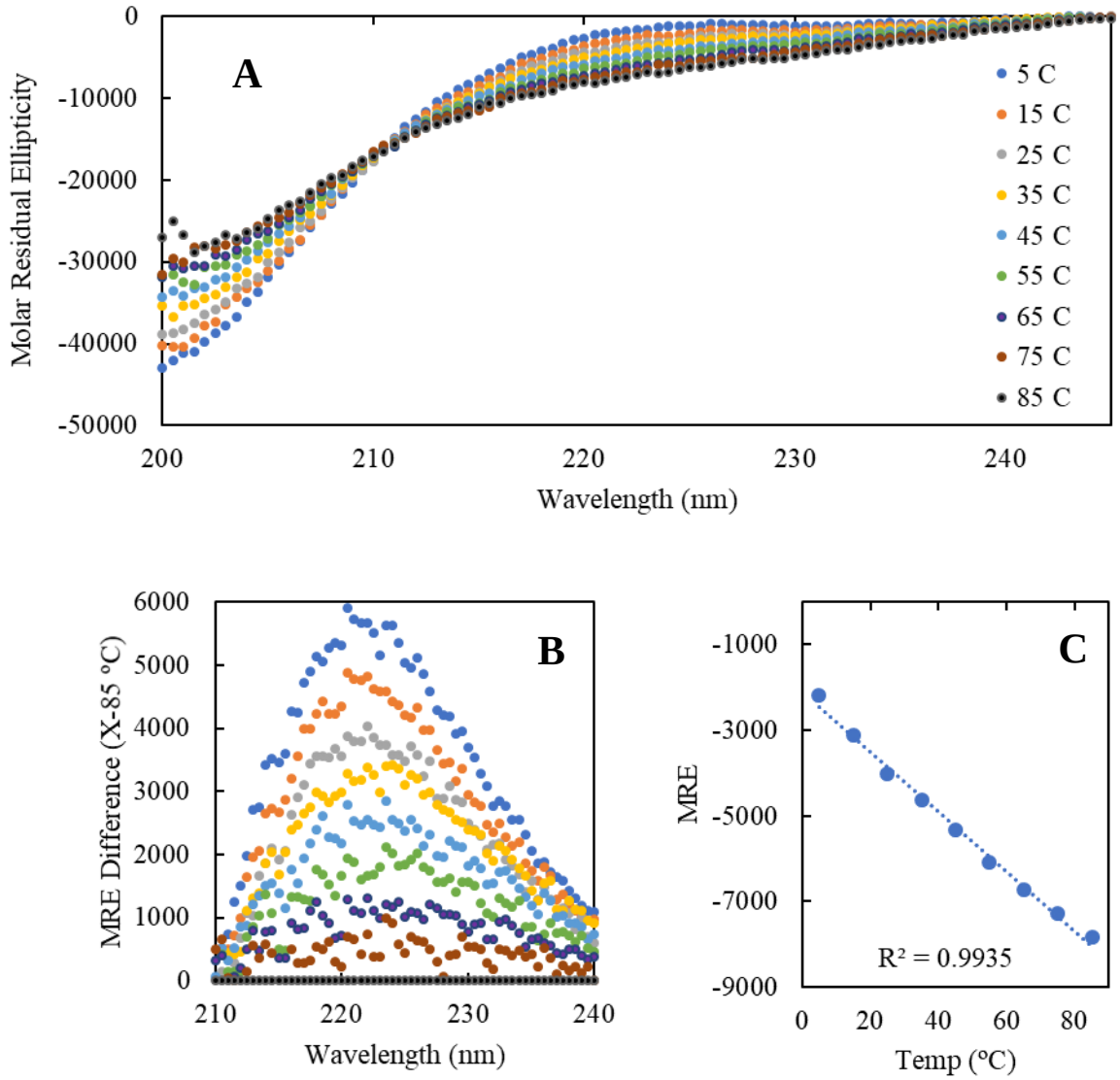

Figure 3.12. CD spectrum of WT p53(1-93) at pH 7.0. (A) CD spectrum was measured at eight temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red), 85 ºC (black). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 85 ºC spectrum, to emphasize the local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

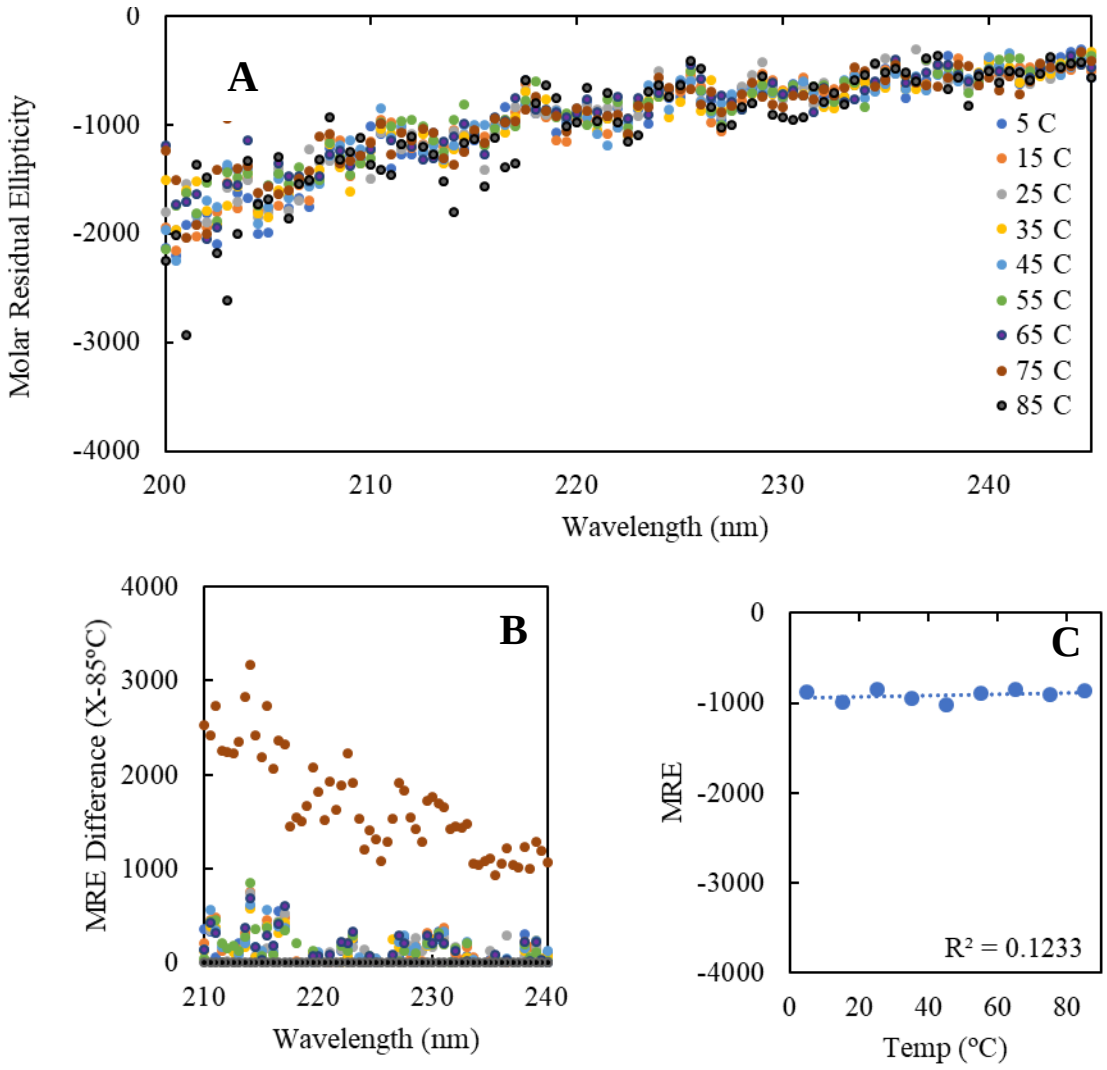

Figure 3.13. CD spectrum of phosphorylated WT p53(1-93) at pH 7.0. (A) CD spectrum was measured at eight temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red), 85 ºC (black). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 85 ºC spectrum, to emphasize the loss of local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

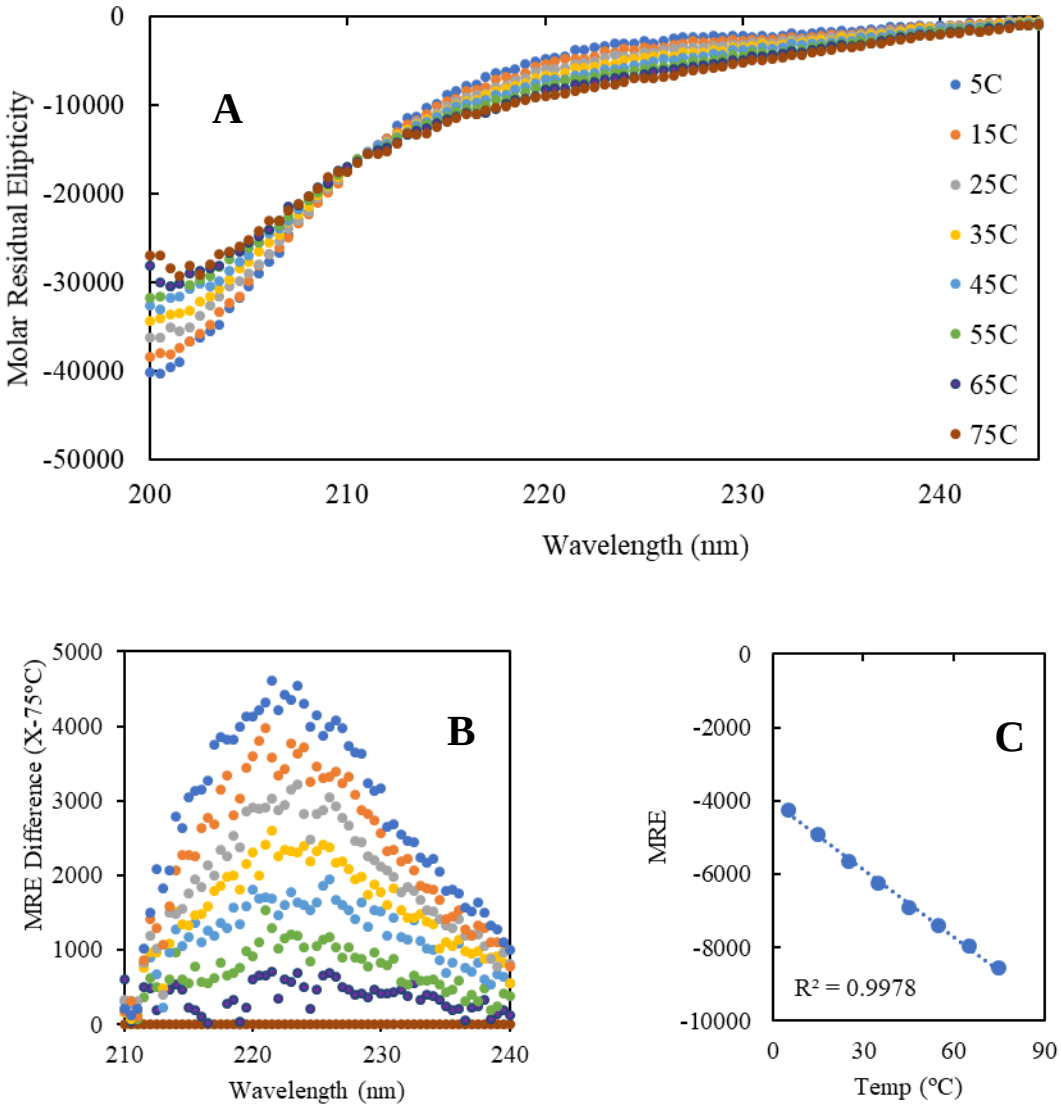

Figure 3.14. CD spectrum of Ala⁻ p53(1-93) at pH 7.0. (A) CD spectrum was measured at eight temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 75 ºC spectrum, to emphasize the local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

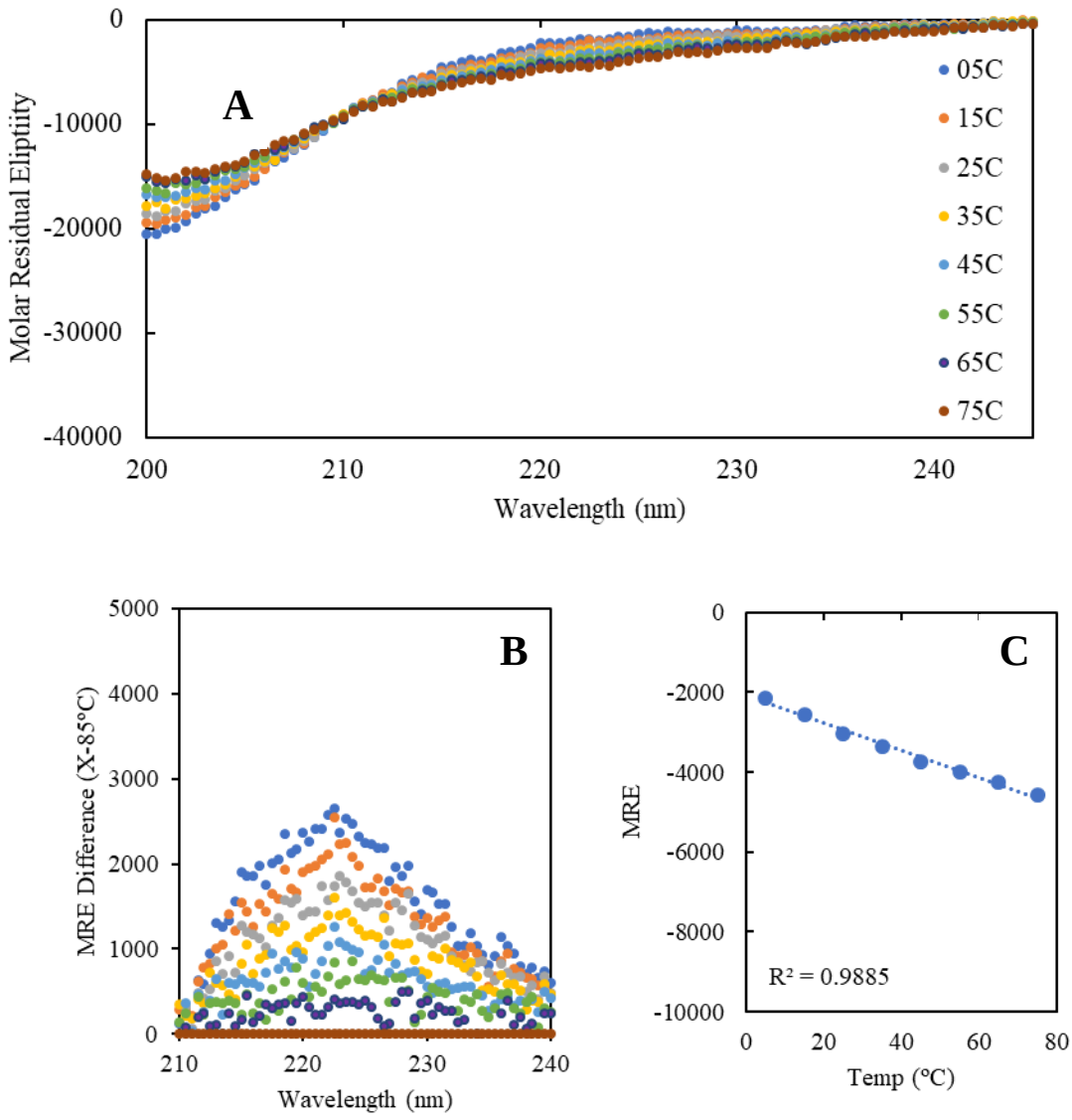

Figure 3.15. CD spectrum of phosphorylated Ala$^-$ p53(1-93) at pH 7.0. (A) CD spectrum was measured at eight temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 75 ºC spectrum, to emphasize the local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

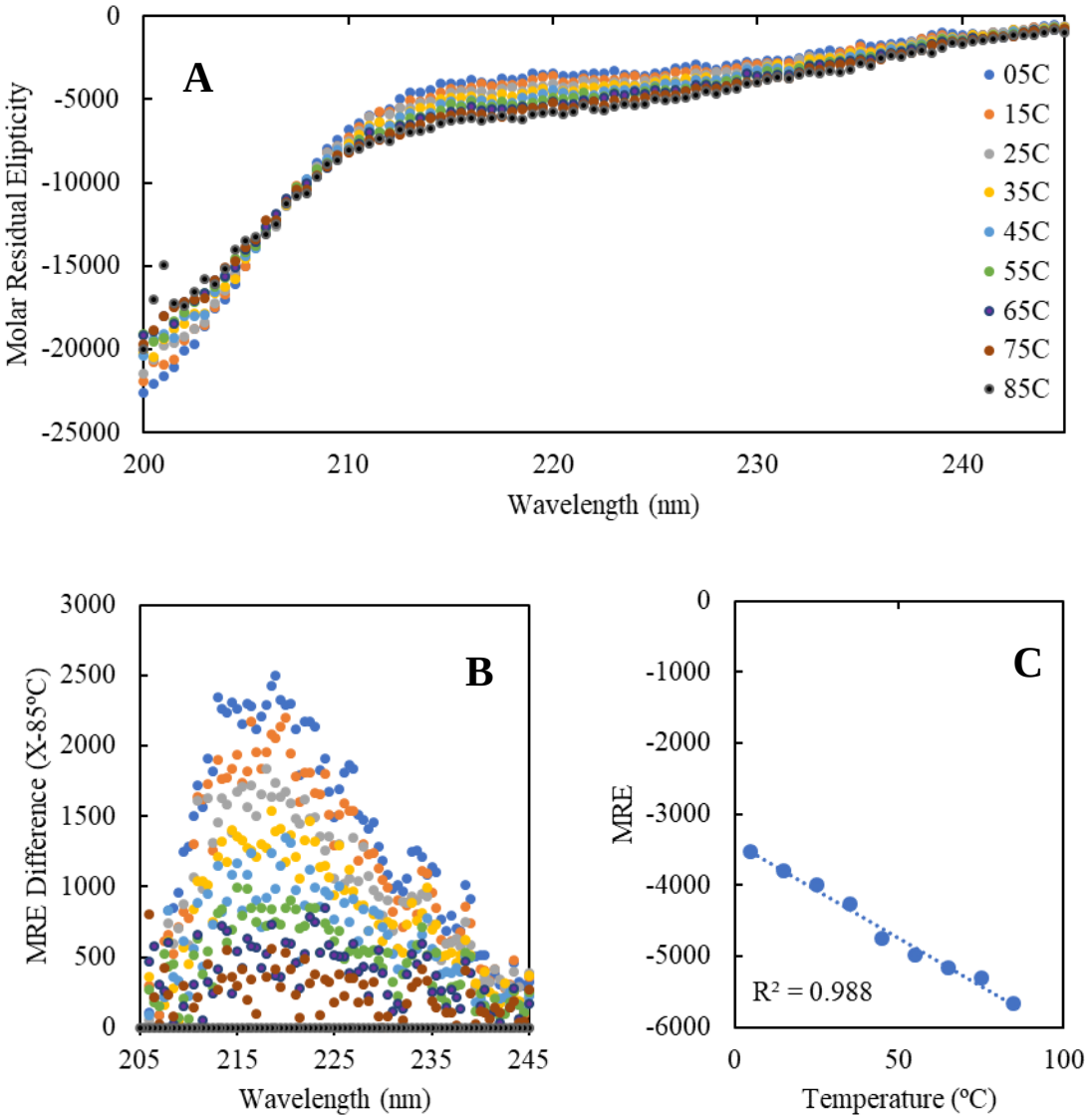

Figure 3.16. CD spectrum of Pro$^-$ p53(1-93) at pH 7.0. (A) CD spectrum was measured at nine temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red), 85 ºC (black). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 85 ºC spectrum, to emphasize the local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

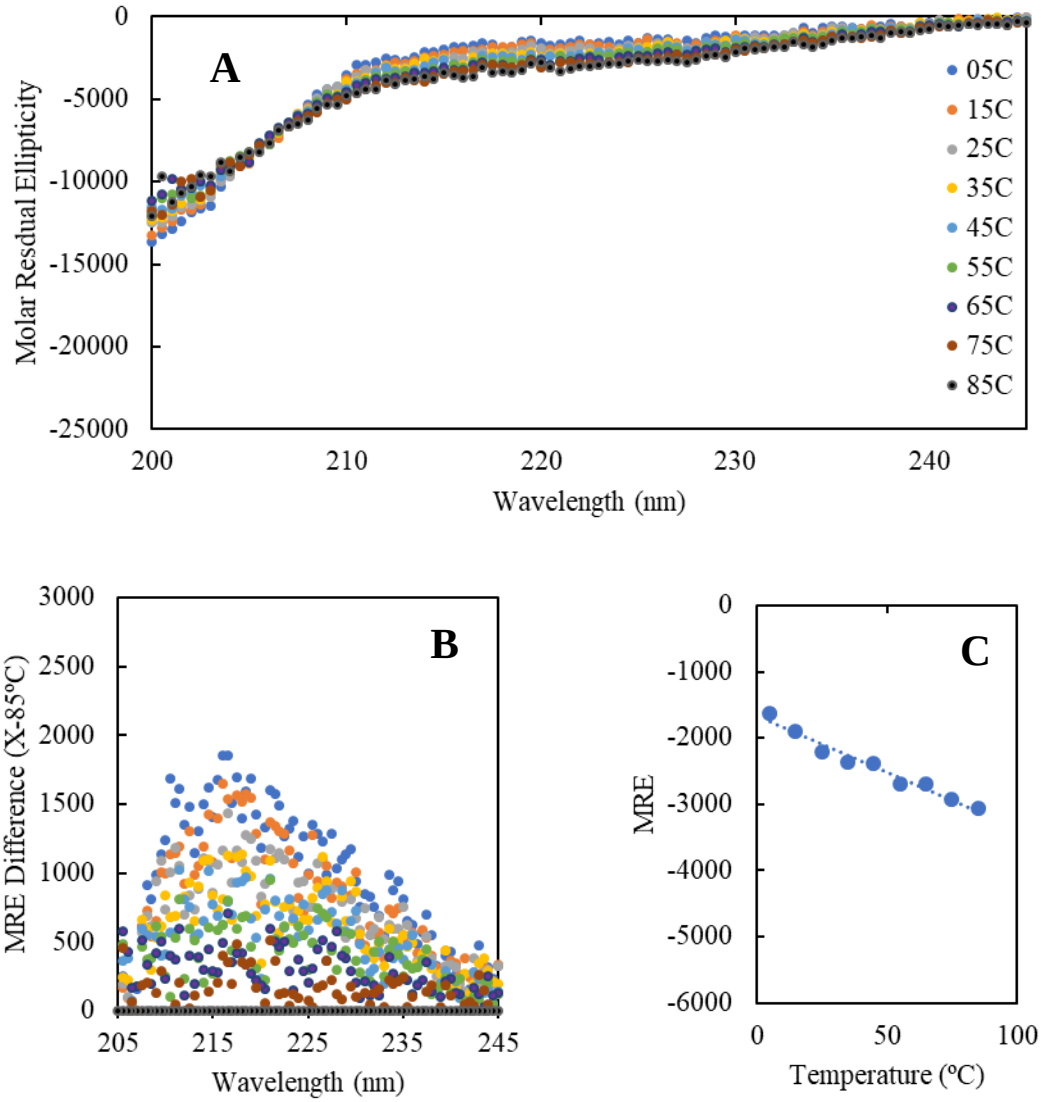

Figure 3.17. CD spectrum of phosphorylated Pro$^-$ p53(1-93) at pH 7.0. (A) CD spectrum was measured at nine temperatures: 5 ºC (dark blue), 15 ºC (orange), 25 ºC (gray), 35 ºC (gold), 45 ºC (light blue), 55 ºC (green), 65 ºC (purple), 75 ºC (brick red), 85 ºC (black). (B) The molar residue ellipticity (MRE) at each temperature, relative to the 85 ºC spectrum, to emphasize the local maximum observed at ~221 nm. (C) Temperature dependence to the MRE at 221 nm.

3.7 Characterization by ESI-MS

Electrospray ionization mass spectrometry (ESI-MS) is a technique that can be used to determine the molecular weight of a protein. During ESI, a purified protein in solvent is placed into a capillary needle before being injected into the mass spectrum apparatus. The capillary needle is surrounded by a steady voltage which applies charge to solvent droplets as the sample is propelled through the needle to create a mist. The droplets continue through an evaporation process in which the charge of the solvent is transferred to the protein. As the charges are concentrated from solvent evaporation, the protein ions are dispelled from the nanodroplets into the gas phase via electrostatic and hydrophobic effects[90]. A spectrum of the different charged ions is given in the form of mass-to-charge ratio (m/z) relative to their intensity (%) based on their flight time in a vacuum. Molecular weights of recombinant protein (M_r) can be calculated using

Equation 3.4 $$\boldsymbol{p = m/z}$$

Equation 3.5 $$\boldsymbol{p_1 = \frac{M_r + z_1}{z_1}}$$

Equation 3.6 $$\boldsymbol{p_2 = \frac{M_r + (z_1 - 1)}{z_1 - 1}}$$

where p_1 and p_2 are adjacent peaks and z_1 is the charge of peak one. To verify if the DNA-PK method for phosphorylation of WT and ALA$^-$ p53(1-93) was successful, ESI-MS was performed on the phosphorylated products. Figures 3.18 and 3.19 are spectra of phosphorylated WT p53(1-93) and phosphorylated Ala$^-$ p53(1-93), respectively. ExPasy ProtParam was used to calculate the molecular weights of each protein for comparison to what was obtained by ESI-MS, and the results are shown in Table 3.4. The ExPasy ProtParam calculated molecular weight of WT p53(1-93) without the 6 x His-tag was

calculated to be 10,123.25 Da, while Ala$^-$ without the 6 x His-tag molecular weight was calculated to be 9,954.92 Da.

An estimated molecular weight for WT p53(1-93) was calculated using the most prominent peaks of the ESI-MS spectra shown in Figure 3.18, averaging 10,230 Da, however the calculated molecular weights from each of the charge states were inconsistent with each other indicating the presence contaminants. For the Ala$^-$ p53(1-93), the calculated molecular weights from each of the charge states were consistent with each other, giving an average molecular weight of 9,964 Da. The experimental molecular weight for the Ala$^-$ p53(1-93) was only 10 Da larger than the theoretical molecular weight indicating that phosphorylation was not successful. The spectra for both samples indicated that although the correct proteins can be overexpressed and removal of the 6x His tag with thrombin digestion was successful, the DNA-PK method for phosphorylation of p53(1-93) may not be the most viable method for obtaining phosphorylated protein for structural studies. One possible improvement to this system would be the inclusion of a phosphatase inhibitor to the DNA-PK reaction, however due to the poor results of the MS characterization of phosphorylated WT and Ala$^-$ p53 using the DNA-PK method, the project was changed to test another method for phosphorylation, called the orthogonal translation system. This new project was taken on prior to MS characterization of Pro$^-$ p53(1-93).

Table 3.4. Molecular weights of WT and Ala$^-$ p53(1-93)

Protein	Theoretical MW (Da) w/ 6x His	Theoretical MW (no Phos) (Da)	Measured MW (Phos) (Da)
WT p53(1-93)	12074.41	10123.25	-
Ala$^-$ p53(1-93)	11906.08	9954.92	9958.5

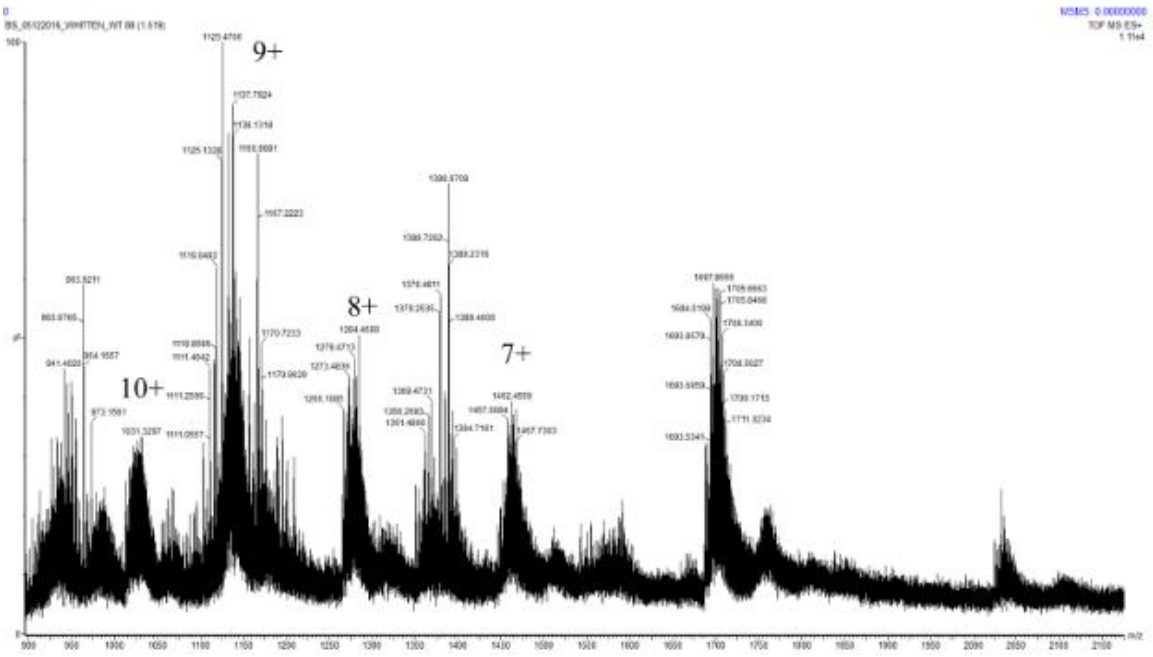

Figure 3.18 Phosphorylated WT p53(1-9) ESI-MS Chromatogram. It is obvious that each charge state is not a single peak but consists of multiple peaks due to contamination. Molecular weights for each charge state are as follows: 7+ (10,230.19), 8+ (10,227.77), 9+ (10,120.24), 10+ (10,300.32). These molecular weights are inconsistent and cannot be used to determine an accurate average molecular weight.

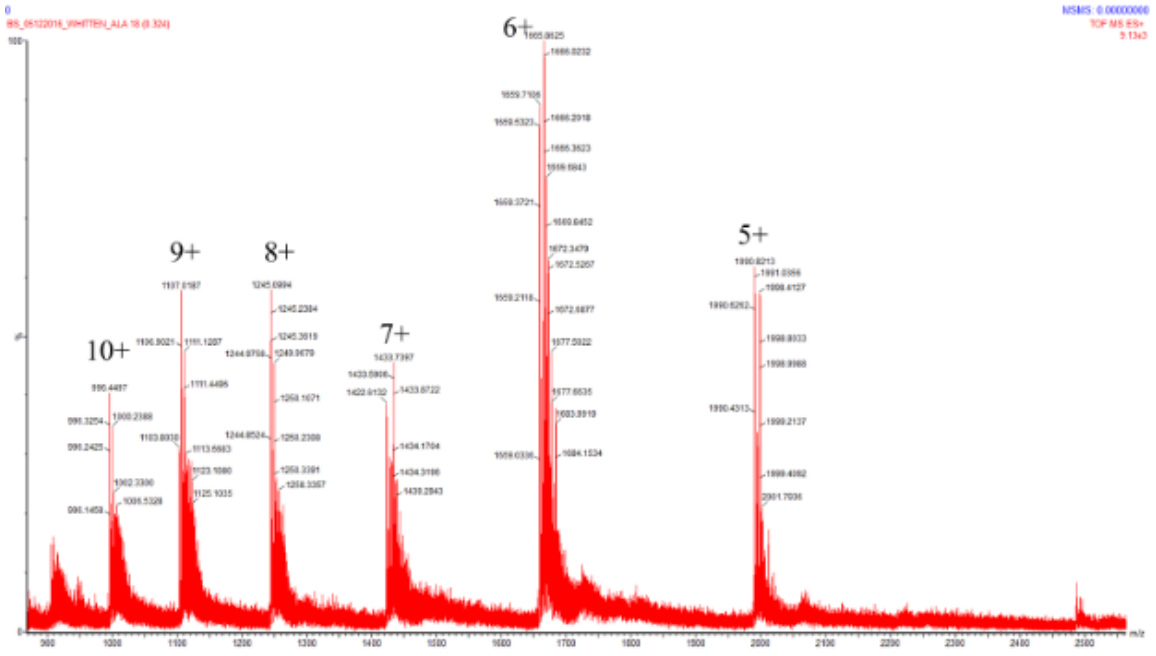

Figure 3.19 Phosphorylated Ala^{-} p53(1-9) ESI-MS Chromatogram. Each charge state gives molecular weights that are consistent with one another. Molecular weights for each charge state are as follows: 5+ (9,949.11), 6+ (9,989.175), 7+ (9,951.29), 8+ (9,952.8), 9+ (9,954.17), 10+ (9,954.5).

4. ORTHOGONAL TRANSLATION SYSTEM FOR PHOSPHORYLATION OF WT P53(1-93)

4.1 Introduction

Our efforts to phosphorylate recombinant p53(1-93) using DNA-PK were mostly unsuccessful and, moreover, did not yield protein at amounts required for structural studies. Since the DNA-PK method of phosphorylation was not successful for us, a phosphoserine (Sep) orthogonal translation system (OTS) was chosen to attempt to initiate site-specific phosphorylation of WT p53(1-93). The SepOTS utilizes two plasmids, whereby one codes for the protein and the other for transfer RNA (tRNA) that transports Sep to the ribosome and a few additional accessory proteins that allow for the use of the non-traditional tRNA. A strain of *Escherichia coli* that has the serine-phosphatase gene (SerB) deleted from its genome was also used.

The phosphoserine orthogonal translation system relies on stop-codon read-through (i.e., reading a non-sense codon as a sense codon) for the incorporation of Sep into a target protein at the UAG amber stop-codon. The first plasmid in the system, named SepOTSλ, codes for the molecular machinery needed for Sep incorporation onto the UAG sites. SepOTSλ codes for a mutant Sep-tRNA ($tRNA^{Sep}_{CUA}$) and Sep-tRNA synthetase (SepRS9) to amino-acylate Sep-tRNA. The gene coding for $tRNA^{Sep}_{CUA}$ (tRNA-SepA37) is inserted in each plasmid four times for increased production. Finally, the plasmid also codes for an elongation factor (EFSep21) that mediates introduction of the amino-acylated Sep-$tRNA^{Sep}_{CUA}$ specifically into the free site in the ribosome for Sep incorporation into recombinant p53(1-93) protein. The SepOTSλ plasmid is a pUC derived plasmid vector with a high copy number of ~500-700 plasmids per cell. The

induction of this plasmid uses a T7 phage RNA polymerase which is highly specific and does not interact with native cellular machinery. A version of the SepOTSλ plasmid, called SupD, is also included as part of the OTS. This variation of the first plasmid in the SepOTS is used to test for the incorporation of serine, rather than Sep, into a UAG modified protein transcript. This plasmid contains a specialized $tRNA^{Ser}_{CUA}$ called an amber suppressor tRNA, that tracks serine to UAG codons during translation.

The second plasmid of the system codes for Sep-p53(1-93), which is WT-p53(1-93) with UAG stop-codons substituting for the serine codons at residue positions 15, 33, and 46. These three serine positions in p53 are phosphorylated *in vivo* in response to genotoxic or oxidative stressors to induce apoptosis[10,69]. The SepOTSλ plasmid should work in conjunction with the plasmid containing the Sep-p53(1-93) gene to produce recombinant p53(1-93) that has been phosphorylated at positions 15, 33, and 46.

Two versions of the Sep-p53(1-93) plasmid were used with each having a different replicon. A replicon is the region in a plasmid that host the origin of replication (ORI) of the plasmid as well as the control elements for its replication. Although the replicons of plasmids are different than that of their host, they rely on host machinery to regulate their replication. One version of the plasmid contained a high copy-number pBR322 ORI similar to that of the pUC ORI of SepOTSλ plasmid. These plasmids are both derivatives of pMB1 plasmids and, as such, are under control of the same negative control element (RNAI) as the other. RNAI is a diffusible negative replication regulator that inhibits replication in these plasmids. This replication regulator is used to maintain the copy number of the plasmid in a cell. It is thought that if plasmids are regulated by the same elements that neither plasmid will reach the optimum copy number and will

eventually separate into different cells, leading to inadequate SepOTS plasmid ratios. In spite of these limitations, this plasmid combination was tested since it has been successfully implemented by other research groups[56,62,91]. Our hypothesis was that high copy numbers would be maintained long enough for sufficient protein to be produced.

The second version of this plasmid contains a p15a ORI which is regulated by the replication initiator protein (Rep). Rep is shown to bind to iterons upstream of the ORI site to induce replication. This mechanism does not interfere with the RNAI regulation machinery of pMB1-based plasmids, so both plasmids should be maintained in the cells for expression. That being said, the p15a plasmid itself is a low-copy number plasmid, found at only 10 – 20 plasmids per cell[92], which may limit the amount of recombinant protein that is ultimately produced.

Moreover, two different *E. coli* cell lines were tried. The first was a derivative of the BL21 line that has the serine phosphatase gene (*serB*) deleted, BL21ΔserB, to prevent Sep hydrolysis. The second was C321.ΔA that has all UAG stop-codons removed from cellular DNA and also is missing the gene for release factor-1 (*RF1*), which recognizes UAG stop codons and initiates termination of translation. Removal of genomic UAG stop codons and *RF1* allows 100% efficiency of read-through. This project tested both cell lines since there are benefits and drawbacks associated with each. Though BL21ΔserB does not reach 100% incorporation efficiency, and mutant tRNA could incorporate phosphoserine into non-target positions[62], its cell growth is often better compared to C321.ΔA. The loss of *RF1* in C321.ΔA can negatively affect cell growth[93] and thus lower recombinant protein amounts are typically produced. Combinations of cell lines and plasmids tested is summed up in Table 4.1. Transformation of two plasmids into each cell

line necessitated that each cell type was made chemically competent, transformed with one plasmid, made chemically competent again, and transformed again with the second plasmid while maintaining selective pressure for cell types, and plasmids.

Table 4.1. Combinations of cell lines and plasmids used to test expression system.

Cell Line	OTS Plasmid	p53(1-93) plasmid
C321.ΔA	SepOTSλ	pBR322
C321.ΔA	SupD	pBR322
C321.ΔA	SepOTSλ	p15a
BL21ΔSerB	SepOTSλ	p15a
BL21ΔSerB	SepOTSλ	pBR322
BL21ΔSerB	SupD	p15a
BL21ΔSerB	SepOTSλ	Ala- p53(1-93)

4.2 Expression of p53(1-93) in C321.ΔA Cell Line

Expression of recombinant Sep p53(1-93) was first attempted using the pBR322 plasmid in *C321.ΔA* cells. Cell cultures were grown at 37 ºC in 4 L of 2xYT (1.0% w/v yeast extract, 1.6% w/v tryptone, 0.5% w/v NaCl, pH 7.0). 2xYT is highly enriched media for bacterial culture that uses approximately double the amounts of yeast extract and tryptone when compared to LB media, hence the name 2xYT. Induction used 1 mM IPTG and 0.2% w/v rhamnose, as suggested by the plasmid manufacturer, and was administered at mid-log growth. To isolate recombinant protein from the cell lysate, nickel affinity chromatography and anion exchange chromatography was performed as specified in Chapter 3 for WT p53(1-93). A small elution peak was observed directly before the imidazole shoulder in the Ni^{+2} chromatogram, from the elution buffer. Post-thrombin digestion, to remove the 6xHis affinity tag, the anion exchange chromatogram unfortunately displayed no discernable elution peak that would indicate Sep p53(1-93) was adequately expressed (Figure 4.1).

Recombinant expression of Sep p53(1-93) was attempted a second time with the temperature at the induction step reduced to 30 ºC. Lowering the induction temperature has been shown to increase the expression of other recombinant proteins due to increasing the protein solubility and a concomitant decrease in protein aggregation[94–97]. As before, there was a small peak directly before the imidazole shoulder from the elution buffer in the Ni^{+2} chromatogram, but the anion exchange results of the second chromatography step again reported no discernable elution peak for Sep p53(1-93) (Figure 4.2). A sample was saved from the anion exchange elution at the position that is normally observed for p53(1-93). This sample was concentrated and dialyzed. UV

spectroscopy of the dialyzed anion exchange elution sample showed a total yield of only 0.037 mg of protein was obtained. SDS-PAGE and Phos-tag SDS-PAGE were used for further analysis of the sample (Figure 4.5). Phos-Tag acrylamide is a modified version of acrylamide that can be used to detect the presence of phosphorylated protein by a change in the protein's electrophoretic mobility when compared to the non-phosphorylated protein. Phos-Tag acrylamide uses an anion trap that binds diphosphate dianion groups more so than any other anionic substituents. When casting a gel, Phos-Tag acrylamide is added to the acrylamide mixture, so that the protein and tag will interact as the protein migrates through the finished gel. If the two divalent metal ions trap phosphorylated proteins during migration slowing the migration velocity, then there should be an apparent shift in the Sep-p53(1-93) band when comparing the Phos-Tag SDS-PAGE to standard SDS-PAGE. The results of Phos-tag SDS-PAGE, compared to the SDS/PAGE analysis, indicated that no protein was present in the concentrated sample at detectable levels, as shown in lane 3 of the Figure 4.5 gel images.

A third attempt at Sep p53(1-93) expression was conducted with an additional temperature reduction during the induction step to 20 ºC. From this change, there was again a small elution peak before the imidazole shoulder in the Ni^{+2} chromatogram, and multiple small, but discernable elution peaks in the chromatogram from the anion exchange step, as shown in Figure 4.3. After concentration and dialysis of this sample, UV spectroscopy showed there was approximately 0.235 mg of protein expressed. Analysis of the sample by SDS-PAGE, shown in lane 4 of the Figure 4.5 images, gave a faint protein band where p53(1-93) is expected to appear but multiple other bands at

different molecular weights were also present suggesting that the sample was mostly protein other than Sep p53(1-93).

4.2.1 Ethanol Effects on Expression

The addition of ethanol to bacterial expression systems has been reported to increase the production of recombinant proteins. Ethanol has been shown to increase membrane fluidity which is associated with higher DnaA activity. Active DnaA initiates DNA replication. Higher membrane fluidity allows ADP to be released from the inactive ADP-DnaA complex which makes it easier for the ATP-DnaA active complex to form. It is proposed by the Chhetri group that the enhancement of DNA synthesis from adding ethanol to an expression system could lead to enhancement of plasmid DNA synthesis and thus increased protein expression. Ethanol is also known to induce a heat shock response when introduced into *E. coli* cultures at 1.0% - 5.0% v/v. The major heat shock response is the increased production of sigma factor 32 (σ^{32}). This sigma factor controls the expression of chaperone proteins that leads to greater solubility of recombinant proteins. It is also proposed that the increase in chaperones can lead to increased recombinant protein expression[98–100].

A fourth expression was also carried out with the induction temperature again at 20 ºC, but with 2.0% v/v ethanol added to the culture at the induction step. Our use of 2.0% v/v ethanol at the induction step followed the protocol of Chhetri and colleagues[36,37] who observed enhanced recombinant expression of Aβ Fusion protein, GlnRS protein, and RPB9 protein. Unfortunately, Sep p53(1-93) expression did not increase by the addition of 2.0% v/v ethanol. Again, there was a small elution peak directly before the imidazole shoulder in the Ni^{+2} chromatogram, indicating possible protein expression, as

well as a small but discernable elution peak in the DEAE chromatogram. After concentration and dialysis, approximately 0.315 mg of protein was detected in the sample by UV absorbance. Analysis of the sample by SDS-PAGE (Figure 4.5) showed expression and contamination levels that were identical to the sample from 20°C induction but without ethanol. Phos-Tag SDS-Page analysis showed no change in protein migration that could be associated with the successful expression of Sep p53(1-93).

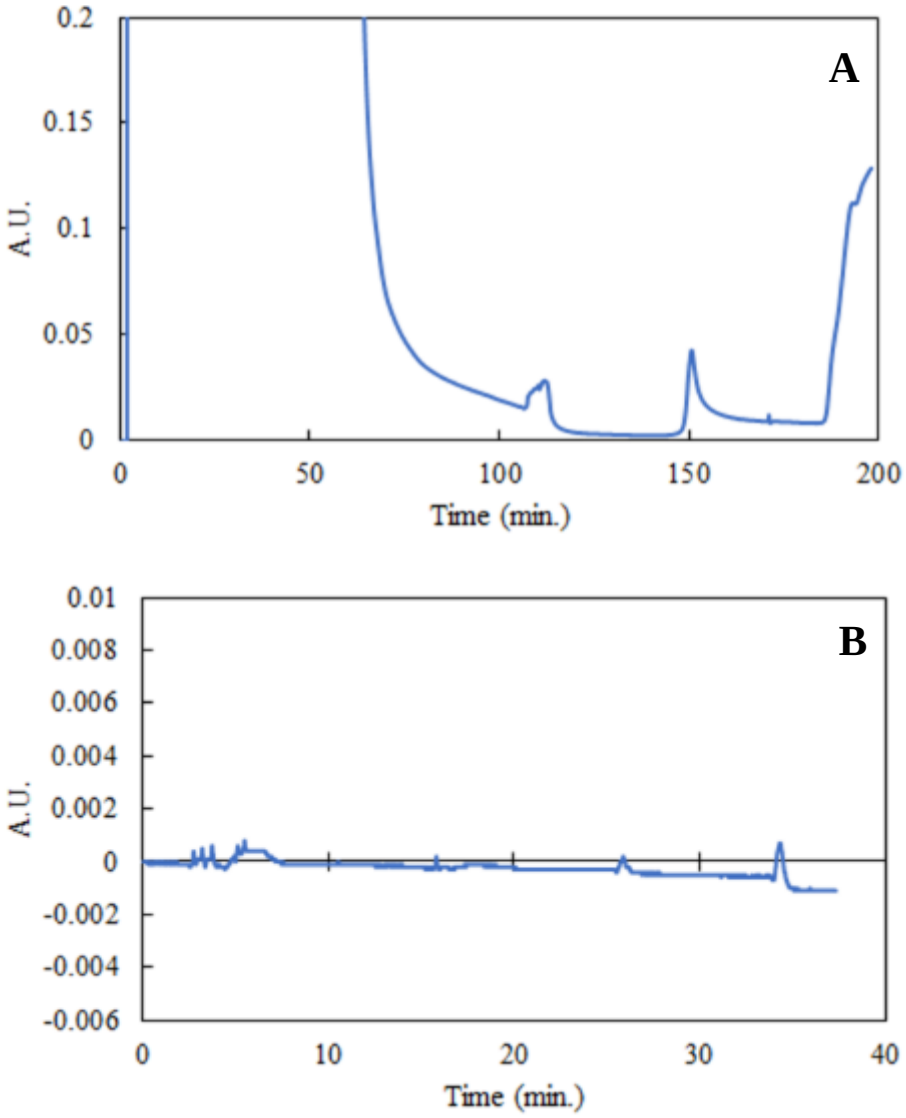

Figure 4.1. Chromatograms showing the purification of Sep p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~110 and 15 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~190 minutes. (**B)** Anion exchange chromatography showing no discernable elution peak of thrombin-digested Sep p53(1-93).

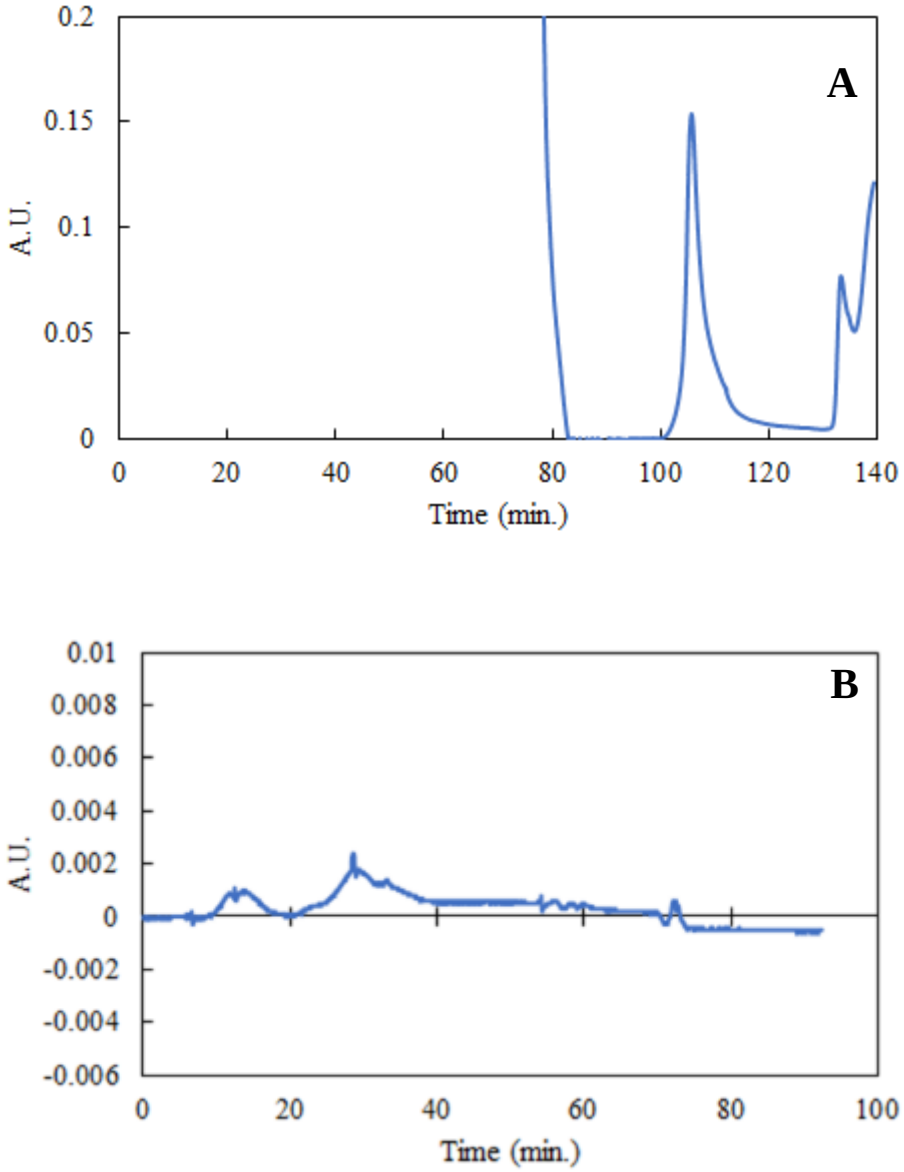

Figure 4.2. Chromatograms showing the purification of Sep p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~105 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~135 minutes. (**B)** Anion exchange chromatography showing no discernable elution peak of thrombin-digested Sep p53(1-93).

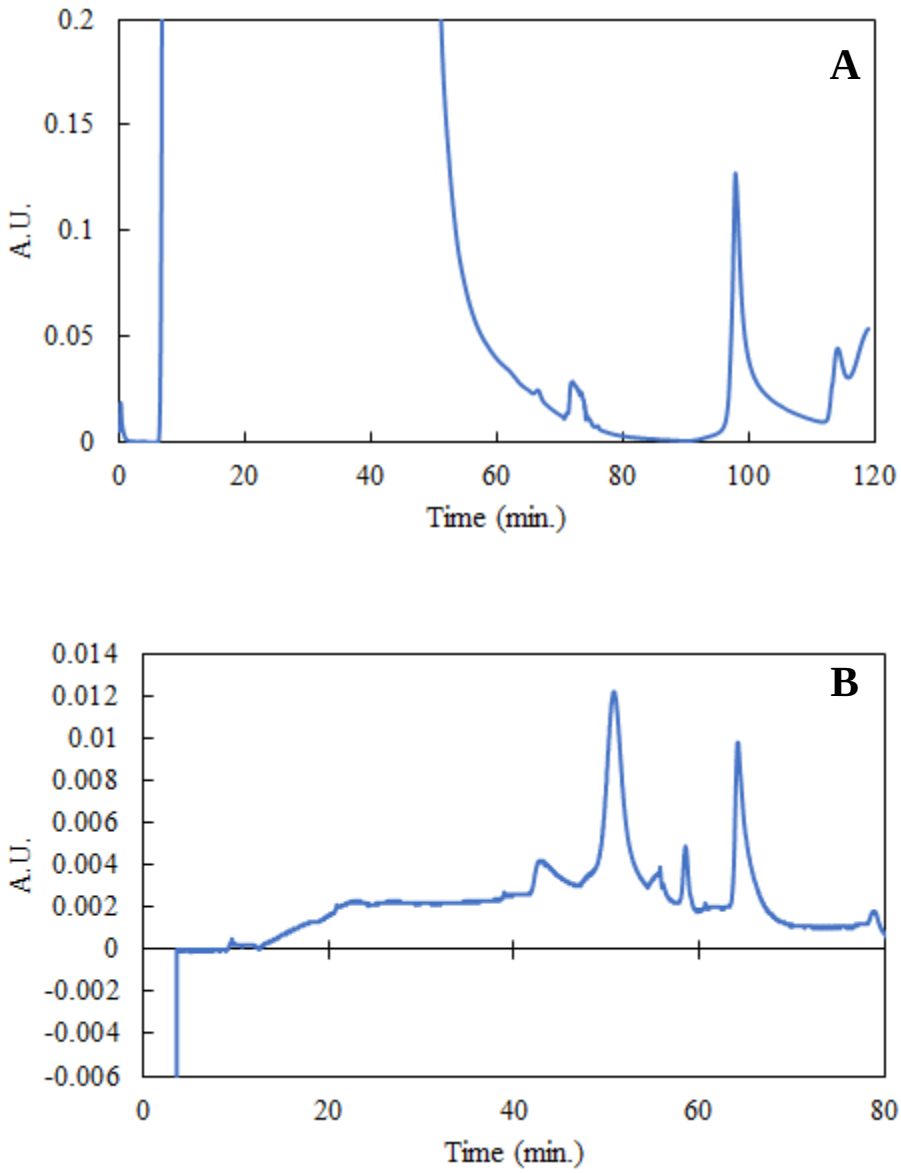

Figure 4.3. Chromatograms showing the purification of Sep p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~70 and 100 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~117 minutes. (**B)** Anion exchange chromatography showing two possible elution peaks of thrombin-digested Sep p53(1-93) at ~50 and 65 minutes.

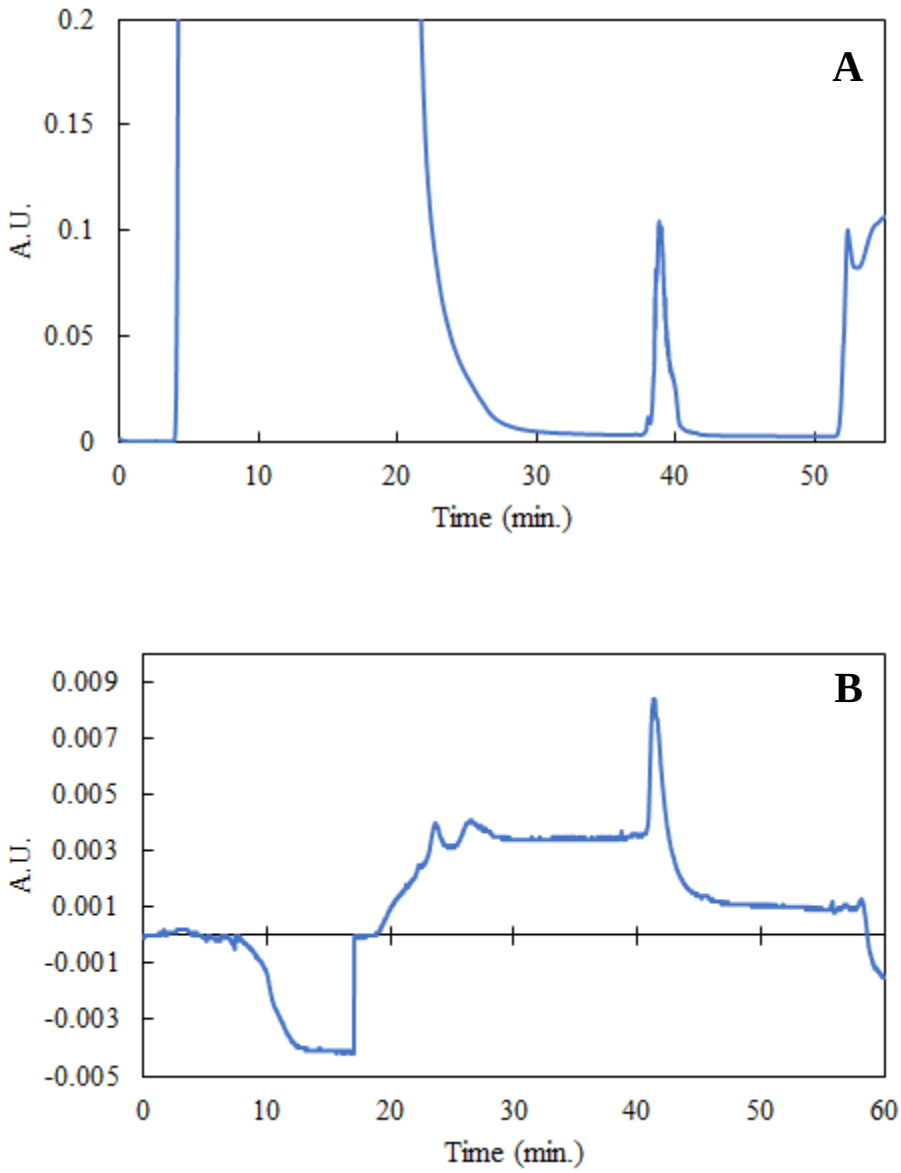

Figure 4.4. Chromatograms showing the purification of Sep p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~39, and p53(1-93) with a 6x His tag eluting from the column at ~52 minutes. (**B)** Anion exchange chromatography showing a possible elution peak of thrombin-digested Sep p53(1-93) at ~40 minutes.

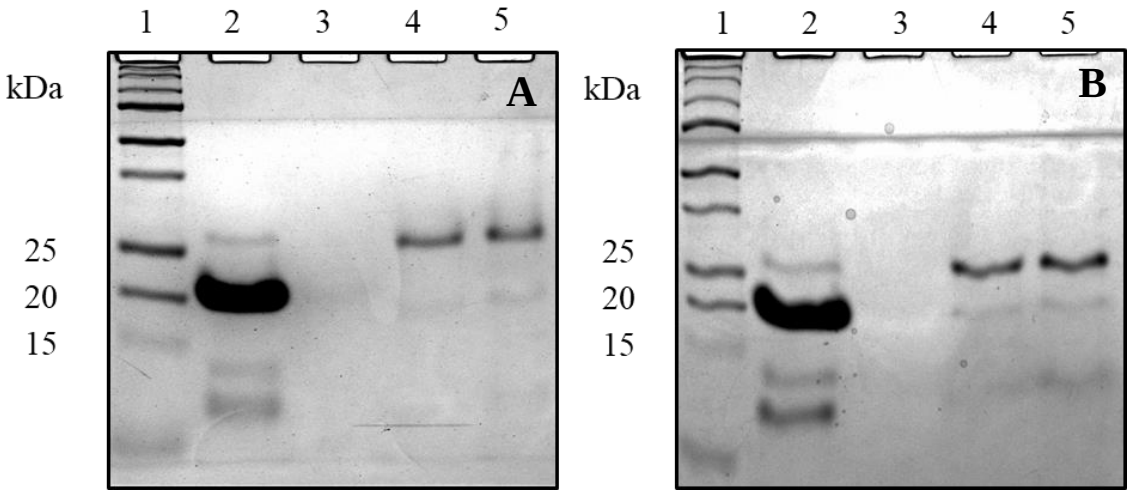

Figure 4.5. SDS-PAGE and Phos-Tag SDS-PAGE analysis of recombinant Sep p53(1-93) purifications. Silver-stained SDS-PAGE and Phos-Tag gels from the results of WT purification. **A.** Standard SDS-PAGE; Lane 1: Protein Standards; Lane 2: Ni-NTA WT eluate sample of a standard expression; Lane 3: Anion exchange Sep eluate samples of 30 ºC expression; Lane 4: Anion exchange Sep eluate samples of 20 ºC expression; Lane 5: Anion exchange Sep eluate samples of ethanol expression at 20 ºC. **B.** Phos-Tag SDS-PAGE; Lane 1: Protein Standards; Lane 2: Ni-NTA WT eluate sample of a standard expression; Lane 3: Anion exchange Sep eluate samples of 30 ºC expression; Lane 4: Anion exchange Sep eluate samples of 20 ºC expression; Lane 5: Anion exchange Sep eluate samples of ethanol expression at 20 ºC.

4.3 Glucose Effects on Cell Growth

For initiation of transcription to begin efficiently, cyclic AMP (CAP) and cyclic AMP receptor protein (CAP) form a CAP/cAMP complex which binds directly upstream of multiple promotors including *lac* and *rhaBAD* promotors. This binding event increases rates of transcription by making it easier for RNA polymerase to bind promotor regions[101]. When glucose is added to growth media, cellular cAMP levels are low as the transport of glucose into the cell negatively effects adenylate cyclase, which is responsible for the production of cAMP. The addition of glucose to growth media before expression induction is often used to keep leaky expression of plasmids down as this leaky expression lowers growth rates, and in some cases, where the protein of interest is toxic to *E. coli*, can stop cell replication before induction is even started[70,101,102].

An addition of 0.08% w/v glucose is already called for in the SepOTS protocol as put forward by Pirman and colleagues[62], however there were still many instances while trying to express Sep-p53(1-93) where the growth rate stalled at a low OD_{600}, and the desired OD_{600} of 0.8-0.9 of the protocol was never reached. Expression systems using IPTG often call for between 0.5% and 1.0% glucose w/v to expression buffers, while systems using rhamnose call for up 0.5% w/v glucose to prevent leaky expression. The next experiment used 5-100 mL samples in 500 mL Erlenmeyer flasks to test if extra glucose in the culture media would provide cells a better environment to reach the desired OD_{600} of between 0.8 and 0.9. Glucose concentrations chosen were 0%, 0.1%, 0.5%, 1.0%, and 2.0% w/v. One-way ANOVA and Tuckey test were used to analyze the data to see if adding more glucose to the culture media would allow cells to reach the desired OD_{600} of between 0.8 and 0.9 that is suggested for the SepOTS method. An average of 3

trials was done. The control at 0% and the 0.1% w/v additions were not significantly better than one other at allowing cell growth. Both the 0% and the 0.1% w/v additions were shown to be less effective at facilitating cell growth than the 0.5%, 1.0% and 2.0 % w/v additions as summarized by Table 4.2 and Figures 4.6 and 4.7. The 0.5%, 1.0% and 2.0 % w/v additions were not significantly different from one another, so the smallest effective concentration of 0.5% glucose w/v was added to the culture media moving forward.

Table 4.2. Effects of Glucose Concentration of Cell Growth.

Glucose Concentration	0	1.5	3	5	6	7
0.0% w/v	0.027	0.071	0.143	0.319	0.544	0.799
0.1% w/v	0.027	0.079	0.166	0.390	0.662	0.850
0.5% w/v	0.027	0.084	0.181	0.510	0.786	1.058
1% w/v	0.027	0.078	0.181	0.490	0.784	1.116
2% w/v	0.027	0.076	0.186	0.520	0.805	1.139

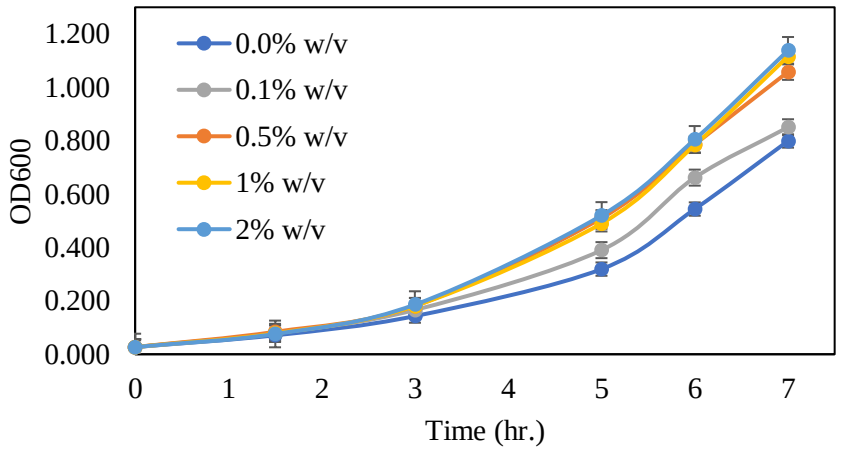

Figure 4.6. Effects of Glucose Concentration on Cell Growth Over Time. 0.0% and 0.1% w/v glucose additions do not do as well in growing cells than the 0.5%, 1.0%, or 2.0% w/v glucose additions.

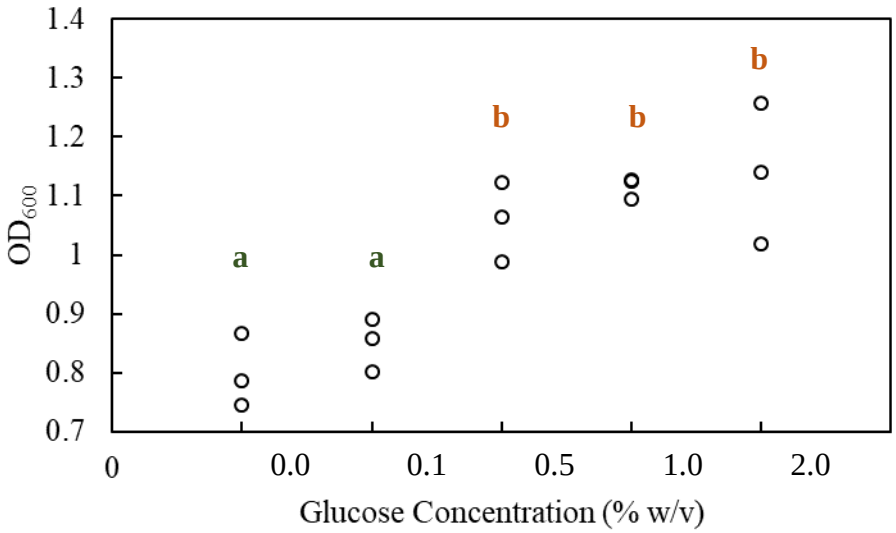

Figure 4.7. Analysis of Glucose Concentrations Impact on Final Cell OD. Group (a) and Group (b) are determined to be statistically different than one another via ANOVA and Tuckey test. Concentrations of 0% and 0.1% w/v glucose belong to group (a) while 0.5%, 1.0%, and 2.0% w/v glucose additions belong to group (b).

4.4 Expression of Sep p53(1-93) in C321.ΔA with Glucose Addition

Expression of recombinant Sep p53(1-93) was then attempted with an added 0.5% w/v glucose to the 2xYT culture media used during expression. Expression volumes were brought down to 1 L. C321.ΔA cells were used again at induction temperatures of 37 ºC, 30 ºC, and 20 ºC with 1 mM IPTG and 0.2% w/v rhamnose as suggested by the plasmid manufacturers. To isolate any expressed protein, nickel affinity chromatography and anion exchange chromatography were performed. There were no discernable elution peaks during the first two expressions at temperature 37 ºC and 30 ºC. There seemed to be a slight elution peak before the formation of an imidazole shoulder in the nickel chromatogram for the expression carried out at 30 ºC, however no elution peak that would indicate protein overexpression was seen during anion exchange chromatography (Figure 4.8).

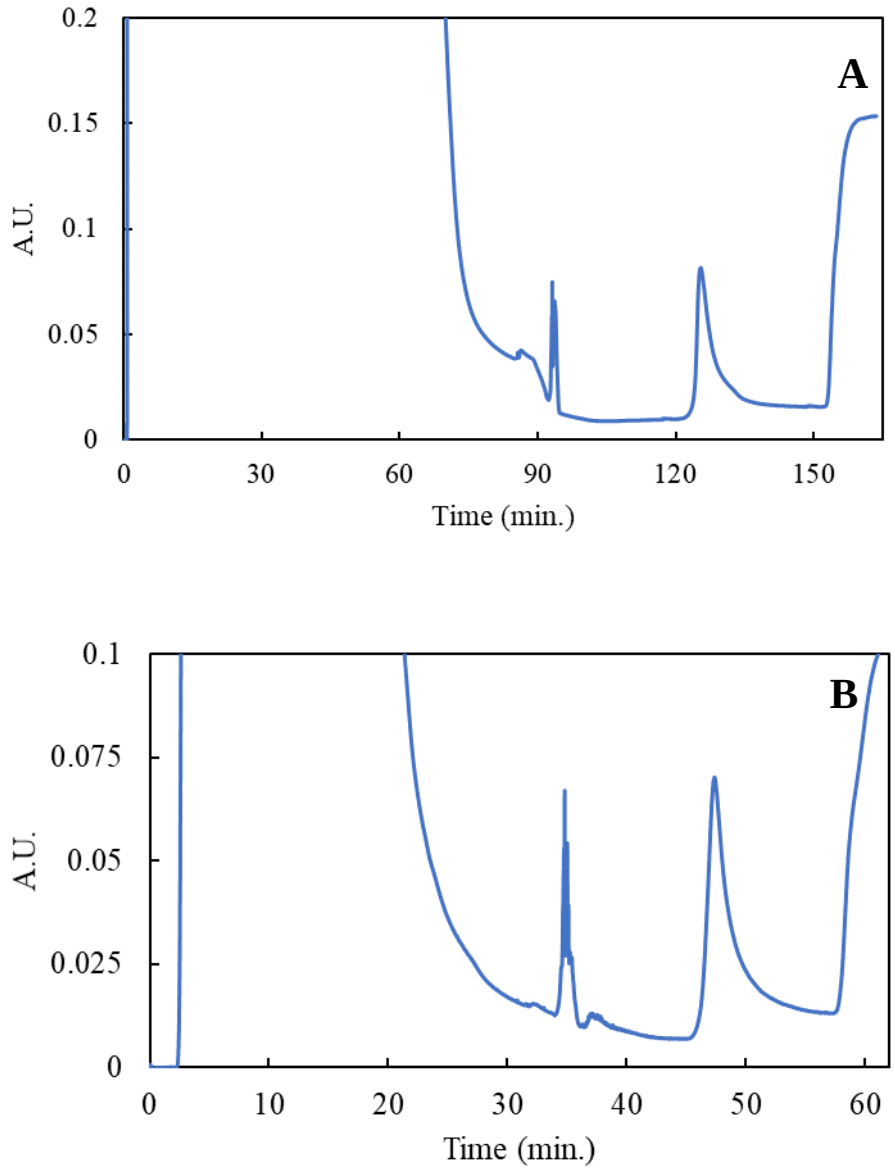

Figure 4.8. Nickle Affinity Chromatograms of attempted protein expressions. The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography for 37 ºC expression. Contaminating proteins eluting at ~90 and 125 mins, but no elution peak for p53(1-93) with a 6x His tag. (**B)** Nickel affinity chromatography for 30 ºC expression. Contaminating proteins eluting at ~35 and 49 mins, but no elution peak for p53(1-93) with a 6x His tag.

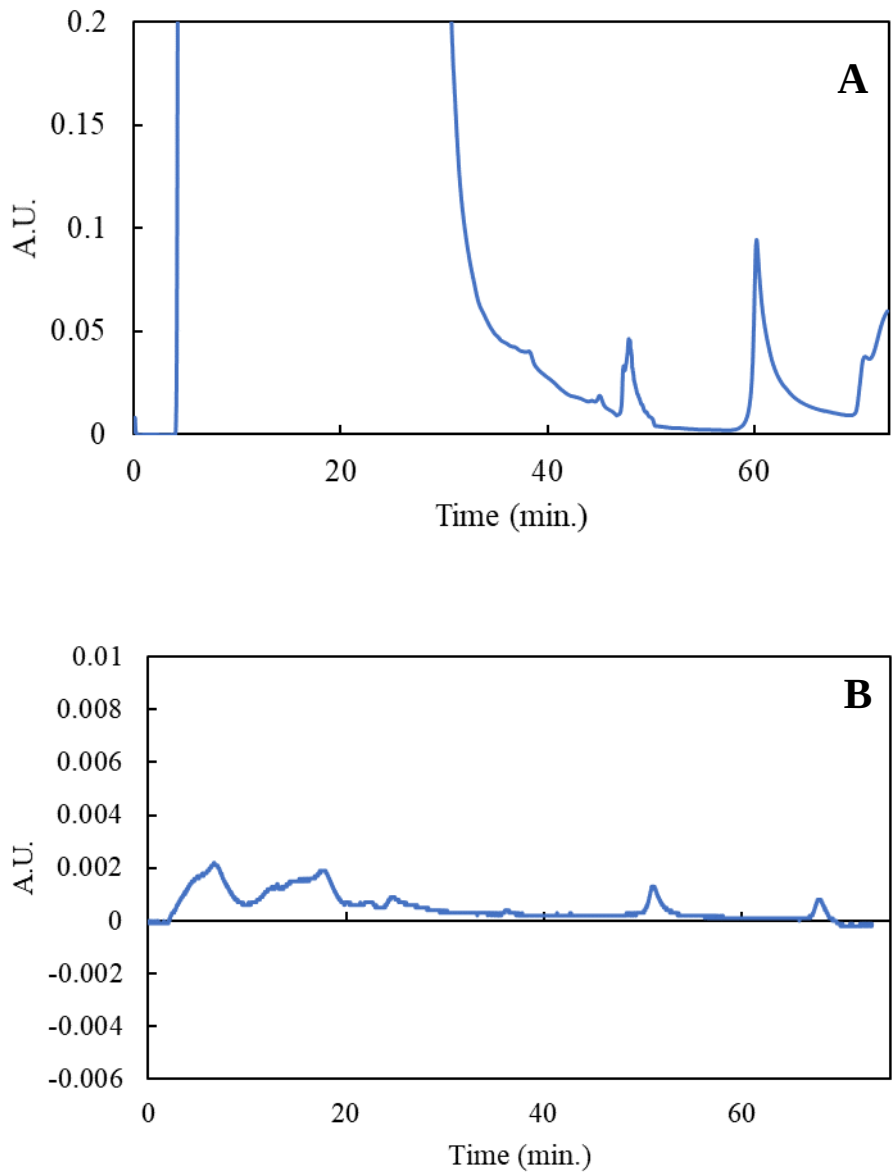

Figure 4.9. Chromatograms showing the purification of Sep p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~45 and 60 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~70 minutes. (**B)** Anion exchange chromatography showing no discernable elution peak for Sep p53(1-93).

4.5 Expression of WT p53(1-93) with SupD plasmid in C321.ΔA

An expression using the SupD plasmid was done next as a positive control to see if the two-plasmid system could express recombinant WT p53(1-93) without phosphorylation. As with the previous attempts at expression, 0.5% w/v glucose was added to the 2xYT culture media used, and the expression volume was kept at 1 L. C321.ΔA cells were used at an induction temperature of 20 ºC with 1 mM IPTG and 0.2% w/v rhamnose as suggested by the plasmid manufacturers. To isolate any expressed protein, nickel affinity chromatography and anion exchange chromatography were performed. The nickel chromatogram for recombinant WT p53(1-93) showed a small peak before the imidazole shoulder that indicates protein was expressed (Figure 4.10). There was one distinguishable peak in the anion exchange chromatogram, that after concentrating and dialyzing showed that only 0.732 mg of protein was produced.

SDS-PAGE followed by silver staining was carried out on the sample that eluted from the nickel and anion exchange columns from the SupD expression attempt. All samples migrated to about the 20 - 25 kDa mark of the protein ladder in lane, which is the expected migration of WT p53(1-93) from previously obtained sample. Digestion of the affinity tag from WT p53(1-93) has been shown to decrease its migration[19], which was observed from the protein obtained by SupD expression (Figure 4.11).

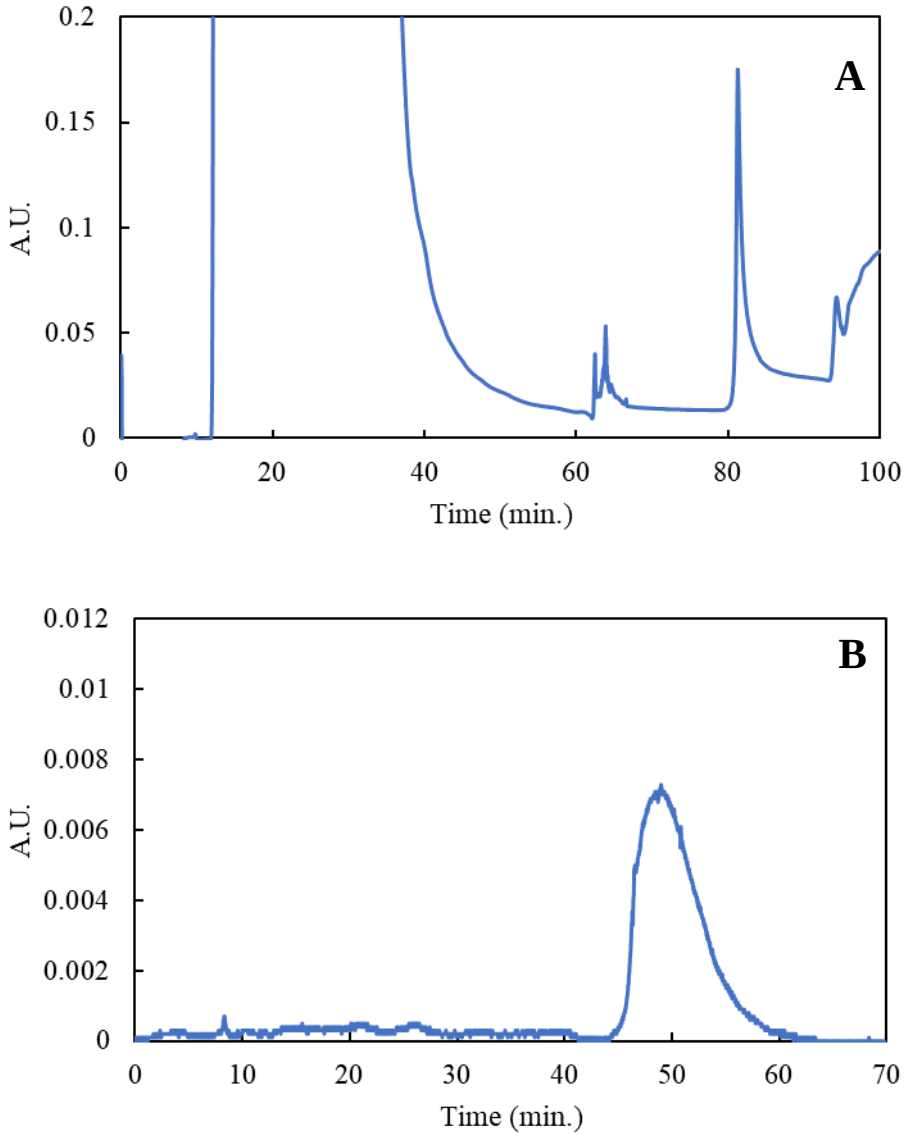

Figure 4.10. Chromatograms showing the purification of SupD WT p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~65 and 85 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~95 minutes. (**B)** Anion exchange chromatography showing a possible elution peak of thrombin-digested SupD WT p53(1-93) at ~50 minutes.

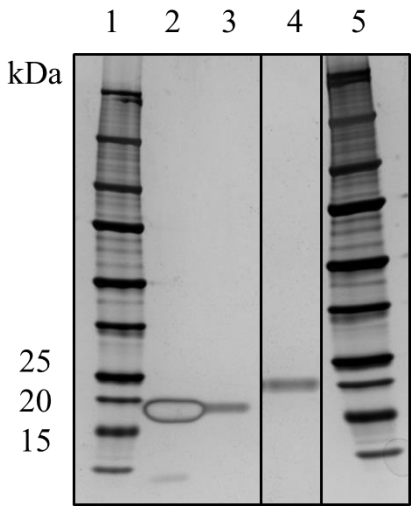

Figure 4.11. SDS-PAGE analysis of recombinant WT p53(1-93) SupD. Silver-stained SDS-PAGE gel from the results of SupD WT purification. Lane 1: Protein Standards; Lane 2: Ni-NTA WT eluate sample; Lane 3: Thrombin digested WT eluate sample; Lane 4: Concentrated and dialyzed anion exchange eluate. Lane 5: Protein Standards.

4.6 Expression of Sep-p53(1-93) p15a plasmid in C321.ΔA

As put forth earlier in this chapter, it is thought that if two plasmids belong to the same compatibility group - using the same replicon machinery – the two plasmids will compete for the same cellular machinery. This competition causes less copies of each plasmid being replicated per cell, and eventually leads to the plasmids being separated into different cells completely. The ORI of the SepOTSλ plasmid and the first Sep p53(1-93) plasmid belong to the same compatibility group which may have been one reason for low yield. The new Sep p53(1-93) p15a that was purchased that contained a p15a ORI that is unrelated to the pUC and pBR322 plasmid group.

In order to cut down on time and guanidine-HCl during lysis of cell pellets, expressions were done at 10-mL volumes. 1-mL samples were taken at multiple times during the induction step of attempted expressions, and the samples run on 14% bis/acrylamide hand-cast gels. Alongside each expression attempt, a negative control was done in which no IPTG or rhamnose was added to the culture. Multiple expression attempts were performed at 37ºC, 30ºC, and 20ºC with rhamnose concentrations at 0.2% w/v, and IPTG concentrations at 1 mM. Figures 4.12 – 4.14 show that there were no changes in any bands widths when comparing the control and induced expressions, which indicates that there was no overexpression of protein at any temperature.

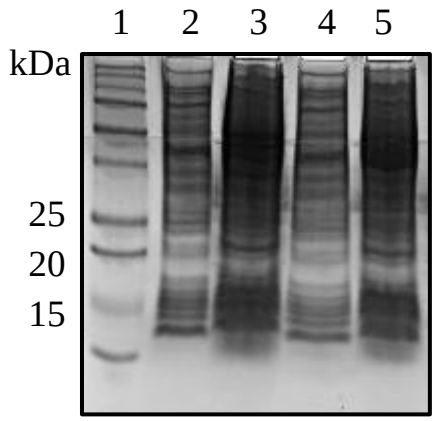

Figure 4.12. SDS-PAGE analysis of Sep p53(1-93) expression in C321.ΔA Cells at 37 ºC. Silver-stained SDS-PAGE gel from the results of attempted Sep p53(1-93) expression in C321.ΔA cells at 37 ºC. Lanes 2 and 3 are samples from a non-induced negative control expression. Lane 1: Protein Standards; Lane 2: 0 hr sample of Control expression; Lane 3: 5 hr sample of Control expression; Lane 4: 0 hr sample of Induced expression; Lane 5: 5 hr sample of Induced expression.

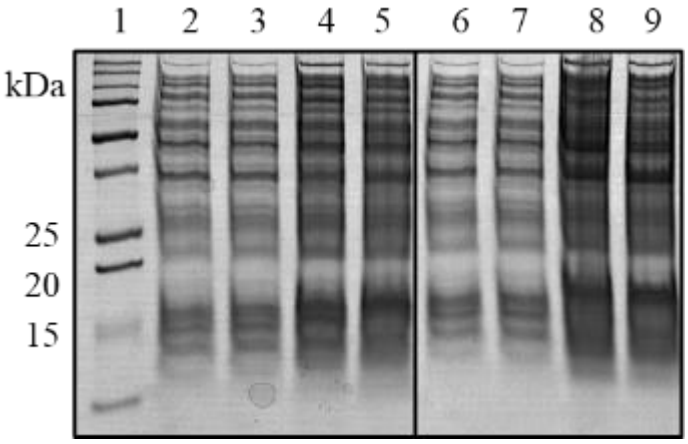

Figure 4.13. SDS-PAGE analysis of Sep p53(1-93) expression in C321.ΔA Cells at 30 ºC. Silver-stained SDS-PAGE gel from the results of attempted Sep p53(1-93) expression in C321.ΔA cells at 30 ºC. Lanes 2 through 5 are samples from a non-induced negative control expression. Lane 1: Protein Standards; Lane 2: 0 hr sample of Control expression; Lane 3: 1 hr sample of Control expression; Lane 4: 5 hr sample of Control expression; Lane 5: 18 hr sample of Control expression.; Lane 6: 0 hr sample of Induced expression; Lane 7: 1 hr sample of Induced expression; Lane 8: 5 hr sample of Induced expression; Lane 9: 18 hr sample of Induced expression.

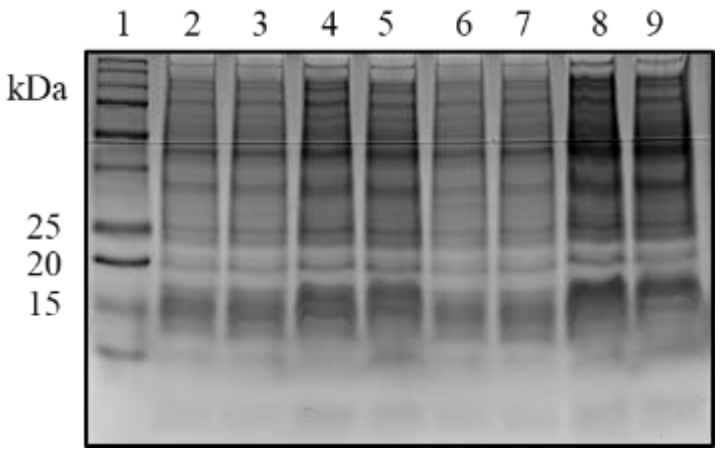

Figure 4.14. SDS-PAGE analysis of Sep p53(1-93) expression in C321.ΔA Cells at 20 ºC. Silver-stained SDS-PAGE gel from the results of attempted Sep p53(1-93) expression in C321.ΔA cells at 20 ºC. Lanes 2 through 5 are samples from a non-induced negative control expression. Lane 1: Protein Standards; Lane 2: 0 hr sample of Control expression; Lane 3: 1 hr sample of Control expression; Lane 4: 5 hr sample of Control expression; Lane 5: 18 hr sample of Control expression.; Lane 6: 0 hr sample of Induced expression; Lane 7: 1 hr sample of Induced expression; Lane 8: 5 hr sample of Induced expression; Lane 9: 18 hr sample of Induced expression.

4.7 Expression of Ala$^-$ p53(1-93) in BL21ΔSerB Cell Line

C321.ΔA cells were originally chosen for the expression of Sep-p53(1-93) due to the lack of any TAG stop codons present on the cellular DNA as well as for the deletion of release factor 1 (RF-1) which mediates the termination of translation at TAG stop codons. However, the deletion of the RF-1 gene has been shown to decrease overall cellular growth and thus, the amount of protein being produced[103]. For this reason, expression using the BL21ΔserB cell line was carried out to see if poor cell growth was a limiting factor for the expression of recombinant protein.

For the first control experiment, an expression of the mutant Ala$^-$ p53(1-93) was carried out. The Ala$^-$ mutant was chosen because it expresses at higher levels then WT p53(1-93) for unknown reasons. Conditions were the same as the Ala$^-$ expression in chapter 3, with the expression temperature set to 37 ºC, and induction with 1 mM IPTG. Both, nickel and anion exchange chromatography showed elution peaks that would indicate of expression of Ala$^-$ p53(1-93) as shown in Figure 4.15. SDS-PAGE lanes 5-7 in Figure 4.16 show that the elution samples during nickel chromatography were not very concentrated. The band lane 8 indicates there was protein purified from the anion exchange column, but UV-vis showed that it was at a concentration of only 0.383 mg total protein.

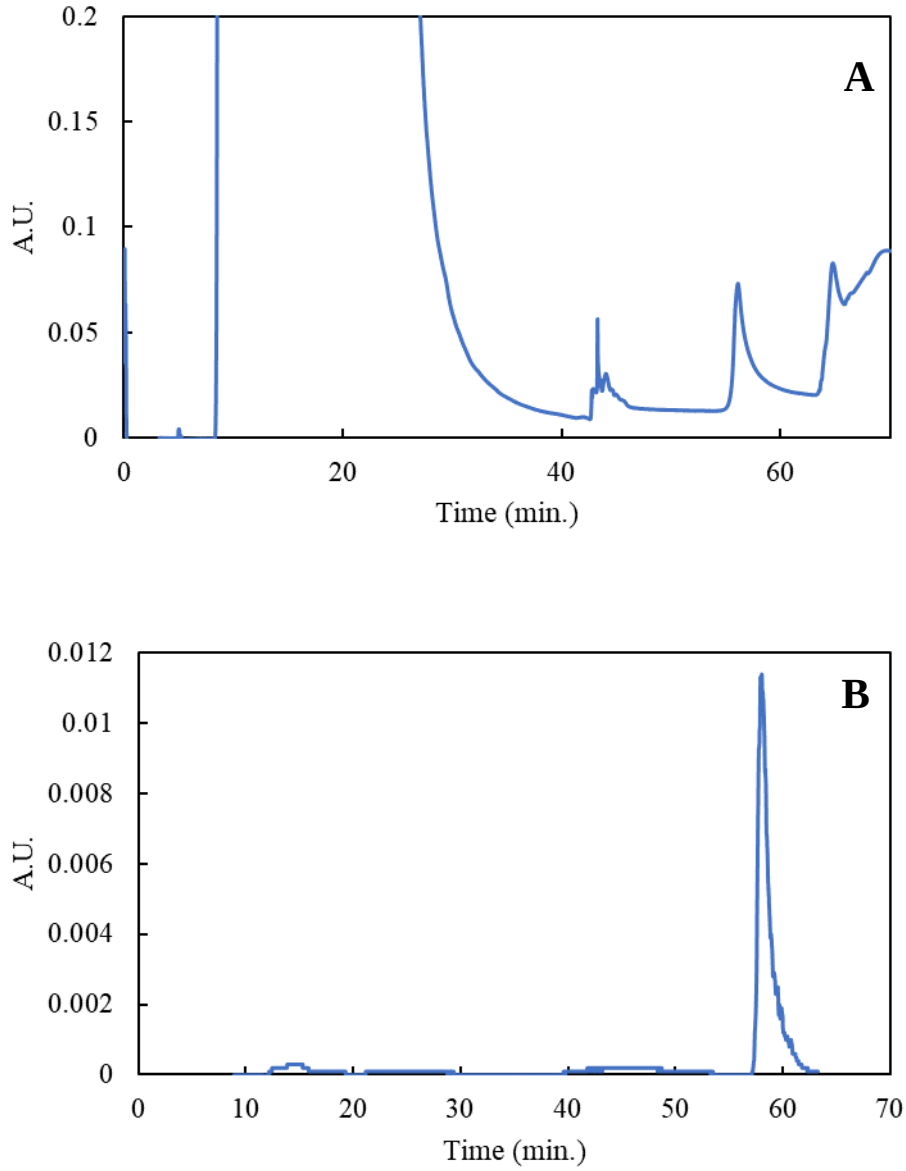

Figure 4.15. Chromatograms showing the purification of Ala⁻ p53(1-93) from C321.ΔA cells. The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~43 and 55 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~65 minutes. (**B)** Anion exchange chromatography showing a possible elution peak of thrombin-digested Ala⁻ p53(1-93) at ~59 minutes.

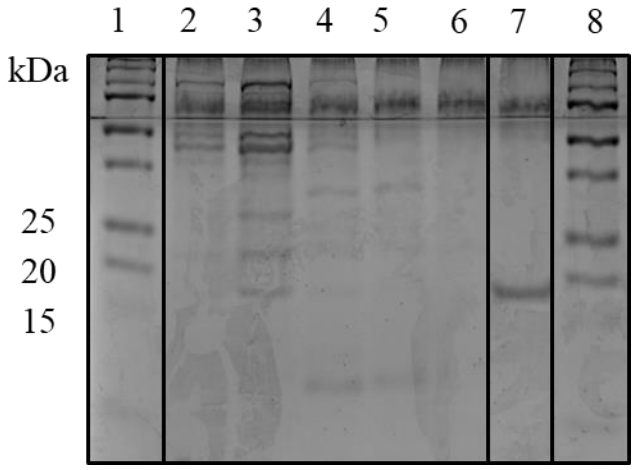

Figure 4.16. SDS-PAGE analysis of recombinant Ala$^-$ p53(1-93) in BL21ΔSerB.

Silver-stained SDS-PAGE gel from the results of Ala$^-$ purification from BL21ΔSerB. Lane 1: Protein Standards; Lane 2 and 3: Ni-NTA Ala$^-$ Wash Buffer 3 samples; Lane 4 - 6: Ni-NTA Ala$^-$ eluate samples; Lane 7: Concentrated and dialyzed anion exchange Ala$^-$ eluate; Lane 9: Protein Standards.

4.8 Expression of Sep-p53(1-93)-p15a in BL21ΔserB Cells

Experiments were done at 10-mL volumes, with expression temperatures 37 ºC, 30 ºC, and 20 ºC. Figure 4.16 shows that after expression at each of these temperatures with 1 mM IPTG and 2.0% rhamnose, only the expression at 30 ºC seemed to indicate any overexpression of protein. This is seen by the increased presence of a protein band in lanes 4 – 6 between the 20 kDa and 25 kDa ladder marks that is disproportionate to an increase in expression that would be simply due to the passage of time. This is where the Sep p53(1-93) is thought to migrate during electrophoresis. The figure also shows the expression in C321.ΔA at 30 ºC in lanes 9 – 11 that shows no significant changes over time as compared to expression in the BL21ΔSerB cell line.

When attempting to scale up the expression of Sep p53(1-93) in BL21ΔSerB cells to 100 mL for the purpose of purifying protein there seemed to be no increase in band width that might be associated with expression of protein when run on a gel and when compared to a negative control expression where no IPTG or rhamnose was added (Figure 4.17). This indicates that the expression is possible, but the system is only viable for expression of Sep p53(1-93) at amounts that would be too low for our structural analysis techniques.

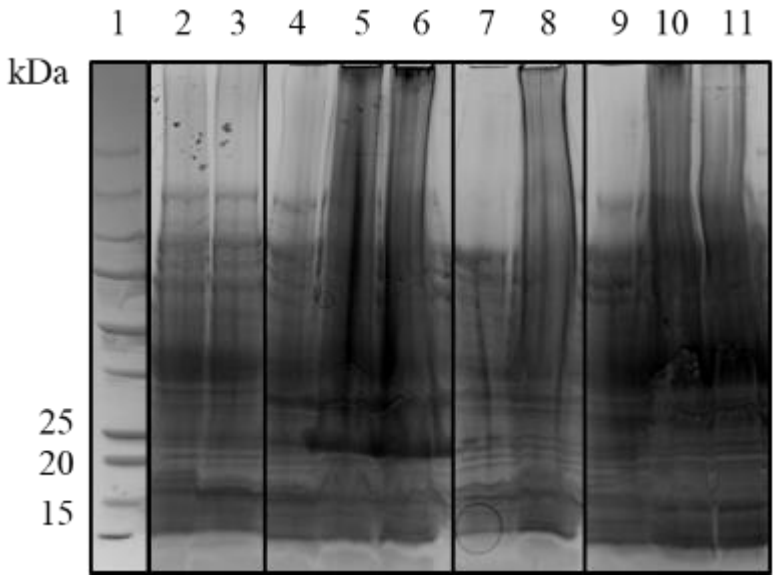

Figure 4.17. SDS-PAGE analysis of recombinant Sep p53(1-93) expressions at 37 ºC, 30 ºC, and 20 ºC. Silver-stained SDS-PAGE gel from the results of a Sep p53(1-93) expressions at 20 ºC ,30 ºC, and 37 ºC from BL21ΔSerB compared to 30 ºC C321.ΔA expression. Lane 1: Protein Standards; Lane 2: 0 hr sample of 20 ºC expression from BL21ΔSerB cells; Lane 3: 18 hr sample of 20 ºC expression from BL21ΔSerB cells; Lane 4: 0 hr sample of 30 ºC expression from BL21ΔSerB cells; Lane 5: 5 hr sample of 30 ºC expression from BL21ΔSerB cells; Lane 6: 18 hr sample of 30 ºC expression from BL21ΔSerB cells; Lane 7: 0 hr sample of 37 ºC expression from BL21ΔSerB cells; Lane 8: 6 hr sample of 37 ºC expression from BL21ΔSerB cells; Lane 9: 0 hr sample of 30 ºC expression from C321.ΔA cells; Lane 10: 5 hr sample of 30 ºC expression from C321.ΔA cells; Lane 11: 18 hr sample of 30 ºC expression from C321.ΔA cells.

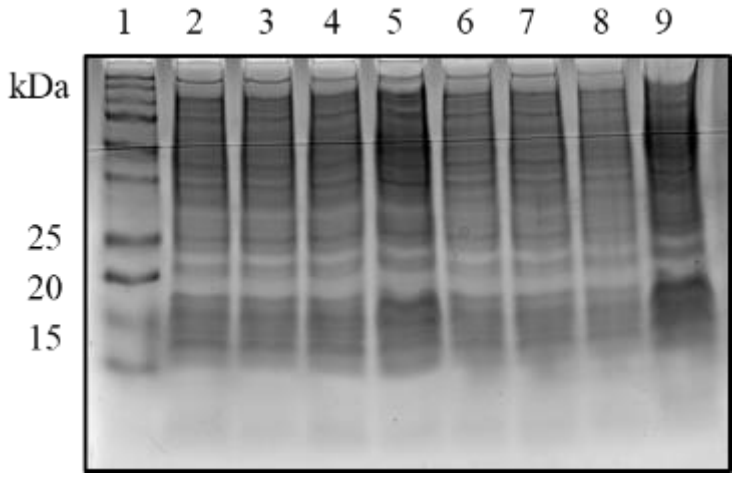

Figure 4.18. SDS-PAGE analysis of Sep p53(1-93) expression in BL21ΔSerB Cells. Silver-stained SDS-PAGE gel from the results of a scaled-up Sep p53(1-93) expression in BL21ΔSerB cells. Lanes 2 through 5 are samples from a non-induced negative control expression. Lane 1: Protein Standards; Lane 2: 0 hr sample of Control expression; Lane 3: 1 hr sample of Control expression; Lane 4: 5 hr sample of Control expression; Lane 5: 18 hr sample of Control expression; Lane 6: 0 hr sample of Induced expression; Lane 6: 3 hr sample of Induced expression; Lane 6: 5 hr sample of Induced expression; Lane 6: 18 hr sample of Induced expression.

4.9 SupD Expression of WT p53(1-93)-p15a Plasmid in BL21ΔserB Cells

An expression using the SupD plasmid was done next as a positive control to see if the SepOTSλ plasmid was the issue in the system. As with the previous attempts at expression, 0.5% w/v glucose was added to the 2xYT culture media used. The expression volume was brought back up to 1 L. An induction temperature of 30 ºC was used for expression with 1 mM IPTG and 0.2% w/v rhamnose as suggested by the plasmid manufacturers. To isolate any expressed protein, nickel affinity chromatography was used to purify 6xHis tagged protein. The new p15a plasmid uses a tobacco etch virus (TEV) cleavage site instead of a thrombin cleavage site. TEV has shown to have more specificity to its cleavage site than thrombin. TEVs cleavage site is Glu-Asn-Leu-Tyr-Phe-Gln-Gly and it cleaves between Gln and Gly. A second nickel column was then used to separate the cleaved protein from the 6xHis tag in the flow through as shown in Figure 4.19. The nickel chromatogram for recombinant WT p53(1-93) showed a small peak before the imidazole shoulder that indicates protein was expressed (Figure 4.19). There was a distinguishable peak in the second nickel column, that after concentrating and dialyzing showed that ~ 0.841 mg of protein was produced (Figure 4.10).

SDS-PAGE followed by silver staining was carried out to on both nickel elution samples from the SupD expression attempt, along-side samples of a WT p53(1-93) from our normal BL21 cell line using a normal WT p53(1-93) plasmid. Lanes 2 and 3 in Figure 4.20 belong to the normal expression system and lanes 4 and 5 contain SupD p53(1-93) sample. All samples travel to about the 25 kDa mark of the protein ladder in lane, which is area WT p53(1-93) is expected to travel. It is shown that the normal expression allows for greater recovery of expressed protein. There is also more protein

shown at the bottom of gel in lane 4 than in lane 2, possibly indicating that either the protein recovered from the BL21ΔSerB expression was broken down easier than normally expressed WT p53(1-93) or that read through of the stop codon is not efficient.

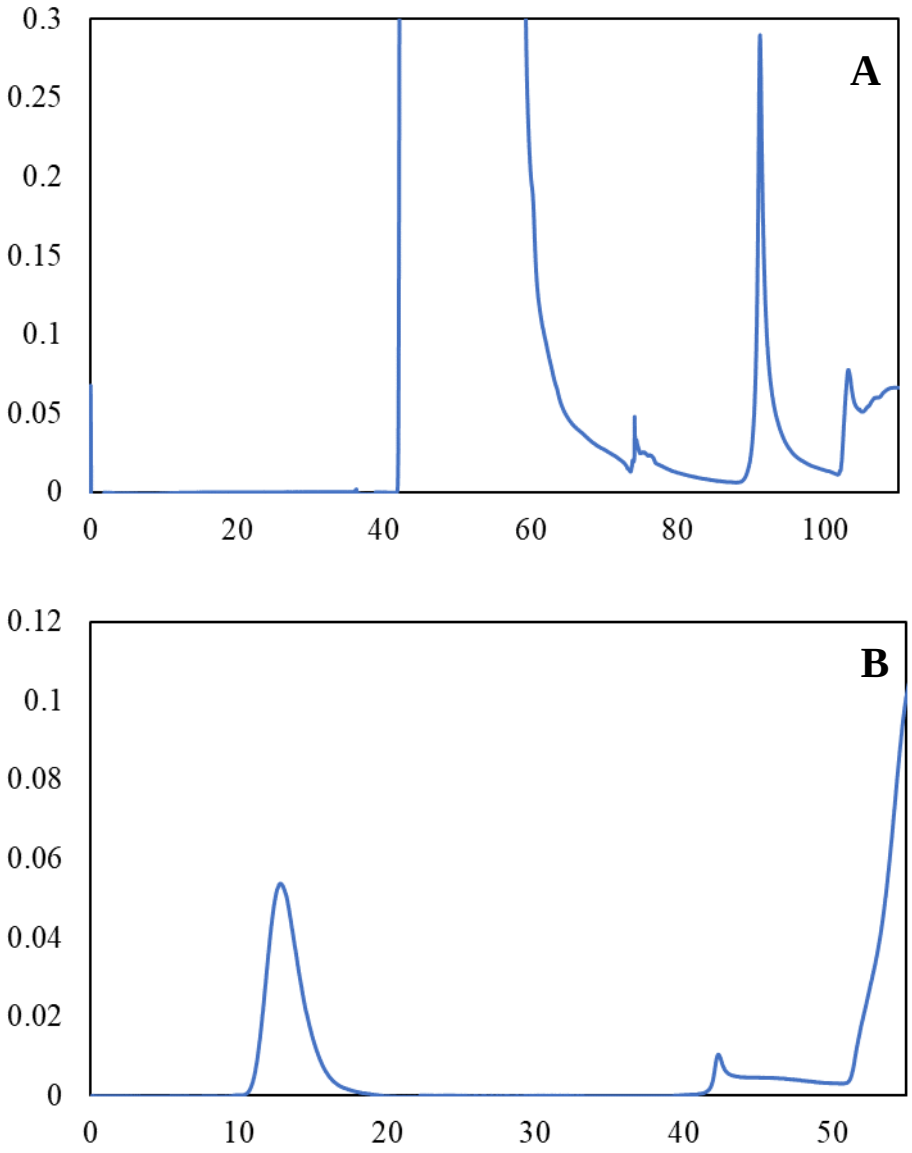

Figure 4.19. Chromatograms showing the purification of SupD WT p53(1-93). The blue line represents absorbance at 280 nm. **(A)** Nickel affinity chromatography with contaminating proteins eluting at ~75 and 95 minutes, and p53(1-93) with a 6x His tag eluting from the column at ~105 minutes. (**B)** Nickel affinity chromatography post TEV digestion showing an elution peak of TEV-digested SupD WT p53(1-93) at ~12 minutes. Possible un-cleaved p53(1-93) eluting at ~ 41 minutes before imidazole spike starts at a ~ 50 minutes.

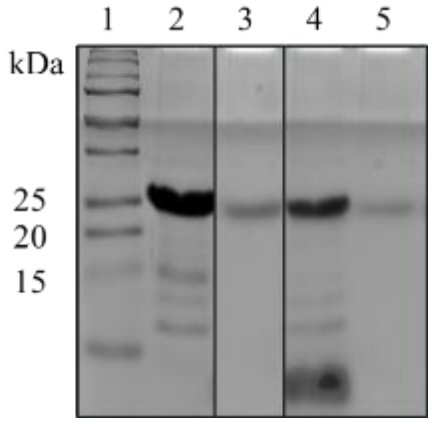

Figure 4.20. SDS-PAGE analysis of recombinant WT p53(1-93) SupD purification and comparison to Standard WT p53(1-93) purification. Silver-stained SDS-PAGE gel from the results of standard and SupD WT purification. Lane 1: Protein Standards; Lane 2: Ni-NTA WT eluate samples of standard protein expression; Lane 3: TEV digested WT eluate sample of standard protein expression; Lane 4: Ni-NTA WT eluate samples of SupD protein expression; Lane 5: TEV digested WT eluate sample of SupD protein expression.

4.10 Discussion

Phosphorylation is one of the most well studied of post-translational modifications and is thought to be the most common mechanism for regulating protein function[66]. One-third of the proteome is said to be regulated by phosphorylation[66]. Most of these phosphorylation sites are located in intrinsically disordered regions of the protein[18,67,104]. IDPs are known to consist of non-interacting charged groups that result in high overall charged state in neutral pH as well as expanded states due to charge-charge repulsion interactions[31]. IDPs primary sequence are usually low complexity and mainly consist of disorder-promoting residues (D, K, R, S, Q, P, and E). It has previously been shown that changes in pH, temperature, and primary sequence substitutions affect the compaction or extension of IDPs. Phosphorylation induces structural and dynamic changes that facilitate the catalytic activity of proteins or allow association or dissociation with neighboring proteins, as in the case of the IDP prostate associated gene 4 (PAGE4)[104]. Kulkarni and associates showed that when PAGE4 is in its native un-phosphorylated state, it has a compact structure that binds to the activator protein 1 (AP1) transcription factor. AP1 is a transcription factor that is associated with many processes including cell growth, apoptosis, and cell differentiation. When PAGE4 is hyperphosphorylated during oxidative stress events, PAGE4 exhibits an expanded, random coil structure. This structure has a lowered affinity to AP1, allowing potentiation of AP1 activity. During phosphorylation of an IDP the additional -2 charge groups is added which is thought to lead to greater expansion or compaction of the phosphorylated protein due to increased charge-charge repulsion interactions or interactions with basic residues, respectively. Though phosphorylation has been shown to facilitate functionally

significant structural changes in IDPs, how phosphorylation effects the conformational biases of IDPs is still not well understood., therefore developing a system from which a model phosphorylated IDP can be purified is necessary for further study.

4.10.1 Phosphorylation Orthogonal Translation System

In the field of protein research, the production of many recombinant proteins over time has led to the development of many systems and technologies that have decreased the difficultly of protein expression and purification. However, during the expression and purification of new recombinant proteins many difficult and strategic choices have to be made. Every protein is different and there is no answer to the many questions that go into coming up with a purification strategy. However, there are evidence-based guidelines for us to follow when working with a new protein. In the case of attempting to purify a model phosphorylated IDP, the method followed was that of the SepOTS.

Throughout chapter 4 the SepOTS was not shown to produce phosphorylated recombinant proteins as has previously been done in the hands of other lab groups[56,62,91]. From our data presented here we show that the cell types associated with SepOTS were capable of producing some non-phosphorylated recombinant proteins. Namely the C321.ΔA cell line allowed for expression of non-phosphorylated WT p53(1-93) in section 4.5 using the two-plasmid system, while the BL21ΔSerB cell line was shown to express both Ala^- p53(1-93) and non-phosphorylated WT p53(1-93) is sections 4.7 and 4.9 respectively. When looking at the amounts of protein recovered from 1-L expressions and comparing them to other expressions done in our lab, the difference in the amount of protein produced is significantly less, with the SepOTS producing less than 1 mg of protein during any expression. There are a number of reasons the expression

phosphorylated p53(1-93) was not successful as well as other directions in which to take this project.

One explanation for low expression of phosphorylated recombinant protein could be the relationship between the inducing agent for the production of p53(1-93) and the ORI sites of the plasmids being used. Both of the p53(1-93) plasmids used for the SepOTS relied on the rhaBAD promotor for production. The induction of the rhaBAD promotor is governed by a cascade in which after the introduction of L-rhamnose, the expression of RhaR regulatory protein is induced. RhaR then induces rhaSR regulatory protein expression. As RhaS accumulates in the cell, rhaBAD expression is induced increasingly over time[105]. In a normal system this expression is considered slow, taking anywhere from 40 to 50 minutes to reach a steady expression rate. In the systems used in this project, time was of the essence as in one plasmid pairing, the plasmids were competing for replication machinery. This competition, along with the slow expression rates of the rhaBAD promoter could be a reason for low expression in the pUC-SepOTSλ/pBR322-p53(1-93) plasmid pairing. In the second plasmid pairing the p15a-p53(1-93) plasmid was a low copy number meaning the slow expression could have further limited the plasmid from expressing recombinant protein. I feel the best option may be to use the pUC/pBR322 plasmid combination with the expression of the p53(1-93) plasmid placed under the control of either IPTG, or tetracycline using the *lac* or *tet-op* promotors respectively. Tetracycline induced expression is a method in which tetracycline, when introduced to an expression culture after the OD_{600} has reach mid-log phase, binds directly to rtTA transcription factors which in turn, bind to the *tet-op* promoter upstream of a gene of interest to allow transcription to move forward. Pairing

high copy number plasmids with fast acting inducers might be able to produce more protein what has been seen. Another hypothesis for poor expression is that perhaps during SepOTS the cells preferred anaerobic conditions for expression of target protein. Since there was some evidence of expression of target protein at smaller volumes in test tubes with poor aeration as opposed to no expression of protein in larger more aerated sample, the possibility of anaerobic environments driving expression would also be explored by expressing in smaller volumes with little shaking. Finally, due to each cell type going through two rounds being made chemically competent, while under selective pressure with antibiotics, it is possible that the cells were too stressed to express protein using the SepOTS method. One way to reduce this stress would be to optimize a double transformation step for introducing two plasmids into one cell type.

Another route for the production of site specific phosphorylated p53(1-93) could be a cell-free protein synthesis (CFPS). During CFPS the same SepOTSλ plasmid that we have been using is induced by itself during regular cell growth to express orthogonal translation machinery. These cells are then collected, lysed, and processed to obtain OTS extract (30% v/v cell extract, 1.2 mM ATP, 0.85 mM each of GTP, UTP and CTP, 34.0 μg/mL folinic acid, 170.0 μg/mL of E. coli tRNA mixture, 100 μg/mL T7 RNA polymerase, 2 mM each of 20 standard amino acids, 0.33 mM NAD, 0.27 mM coenzyme-A (CoA), 1.5 mM spermidine, 1 mM putrescine, 4 mM sodium oxalate, 130 mM potassium glutamate, 10 mM ammonium glutamate, 12 mM magnesium glutamate, and 33 mM phosphoenolpyruvate). Reactions are conducted by adding 13.3 μg/mL of the plasmid containing the protein of interest to the reaction and setting the reaction in 30 °C incubator for 20 hours without shaking. This plasmid could be the pBR322 plasmid we

are already using. Once again, the phosphate group incorporation takes place at designated Amber stop codons. This method has shown to produce active milligram quantities of phosphorylated human MEK-1 by Oza and colleagues[55].

4.11 Conclusions

The data presented here show that SepOTS and its associated cell lines were able produce small amount of non-phosphorylated recombinant protein. However, the target protein – site specific phosphorylated p53(1-93) – was unable to be obtained. Chromatograms and SDS-PAGE gels suggest that the system was susceptible to changes of temperature and plasmid/cell line combinations, but not in a way that produced a phosphorylated model IDP. That being said, there are still changes that could be made to the current, and a different, cell-free system that could be used to produce recombinant phosphorylated p53(1-93).

REFERENCES

1. Duan, G. & Walther, D. The roles of post-translational modifications in the context of protein interaction networks. *PLoS Comput. Biol.* **11**, e1004049–e1004049 (2015).
2. Minguez, P., Letunic, I., Parca, L. & Bork, P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. *Nucleic Acids Res.* **41**, D306–D311 (2013).
3. Minguez, P. *et al.* Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* **8**, 599–599 (2012).
4. Schweppe, R. E., Haydon, C. E., Lewis, T. S., Resing, K. A. & Ahn, N. G. The Characterization of Protein Post-Translational Modifications by Mass Spectrometry. *Acc. Chem. Res.* **36**, 453–461 (2003).
5. Beck-Sickinger Annette G. & Mörl Karin. Posttranslational Modification of Proteins. Expanding Nature's Inventory. By Christopher T. Walsh. *Angew. Chem. Int. Ed.* **45**, 1020–1020 (2006).
6. Soussi, T. & Wiman, K. G. TP53: an oncogene in disguise. *Cell Death Differ.* **22**, 1239–1249 (2015).
7. Belyi, V. A. *et al.* The Origins and Evolution of the p53 Family of Genes. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).
8. Momand, J., Zambetti, G. P., Olson, D. C., George, D. & Levine, A. J. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69**, 1237–1245 (1992).
9. Joerger, A. C. & Fersht, A. R. The Tumor Suppressor p53: From Structures to Drug Discovery. *Cold Spring Harb. Perspect. Biol.* **2**, a000919 (2010).

10. Lee, J.-H., Kim, H.-S., Lee, S.-J. & Kim, K.-T. Stabilization and activation of p53 induced by Cdk5 contributes to neuronal cell death. *J. Cell Sci.* **120**, 2259 (2007).
11. Humphrey, S. J., James, D. E. & Mann, M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Spec. Issue Syst. Approach Metab. Dis.* **26**, 676–687 (2015).
12. Whitmarsh, A. J. & Davis*, R. J. Regulation of transcription factor function by phosphorylation. *Cell. Mol. Life Sci. CMLS* **57**, 1172–1183 (2000).
13. Hafner, A., Bulyk, M. L., Jambhekar, A. & Lahav, G. The multiple mechanisms that regulate p53 activity and cell fate. *Nat. Rev. Mol. Cell Biol.* **20**, 199–210 (2019).
14. Jenkins, L. M. M., Durell, S. R., Mazur, S. J. & Appella, E. p53 N-terminal phosphorylation: a defining layer of complex regulation. *Carcinogenesis* **33**, 1441–1449 (2012).
15. Ashcroft, M., Kubbutat, M. H. & Vousden, K. H. Regulation of p53 function and stability by phosphorylation. *Mol. Cell. Biol.* **19**, 1751–1758 (1999).
16. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradović, Z. Intrinsic Disorder and Protein Function. *Biochemistry* **41**, 6573–6582 (2002).
17. Bah, A. & Forman-Kay, J. D. Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *J. Biol. Chem.* **291**, 6696–6705 (2016).
18. Bah, A. *et al.* Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106 (2014).
19. 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).

20. 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).
21. 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).
22. English, L. R., Tilton, E. C., Ricard, B. J. & Whitten, S. T. Intrinsic α helix propensities compact hydrodynamic radii in intrinsically disordered proteins. *Proteins Struct. Funct. Bioinforma.* **85**, 296–311 (2017).
23. Uversky, V. N. Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol.* **43**, 1090–1103 (2011).
24. Romero, P. *et al.* Sequence complexity of disordered protein. *Proteins* **42**, 38–48 (2001).
25. Wootton, J. C. & Federhen, S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).
26. Oldfield, C. J. & Dunker, A. K. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu. Rev. Biochem.* **83**, 553–584 (2014).
27. Babu, M. M., van der Lee, R., de Groot, N. S. & Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **21**, 432–440 (2011).
28. Protter, D. S. W. *et al.* Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep.* **22**, 1401–1412 (2018).
29. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18 (2014).

30. Bernadó, P., Bertoncini, C. W., Griesinger, C., Zweckstetter, M. & Blackledge, M. Defining Long-Range Order and Local Disorder in Native α-Synuclein Using Residual Dipolar Couplings. *J. Am. Chem. Soc.* **127**, 17968–17969 (2005).
31. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533 (2002).
32. Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L. & Rexach, M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2450–2455 (2003).
33. Davis, L. K., Ford, I. J., Šarić, A. & Hoogenboom, B. W. Intrinsically disordered nuclear pore proteins show ideal-polymer morphologies and dynamics. *bioRxiv* 571687 (2019) doi:10.1101/571687.
34. Schoch, R. L., Kapinos, L. E. & Lim, R. Y. H. Nuclear transport receptor binding avidity triggers a self-healing collapse transition in FG-nucleoporin molecular brushes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16911–16916 (2012).
35. Labokha, A. A. *et al.* Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. *EMBO J.* **32**, 204–218 (2013).
36. Dumetz, A. C., Chockla, A. M., Kaler, E. W. & Lenhoff, A. M. Protein phase behavior in aqueous solutions: crystallization, liquid-liquid phase separation, gels, and aggregates. *Biophys. J.* **94**, 570–583 (2008).
37. Uversky, V. N., Kuznetsova, I. M., Turoverov, K. K. & Zaslavsky, B. Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. *FEBS Lett.* **589**, 15–22 (2015).

38. Elbaum-Garfinkle, S. *et al.* The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci.* **112**, 7189 (2015).

39. Uversky, V. N. Protein intrinsic disorder-based liquid–liquid phase transitions in biological systems: Complex coacervates and membrane-less organelles. *Complex Coacervation Princ. Appl.* **239**, 97–114 (2017).

40. Zhu, J., Zhang, S., Jiang, J. & Chen, X. Definition of the p53 Functional Domains Necessary for Inducing Apoptosis. *J. Biol. Chem.* **275**, 39927–39934 (2000).

41. Meek, D. W. & Anderson, C. W. Posttranslational Modification of p53: Cooperative Integrators of Function. *Cold Spring Harb. Perspect. Biol.* **1**, (2009).

42. Kruse, J.-P. & Gu, W. SnapShot: p53 posttranslational modifications. *Cell* **133**, 930–30.e1 (2008).

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

44. Theillet, F.-X. *et al.* Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* **54**, 217–236 (2012).

45. Kussie, P. H. *et al.* Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain. *Science* **274**, 948 (1996).

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

47. Bista, M., Freund, S. M. & Fersht, A. R. Domain-domain interactions in full-length p53 and a specific DNA complex probed by methyl NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 15752–15756 (2012).

48. Ithuralde, R. E. & Turjanski, A. G. Phosphorylation Regulates the Bound Structure of an Intrinsically Disordered Protein: The p53-TAZ2 Case. *PLOS ONE* **11**, e0144284 (2016).

49. Wang, X. W. *et al.* p53 modulation of TFIIH–associated nucleotide excision repair activity. *Nat. Genet.* **10**, 188–195 (1995).

50. Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* **49**, 9964–9971 (2010).

51. Bochkareva, E. *et al.* Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15412–15417 (2005).

52. Okuda, M. & Nishimura, Y. Real-time and simultaneous monitoring of the phosphorylation and enhanced interaction of p53 and XPC acidic domains with the TFIIH p62 subunit. *Oncogenesis* **4**, e150–e150 (2015).

53. Ji, X., Huang, Q., Yu, L., Nussinov, R. & Ma, B. Bioinformatics study of cancer-related mutations within p53 phosphorylation site motifs. *Int. J. Mol. Sci.* **15**, 13275–13298 (2014).

54. English, L. R. ELECTROSTATIC EFFECTS ON THE STRUCTURE OF INTRINSICALLY DISORDERED PROTEINS. (Texas State University, 2016).

55. Oza, J. P. *et al.* Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun.* **6**, 8168 (2015).

56. Park, H.-S. *et al.* Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science* **333**, 1151–1154 (2011).

57. Giffin, W. *et al.* Sequence-specific DNA Binding and Transcription Factor Phosphorylation by Ku Autoantigen/DNA-dependent Protein Kinase: PHOSPHORYLATION OF Ser-527 OF THE RAT GLUCOCORTICOID RECEPTOR. *J. Biol. Chem.* **272**, 5647–5658 (1997).
58. Lees-Miller, S. P., Sakaguchi, K., Ullrich, S. J., Appella, E. & Anderson, C. W. Human DNA-activated protein kinase phosphorylates serines 15 and 37 in the amino-terminal transactivation domain of human p53. *Mol. Cell. Biol.* **12**, 5041–5049 (1992).
59. Jette, N. & Lees-Miller, S. P. The DNA-dependent protein kinase: A multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Prog. Biophys. Mol. Biol.* **117**, 194–205 (2015).
60. Lees-Miller, S. P. & Meek, K. Repair of DNA double strand breaks by non-homologous end joining. *DNA Repair Pathw. Assoc. Hum. Dis.* **85**, 1161–1173 (2003).
61. Trigon, S. *et al.* Characterization of the Residues Phosphorylated in Vitro by Different C-terminal Domain Kinases. *J. Biol. Chem.* **273**, 6769–6775 (1998).
62. Pirman, N. L. *et al.* A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation. *Nat. Commun.* **6**, 8130 (2015).
63. Wang, K. *et al.* Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **6**, 393 (2014).
64. Chen, Z. & Cole, P. A. Synthetic approaches to protein phosphorylation. *Curr. Opin. Chem. Biol.* **28**, 115–122 (2015).

65. Lajoie, M. J. *et al.* Genomically Recoded Organisms Expand Biological Functions. *Science* **342**, 357 (2013).

66. Cohen, P. The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem. Sci.* **25**, 596–601 (2000).

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

68. 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).

69. Zhang, J., Krishnamurthy, P. K. & Johnson, G. V. W. Cdk5 phosphorylates p53 and regulates its activity. *J. Neurochem.* **81**, 307–313 (2002).

70. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol.* **5**, 172 (2014).

71. Wang, S. *et al.* The catalytic subunit of DNA-dependent protein kinase selectively regulates p53-dependent apoptosis but not cell-cycle arrest. *Proc. Natl. Acad. Sci.* **97**, 1584 (2000).

72. Folta-Stogniew, E. & Williams, K. R. Determination of molecular masses of proteins in solution: Implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *J. Biomol. Tech. JBT* **10**, 51–63 (1999).

73. La Verde, V., Dominici, P. & Astegno, A. Determination of Hydrodynamic Radius of Proteins by Size Exclusion Chromatography. *Bio-Protoc.* **7**, e2230 (2017).

74. Uversky, V. N. Intrinsically Disordered Proteins and Their Environment: Effects of Strong Denaturants, Temperature, pH, Counter Ions, Membranes, Binding Partners, Osmolytes, and Macromolecular Crowding. *Protein J.* **28**, 305–325 (2009).

75. Hynes, T. R. & Fox, R. O. The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution. *Proteins Struct. Funct. Bioinforma.* **10**, 92–105 (1991).

76. Saito, R., Sato, T., Ikai, A. & Tanaka, N. Structure of bovine carbonic anhydrase II at 1.95 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 792–5 (2004).

77. 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).

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

79. Rademann, K. P. W. Atkins: Physical Chemistry, 4th Edition, Oxford University Press, Oxford, ISBN 0–19–855284-X, 1990. 995 Seiten, Preis: £ 19.50 (Paperback). *Berichte Bunsenges. Für Phys. Chem.* **94**, 1171–1171 (1990).

80. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).

81. 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).

82. Woody, R. W. Theory of Circular Dichroism of Proteins. in *Circular Dichroism and the Conformational Analysis of Biomolecules* (ed. Fasman, G. D.) 25–67 (Springer US, 1996). doi:10.1007/978-1-4757-2508-7_2.

83. Whitten, S. T., Yang, H.-W., Fox, R. O. & Hilser, V. J. Exploring the impact of polyproline II (PII) conformational bias on the binding of peptides to the SEM-5 SH3 domain. *Protein Sci.* **17**, 1200–1211 (2008).

84. Lakshminarayanan, R., Fan, D., Du, C. & Moradian-Oldak, J. The Role of Secondary Structure in the Entropically Driven Amelogenin Self-Assembly. *Biophys. J.* **93**, 3664–3674 (2007).

85. Creamer, T. P. Left-handed polyproline II helix formation is (very) locally driven. *Proteins Struct. Funct. Bioinforma.* **33**, 218–226 (1998).

86. Gursky, O. & Atkinson, D. Thermal unfolding of human high-density apolipoprotein A-1: implications for a lipid-free molten globular state. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2991–2995 (1996).

87. Richer, B. C. & Seeger, K. The hinge region of type VII collagen is intrinsically disordered. *Matrix Biol.* **36**, 77–83 (2014).

88. Brown, A. M. & Zondlo, N. J. A Propensity Scale for Type II Polyproline Helices (PPII): Aromatic Amino Acids in Proline-Rich Sequences Strongly Disfavor PPII Due to Proline–Aromatic Interactions. *Biochemistry* **51**, 5041–5051 (2012).

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

90. Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **85**, 2–9 (2013).

91. Rogerson, D. T. *et al.* Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* **11**, 496 (2015).

92. Hiszczyńska-Sawicka, E. & Kur, J. Effect ofEscherichia coliIHF Mutations on Plasmid p15A Copy Number. *Plasmid* **38**, 174–179 (1997).

93. Wannier, T. M. *et al.* Adaptive evolution of genomically recoded Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3090–3095 (2018).

94. Ferrer, M., Chernikova, T. N., Yakimov, M. M., Golyshin, P. N. & Timmis, K. N. Chaperons govern growth of Escherichia coli at low temperature. *Nat Biotechnol* **21**, (2003).

95. Kandror, O. & Goldberg, A. L. Trigger factor is induced upon cold shock and enhances viability of Escherichia coli at low temperatures. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4978–4981 (1997).

96. Vasina, J. A. & Baneyx, F. Recombinant protein expression at low temperatures under the transcriptional control of the major Escherichia coli cold shock promoter cspA. *Appl. Environ. Microbiol.* **62**, 1444–1447 (1996).

97. Qing, G. *et al.* Cold-shock induced high-yield protein production in Escherichia coli. *Nat. Biotechnol.* **22**, 877 (2004).

98. Chaudhuri, S., Jana, B. & Basu, T. Why does ethanol induce cellular heat-shock response? *Cell Biol. Toxicol.* **22**, 29–37 (2006).

99. Chhetri, G., Kalita, P. & Tripathi, T. An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX* **2**, 385–391 (2015).

100. Kusano, K., Waterman, M. R., Sakaguchi, M., Omura, T. & Kagawa, N. Protein Synthesis Inhibitors and Ethanol Selectively Enhance Heterologous Expression of P450s and Related Proteins in Escherichia coli. *Arch. Biochem. Biophys.* **367**, 129–136 (1999).

101. Novy, R. & Morri, B. *Novagen, Inc. Report 'use of glucose to control basal expression in the pET system'*. vol. 13 (2001).

102. Giacalone, M. J. *et al.* Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* **40**, 355–364 (2006).

103. Mukai, T. *et al.* Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. *Sci. Rep.* **5**, 9699–9699 (2015).

104. Kulkarni, P. *et al.* Phosphorylation-induced conformational dynamics in an intrinsically disordered protein and potential role in phenotypic heterogeneity. *Proc. Natl. Acad. Sci.* **114**, E2644 (2017).

105. Egan, S. M. & Schleif, R. F. A Regulatory Cascade in the Induction of rhaBAD. *J. Mol. Biol.* **234**, 87–98 (1993).