PURIFICATION AND ANALYSIS OF SINGLE-STRANDED BINDING PROTEIN TAIL MUTANTS

By

Matthew Franciskovich

APPROVED BY SUPERVISORY COMMITTEE:

Sheena D'Arcy, Chair

Steven O. Nielsen

Gabriele Meloni

Jeremiah J. Gassensmith

PURIFICATION AND ANALYSIS OF SINGLE-STRANDED BINDING PROTEIN TAIL MUTANTS

by

MATTHEW FRANCISKOVICH, BS

THESIS

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the degree of

MASTER OF SCIENCE IN

CHEMISTRY

THE UNIVERSITY OF TEXAS AT DALLAS

August 2020

ACKNOWLEDGEMENTS

I would like to thank the many people who have influenced my life up until this point; I would certainly not be the person I am today without their guidance, teachings, and love. From the smallest gesture to the largest impact, I thank each and every person who had a hand in helping me find my path, and to every person who will teach me in the future.

I thank my family, blood and covenant, for all the happiness and support you have given me. To all my parents, Linda Ronco, Pablo Ronco, and Scott Franciskovich, who always pushed me to reach further beyond what I could see, knowing I could reach higher, and showing me that the sky will never be the limit as well as there is always a lesson to learn from someone. Through grit and brimstone, they helped me to have a happy upbringing and ensured I would be taken care of, even at the cost of their own comfort. I would not have taken the path I did without the sacrifices they made to secure my dreams. To my older sister and brother, Mandy and Tim, who always had my back and gave me support when I needed it, even when I didn't want it. They have seen me at my lowest as well as at my highest, and I look forward to a future where we all stand together as we always have. The two people I share more experiences with than anyone else, I thank you for the lessons you taught me. To my best friends, Damian, Ovidiu, Jacob, Matt, Gabriel, and Stephanie who are my brothers and sister in spirit, having laughed, cried, challenged, and shared with me all the good and bad times over the past 15 years. To my loving girlfriend, Jill, who has been with me through peace and calamity, having my back just as much as I have hers. Thank you for putting up with my endless shenanigans and laughing with me when I do something dumb.

I would also like to thank the wonderful people of the D'Arcy lab for all their help these past two years. They have seen me fall down many times and have always helped me back up. I especially thank Dr. Sheena D'Arcy for her patience and assistance in walking through the sometimes convoluted master's degree process with me, helping and communicating with me, even if it wasn't something I wanted to hear. I am extremely grateful for all your help. I would also like to thank my lab-mates, Dr. Kyle Murray who first introduced me to the lab, Dr. Prithwijit Sarkar who walked me through basically everything, and Noushin Akhavantabib, Naifu Zhang, Darby Ball, Katayoun Morakabi, Joy Shaffer, Kiran Tripathi, Trang "Kelly" Nguyen, Oladimeji Sunday Olaluwoye, Katie Blackwell, Victoria Neumann, Adam Timmerman, and Henock Abera. You have all made this truly feel like a family to me and allowed me to open up from the shy person I was when I first got here. My only regret is that we didn't have more time together and that the last few months of my time couldn't be spent with all of you in the lab. Thank you all for making my time here memorable.

Also, a big thank you to the Dr. Winkler lab for help in crystallization for small trays, and thank you to Dr. McCandless for his help in analyzing cryoconditions. Without your help, I would not have been able to get this far in my project. Your kindness and patience spoke volumes to me, and I am extremely grateful to you.

July 2020

PURIFICATION AND ANALYSIS OF SINGLE-STRANDED

BINDING PROTEIN TAIL MUTANTS

Matthew Franciskovich, MS
The University of Texas at Dallas, 2020

Supervising Professor: Sheena D'Arcy

Single stranded binding protein (SSB) is a prokaryotic DNA protein that binds to single stranded DNA during times when the DNA is rendered from its double stranded form during times of genetic recombination or DNA damage in order to stabilize and protect it from further unnecessary harm. The protein exists as a tetramer with each monomer being made of an N-terminal and C-terminal domain. The C-terminal domain is made of two smaller sub-domains, both of which have yet to resolve properly in a crystal structure, named the intrinsically disordered linker and the acidic tip, with limited understanding on how they function and relate to other proteins and SSB itself. Due to the disordered nature of its C-terminal domain limiting the ability to yield a concise crystal structure, much of the function and nearly all of the structure of the C-terminal domain has yet to be identified. While some function has been determined for these disordered regions, its relationship with other binding partners, DNA, and itself have yet to be fully determined.

Through the use of SSB mutants D4A4(where aspartic acid residues 191, 193, 194, and 195 are changed to alanines) and ΔC8 (where residues 191 through 198 are removed) in purification experiments and analytical techniques such as size exclusion chromatography, x-ray

crystallography, and multi-angle light scattering, some potential hypotheses were built based on the resulting data. Overall, D4A4 showed a higher ability to crystallize compared to is counterpart but was plagued with difficulty when trying to concentrate to high amounts. It is hypothesized that D4A4 was more susceptible to degradation and/or aggregation compared to ΔC8 due to many issues seen throughout the purification process such as D4A4's band in each gel splitting in two and increased difficulty concentrating, but no definite statement could be made regarding this possibility. Conditions tested during x-ray crystallography provided useful information for use in future crystallization attempts; however, because a crystal structure failed to be uncovered, more progress must be made to establish the best conditions for crystallization. Initial tests show that HDX is a valid approach for comparing these mutants to the full-length protein to uncover structural differences. The speculative function of these areas of interest can only be confirmed through more crystallography studies and additional experiments with binding partners and DNA used in tandem with other SSB mutants. Altogether, the purifications and initial results in this thesis lay a solid foundation for future work assessing the role of the SSB C-terminal domain.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 DNA Packaging in Eukaryotes - Chromatin

Within the nucleus of nearly every eukaryotic cell lies its DNA, most commonly found in its tightly packed, supercoiled form of chromatin. Chromatin exists in two primary forms, heterochromatin and euchromatin, which are the tightly packed and loosely packed forms of chromatin, respectively.[1] Due to its tightly wound structure, heterochromatin normally cannot be transcribed, while euchromatin is easily transcribed due to its loosely wound state.[2] In eukaryotes, chromatin fibers are primarily comprised of nucleosomes, octameric histone proteins wrapped by 145-147 base pairs of DNA, which are then supercoiled to fit inside a cell's nucleus. The histone octamer is made up of two copies each of histone subunits H2A, H2B, H3, and H4.[3] There also exists a single histone H1 that commonly lies on the outside surface of the nucleosome with the linker DNA.[4] Prokaryotes do not exhibit the same genetic packaging as eukaryotes, lacking much of the machinery that eukaryotic cells take advantage of, and as a result, their method of DNA storage is different.

1.2 DNA Packaging in Prokaryotes - Nucleoid

Much of what is known about prokaryotic cells came from intensive study on Escherichia coli (E. Coli), a bacterium found in the human colon. Contrary to eukaryotes with their linear DNA and histones, prokaryotes' chromosomes are a circular piece of DNA, with only those from archaea having histones for nucleosomal packaging.[5,6] Their cells also lack a nucleus and nucleolus, meaning the chromosome is found in the cytoplasm of the cell, but this does not mean it is without

control. The chromosome is localized into a tightly wound structure known as a nucleoid by intense supercoiling with the assistance of various proteins such as DNA gyrase, topoisomerase I, and HU protein (a protein similar to histone H2B).[7] HU and topoisomerase I help to create the sharp bends in the DNA necessary for supercoiling, organizing it into multiple conformations which are then supercoiled around tetramers of HU, similarly to how eukaryotic DNA wraps around histones.[8] After it is finished, DNA gyrase and topoisomerase I help to maintain the supercoiled state. Not all the DNA in a prokaryote lies within the nucleoid, some smaller circular pieces exist as plasmids.[9,10] These additional pieces of genetic information are not large enough to convey essential commands for life, instead carrying information for various genes that may become beneficial given certain circumstances. Even though the nucleoid is tightly wound, the rate of replication and transcription are much higher than those found in eukaryotes. While the nucleoid is free in the cytoplasm, it is too tightly wound for transcription and replication enzymes to access the DNA.[2] When the DNA needs to be replicated or transcribed, small regions of the chromosome extend from the nucleoid to unwind and associate with ribosomes and other enzymes.[8]

1.3 DNA

DNA stands for deoxyribonucleic acid and is the genetic code for all living organisms on earth. As previously mentioned, in eukaryotes, it is stored in the nucleus compacted in chromatin, while prokaryotes instead possessing a single chromosome formed from a singular circular DNA located in a region of the cell known as a nucleoid.[1] Commonly in a double stranded form, DNA is made of two polynucleotide chains coiling around each other to form a coiled ladder known as the double helix.[11] Each rung on this ladder is made of a single nucleotide made from a deoxyribose sugar molecule, a phosphate group, and one of four different nucleobases. These bases are

cytosine, guanine, adenine, and thymine, and are commonly abbreviated as C, G, A, and T, respectively.[12] On each strand of the helix, multiple nucleotides covalently bind to one another via phospho-diester linkages between the sugar of one nucleotide and the phosphate group of the adjacent nucleotide. At the edge of each nucleotide is one of the four bases previously mentioned, each nucleobase pairs complimentarily with another, with C and A pairing with G and T, respectively. Each strand of DNA has distinct directionality, with a 3' and 5' end, and when two strands bind to one another, they do so in an anti-parallel manner with the 3' end of one strand pairing with the 5' end of the other. The opposing DNA strands bind to one another through hydrogen bonding between their base pairs, forming the double stranded DNA helix.[13] The double stranded form of DNA is stable; however, to access the DNA and read its genetic code for other biological processes, the strands must be unwound and separated to access the bases and multiply, interpret, or repair them.

1.4 DNA Upkeep

For the cells of each organism to grow and multiply, they must each contain their own copy of DNA. Because the copy of DNA that every daughter cell has must be identical to that of its parent cell, the DNA must be replicated very specifically. During DNA replication, various enzymes present within the nucleus act to unwind and stabilize the DNA, while others function to read the bases present and bind their compliment nucleotides to form new DNA strands.[14] This process needs to be very precise, and as a result has many different proofreading steps to ensure proper base pairing. The DNA is opened by DNA helicase creating what's known as a replication bubble, where DNA polymerase attaches to on each strand and begins to form the new complimentary strands, reading the DNA in a 3' to 5' direction.[15] Following complete DNA

replication, the cells then undergo mitosis, creating new daughter cells.[16] Replication produces identical genomes in cells within an organism; however, this genetic code is not completely passed down to offspring, instead undergoing DNA recombination.

In eukaryotes, when the genetic code is prepared and packaged within the gametes of an organism, it undergoes a form of DNA recombination, known as meiosis, to diversify the genetic code of the offspring. Similar to mitosis, meiosis begins with DNA replication forming two pairs of sister chromatids. Following this, some of the chromatids undergo double stranded DNA breaks, resulting in 5'-duplex junctions with 3'-single stranded DNA overhangs.[13] Another sister chromatid then forms a double holiday junction with the broken DNA, creating a crossover between the two DNA strands.[17] Once resolved, the resulting DNA molecules have parts of the other within its code. This is preformed between each pair of sister chromatids, resulting in four possible genetically distinct DNA strands within the gametes.[18] While recombination plays a large role in maintaining genetic diversity within offspring, it is also important in DNA repair pathways, especially in prokaryotes, which do not undergo meiosis.[19] While double stranded breaks are present in meiosis, they can also result from outside influences like anticancer chemotherapeutic drugs and ionizing radiation, as well as during replication if the DNA is in a weakened state.[20]

There exists three methods of repairing double stranded breaks; however only one focuses primarily on recombination, the appropriately named homologous recombination.[21] For homologous recombination to take place, an identical or nearly identical DNA sequence is used as a template to repair the break. A sister chromatid is used if the break happens right after replication, and a homologous chromosome at other times.[22] The enzymes used for this process are very similar to those used for the crossover event in meiosis. Homologous recombination is commonly lost in

cancerous cells, as well as other DNA damage responses, although to a lesser degree.[23] Due to the importance of DNA proliferation, health maintenance, and genomic diversity, there are many factors working to maintain the integrity of our genetic code, especially during periods of increased risk. As mentioned previously, DNA must be unwound and its strands separated to be interacted with by much of the cell's machinery, but the strands of DNA can become separated and exposed due to external and internal damage. During times such as these, certain proteins can interact with the DNA to help stabilize the exposed and reactive single stranded DNA.[24] One protein of importance that does this is single-stranded binding protein (SSB).

1.5 SSB

SSB is a 178-residue protein with a molecular weight of 18975 Da and a pI of 5.44 present in E. coli. It plays two essential roles within the cell: stabilizing single stranded DNA intermediates during DNA processing, and mediating binding to up to 14 DNA binding proteins that constitute the SSB interactome.[25] In its bacterial active state, SSB is a tetramer with each monomer consisting of two dominant domains which can be defined by proteolytic cleavage, the first being the N-terminal domain. Characterized by the first 115 residues, the N-terminal domain is the more conserved domain of the two, with pI of 8.01. The second domain is the C-terminal domain made from the last 63 residues, which can be further broken up into two sub-domains, the intrinsically disordered linker (IDL) and the acidic tip.[26] The IDL exists from residue 116 to 170, and has a pI of 9.6, while the acidic tip makes up the final eight residues and has a pI of 3.32.[25] While there does not currently exist a crystal structure showing SSB in its entirety, what is shown confers some of the secondary structure for the protein within the N-terminal domain.

There are seven aligned beta sheets and one alpha helix in the N-terminal domain, which are located as follows: β1 from Val5 to Leu14, β2 from Glu19 to Tyr22, β3 from Val29 to Ser39, β4 from Gln51 to Leu59, β5 from Gln76 to Trp88, β6 from Arg96 to Val101, and β7 from Gly107 to Met111. The alpha helix ranges from Lys62 to Tyr70.[27] The most complete crystal structure to date fails to show the entirety of the protein sequence, only going as far as residue 145 for nearly each subunit, as a result the full length of the C-terminal domain does not show, including the complete acidic tip. Each of the domains play a role in its relationship with DNA and its binding partners. The N-terminal domain binds to DNA in a non-specific manner via the wrapping of single stranded DNA around the SSB tetramer using a large network of electrostatic interactions with the DNA phosphor-diester backbone and base stacking interactions with the nucleotide bases.[28]

The function of the C-terminal domain is not fully understood, with only minor evidence helping to direct our understanding of its function, but one important understanding is that it plays a critical role in the release of DNA from SSB.[29] After deleting residues 132 through 137 of the IDL, interactions with certain subunits of the DNA polymerase III complex were eliminated.[30] Residues 116 through 174 are highly disordered, even when bound to single stranded DNA, adopting various conformations.[29] As for the acidic tip, studies using the SSB mutant SSB ΔC8 (where the last eight residues of the protein are eliminated) have suggested that 171 through 178 are important for mediating target protein binding.[31] With all these data, it is clear that overall, cooperativity is critical for the function of SSB, enabling quick single stranded DNA binding, leading to its protection and/or conversion into a more proper substrate for processing. It is also understood that IDL is essential to two important areas of SSB function: the aforementioned cooperative single stranded DNA (ssDNA) binding and its interaction with partner proteins.[32]

Meanwhile, even though the data shows that the acidic tip is important with respect to target protein binding, this does not mean that the acidic tip is also a protein-protein interaction domain. While removal virtually eliminates target protein binding, removing it also effects cooperative binding to ssDNA as well as the stability of the SSB tetramer.[25] Needless to say, with so much still not understood about SSB, great effort needs to be made to purify, crystalize, and analyze the protein to completely understand its function and importance biologically.

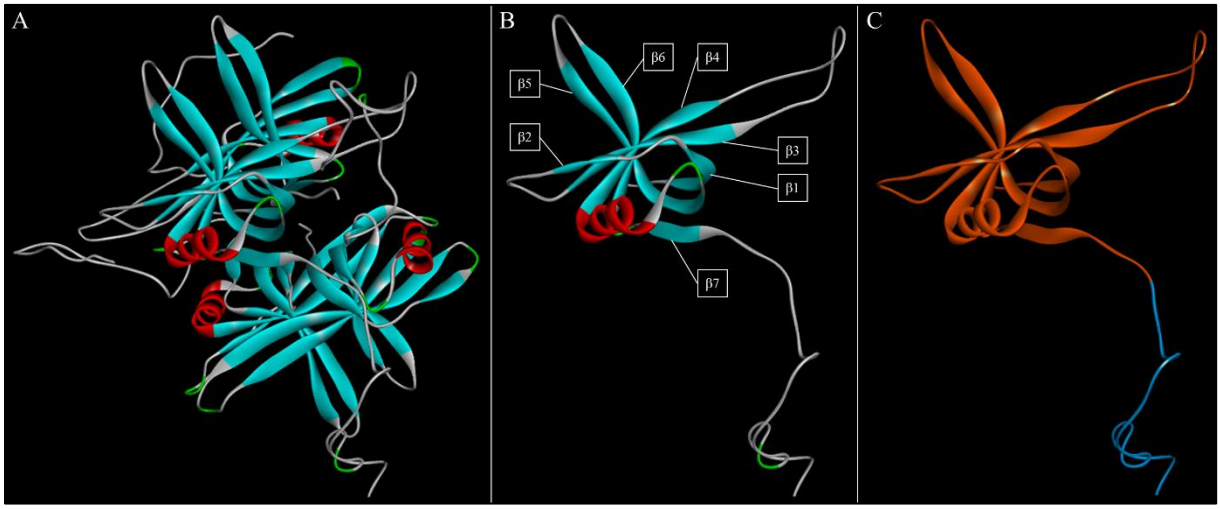

Figure 1. Crystal Structure of SSB (PDB 1QVC). (A) SSB Tetramer. (B) SSB Monomer with β-sheets 1 through 7 labelled. (C) Division of subunits showing N-terminal domain in orange and C-terminal domain in blue.[27]

It is the goal of this research project to uncover some of the information leading to the determination of a complete crystal structure in order to determine the function the IDL of SSB. To do this, an in vitro approach was adopted to compare and contrast two SSB mutants: D4A4 which has its last four aspartic acid residues within the acidic tip changed to alanines, and ΔC8 which completely lacks the last eight residues making up the acidic tip. Both of these proteins also contain an N-terminal his-tag for aid in purification. The aim of this project was to purify these proteins and analyze their structures using light scattering, x-ray crystallography, and hydrogen-

deuterium exchange coupled to mass spectrometry. Some of the specific questions we tackle were how removing the last eight residues in the acidic tip in ΔC8 and how altering the last four aspartic acid residues to alanine's effected the function, stability, and interactions of SSB. We also aim to derive possible function of the IDL and acidic tip from the data obtained from our experimental approaches outlined below.

1.6 Protein Purification – Ni-affinity Chromatography

There are many techniques used for purification of a target protein; however, first the cells from which the protein of interest grows must be lysed and their contents extracted. Based on the nature of the protein, it exists in the supernatant solution following lysing and centrifugation. With all the larger, more dense objects of the cell out of the solution, this leaves mainly proteins, but purifying one protein out of the large collective that now exist within the supernatant can be a challenging task.[33] This is where taking advantage of the target proteins properties can come in handy. Each protein has various qualities that naturally distinguish it from others within the cell, from size, charge, pI, mass, and length, all of these and more can be used to purify a protein out of a mixture. But for a more specific approach, an affinity tag can be added to the protein by using certain plasmids or vectors. These vectors cause the bacteria to produce a certain molecule a certain way.[34]

An example of one such vector is one that genetically attaches a polyhistidine tag to one of the ends of the protein, most commonly with six histidine residues, known as a Histag.[35] With each copy of the protein containing a Histag, the supernatant solution can then be sent through an affinity column filled with a nickel NTA resin. When the supernatant is sent through the column,

the histidine residues bind the nickel in a specific manner via chelation, locking the proteins in the column until they are washed out with an elution buffer, commonly one with a high concentration of imidazole around 500 mM to completely clean out all protein from the column.[36] The process as a whole is known as Ni-affinity chromatography, with the columns used named Histrap, and is an effective means of purification, but not a perfect one. HisTrap is often a preliminary purification technique because polyhistidine columns tend to hold on to other proteins as impurities. Additional purification steps must be performed following this to ensure a pure protein sample.

1.7 Protein Purification – Size Exclusion Chromatography

Because most of the proteins left over from solution are non-uniform in size and shape, separating based on these properties can yield a much purer sample. Size exclusion chromatography (SEC) is a purification technique that does just that, purifying a sample based on the sizes and shapes of the proteins and large molecules within it.[37] The separation is dependent on the resin being used within the size exclusion column, which contain small beads made up of agarose, polyacrylamide, or dextran polymers sold under the names Sepharose, Sephacryl, or Sephadex, respectively.

When the sample is sent through the column, the various proteins within the solution flow at different rates depending on their size and shape. Pores present within the beads play a large role in these different flow rates because the smaller proteins will get caught within the pores when flowing down the column, and the smaller the protein, the more often it will get caught by the pores of the beads throughout the column.[38] Because of this relationship with the pores, the larger proteins and other macromolecules flow faster than the smaller ones, giving a size gradient in all

the fractions collected. Various beads with different pore sizes exist, these pores will have different size thresholds depending the protein of interest ranging from the Sephadex G-10 with a threshold of 0 to 700 Da, with much larger ones such as Sepharose 2B with a range of 2×10^6 to 25×10^6 Da.[39] Using control diagrams for a given bead type, the shape can also be found if the data does not support what is to be expected, this can happen when a protein may exist as an oligomer or when it denatures in the SEC buffer, giving higher than expected results. The data given from the SEC shows proteins of various sizes as well as nucleic acids and how prevalent they are in the solution. If the shape and subunit organization of the protein of interest is known, then the appropriate fractions can be collected to have a potentially pure protein sample. Following this, the sample can then be concentrated further to prepare the samples for analytical observation.

1.8 Light Scattering and X-ray Crystallization

Following protein purification and concentration, the protein is then ready to be analyzed with more precise analytical techniques. Size exclusion chromatography paired with multi-angle light scattering (SEC-MALS) is one technique that allows for the determination of the absolute molar mass based on how light is scattered by the solution. The laser is shone on the sample as it flows through the system at a variety of angles, the data obtained is dependent on the angle of the resulting light as it passes through the sample. The intensity of the resulting light scattering is proportional to the molecules molecular weight, as well as the protein concentration of the solution.[40] Concentration can be measured by ultraviolet absorption or differential refractive index (dRI).[41]

Monomeric proteins put through SEC-MALS are separated from the other impurities found within the system such as fragments, soluble oligomers, and aggregates. The SEC plays an important role in this technique, as without separating the impurities, MALS is unable to accurately identify the protein of interest and will give an incorrect molecular weight. If the resulting molecular weight for the monomers, dimers, etc. can be seen in the data analysis of SEC-MALS, then the percentage of each can be deconvoluted from the average by relating this data to the weighted average.[42] The measurement capabilities of MALS for absolute molar mass is incredibly wide, being able to measure independent of structure and shape from 200 Da to hundreds of millions of Da. From this kind of data, the mass, size, and degradative health of the protein can be assessed to a high degree. On top of size and mass measurements, MALS can also provide insight into protein-protein interactions, molecular branching, and molecular conformation.[43,44]

With a highly pure protein, the sample may then undergo crystallization. Protein crystallization is necessary to create a crystal structure of the protein, used for understanding the various areas of the protein and helping to assign functional groups in tandem with a sequence. Knowing what causes protein crystallization is important, as there are a variety of factors that influence crystallization, from pH and concentration to the kind of salts that are used and temperature in the suspension solution.[45] All of these different parameters are systematically matched against one another at varying levels, allowing for as much information as possible to be collected to aid in the crystallization process. Once an optimized set of conditions can be found, they can be used to provide the best possible crystals for use in structural studies. Because of the large number of variables to consider when carrying out these studies, intense consideration must be had to design the most effective and appropriate crystallization solution.[46] The length of

crystallization is not consistent across all proteins, some might crystallize quickly, while others may need much longer to form high enough quality crystals to be used. The purer the protein, the more likely the proteins will crystallize. Even if a crystal is not obtained, some information can be obtained using other approaches such as hydrogen-deuterium exchange (HDX).

1.9 Hydrogen-Deuterium Exchange

HDX is an analytical technique used to research protein behaviour by monitoring the link between structure, conformation, and function. It does this by measuring the change in mass as a result of isotopic exchange between the backbone amide hydrogens of proteins with the deuterated solvent. The rate of exchange is dependent on how the protein folds as well as its amino acid sequence, and can reveal a myriad of information relating to protein conformation, protein binding, allosteric effects, folding dynamics, intrinsic disorder, and much more.[47–49] HDX is paired with mass spectrometry (MS) because of its effectiveness in accommodating large proteins greater than 100 kDa while being able to accept low concentrations and monitor minute changes in mass.[47] Protein conformation affects exchange rates between hydrogen and deuterium, as a result, measuring this rate can also reveal information about the protein with relation to protein interactions, modifications, temperature, pH, and other properties.[48] Because of this intense dependence on other factors, maintaining consistent experimental conditions is paramount to the integrity of the results. Many controls are required to ensure the reaction proceeds properly so as not to introduce possible bias or produce imprecise results. To confirm the precision of the results for accurate data acquisition, repeats of experiments must be performed, with multiple samples being prepared in parallel to limit the introduction of variety in sample preparation.[47]

Prior to beginning an HDX experiment, a few buffers need to be designed, primarily the equilibration (E), Quench (Q), and Labelling (L) buffers. The E buffer is the same buffer as the protein stock. The Q buffer is needed to quench the HDX reaction, denature the protein, and reduce any disulfide bonds in the protein. The L buffer is identical to the E buffer, but with D_2O in place of H_2O. After the HDX reaction and quenching, the protein is digested on a pepsin column, where it is broken down into peptides. It then flows through a trap and an analytical C18 column where reverse-phase separation occurs as a non-polar solution elutes, polar molecules first, followed by less polar or nonpolar molecules.[50]

The resulting samples go into the MS, where they undergo electrospray ionization (ESI) to give a charge to all possible molecules. Those that are not charged are then filtered out in the ion filter, with the rest of the sample moving on to the quadrupole, another purification section that separates based on the mass to charge ratio (m/z) of each molecule flowing through.[51] The molecules then go through a mobility chamber that puts the sample through gas phase electrophoresis, separating by size with smaller surface area molecules processing faster due to less collisions with the gas.[52] The molecules are then passed through a collision cell which applies voltage to the peptides and fragments them before finally arriving at the ToF, or time of flight, section which accelerates the ions using an electric field within a vacuum. Each ion with the same charge will have the same kinetic energy, but with varying velocities based on their m/z, with heavier ions having lower speeds. The difference in time that it takes for each ion of the same charge to reach the detector is measured and used to identify the molecule.[53] After injections are complete, the data given will relay how much of the protein was able to exchange with deuterium with peptide level resolution. The data can then be used to establish protein structure,

conformation, dynamics and many more, which can be used together with a possible crystal structure to establish the functional areas of each part of the protein.

CHAPTER 2
METHODS

2.1 Small Scale Protein Expression Test in E. coli

All cell strains were provided by Dr. Piero Bianco of the University of Buffalo, Jacobs School of Medicine and Biomedical Sciences in Buffalo, New York. All reagents used were supplied by Sigma Aldrich and Fisher Scientific. All LB broth made consisted of NaCl, tryptone, yeast extract, and diH_2O, followed by rigorous autoclaving. Agar media similarly was made from NaCl, tryptone, yeast extract, and diH2O with the addition of agar, and was also autoclaved prior to use. All gels used were SDS-PAGE gels created were made from diH_2O, 40% w/v acrylamide 37.5:1 mix, Tris buffer, 10% SDS, 10% ammonium persulfate, and TEMED, with Tris pH 8.8 used for the resolving gel and Tris buffer pH 6.8 for the stacking gel. Every molecular ladder used in SDS-PAGE gels were Thermo Scientific Spectra Multicolor Broad Range Protein Ladders Prior to any large-scale expression, a small-scale expression test was performed on the SSB mutants D4A4 and ΔC8. Both of their molecular weights and pI's are different from that of wild type due to the presence of his tags on their N-termini. D4A4 has a molecular weight of 20,962.28 Da and a pI of 9.77, while ΔC8 has a molecular weight of 20,173.33 Da and a pI of 9.77.

Four expression tubes were filled with 5 mL of LB broth, with two of them having a final 1x ampicillin concentration to accommodate the ampicillin resistant D4A4 vector, and the other two having a final 1x kanamycin concentration for the ΔC8 vector. Four p100 pipette tips were briefly run through the flame of an ethanol lamp to sterilize them before being used to take up a single colony of bacteria (two colonies from each mutant) from the petri dishes shipped by Dr. Bianco to UTD. Each tip was dropped into the expression tubes and allowed to incubate at 37^oC

for roughly five hours until the OD_{600} reached 0.6, shaking at 180 rpm. Following this, 900 μL of each mutant sample was taken out and placed into small storage tubes along with 100 μL of 100% glycerol and flash frozen in liquid nitrogen (LN2) before storing in a -80°C freezer to create a cell stock for future use.

1mL of each sample was taken out and centrifuged for five minutes at 12,000 rpm. The supernatant solution was pipetted out and the uninduced pellet was stored in a -20°C freezer. The bacteria remaining in the expression tubes were then treated to an IPTG stock solution for a final concentration of 0.3875 mM IPTG to induce the cells and begin protein production. These samples were kept in the incubator for an additional three hours at 37°C, shaking at 180 rpm. Following the incubation, 1 mL of each sample was once again taken out and centrifuged using identical parameters, keeping the pellet and freezing it again. The following day, the pellets were thawed and resuspended in 200 μL of 1x Bugbuster sample for a final concentration of 0.2x (1/5 of the culture volume). The resuspended pellets were then placed on a rotating mixer at room temperature for 20 minutes. 10 μL of each sample were pipetted along with 10 μL of loading dye into individual tubes, with duplicates made for each sample. The samples were then boiled in their tubes at 90°C for 10 minutes before loading the samples onto an 18% gel and running them through electrophoresis at 220 V for 70 minutes. Throughout this experiment, every gel following electrophoresis will be placed into Coomassie blue stain and shaken for 15 minutes, removing the stain and then placing it into 100% destain (1 L of destain is made from 500 mL of methanol, 400 mL of diH_2O, and 100 mL of glacial acetic acid) for an additional 15 minutes, followed by a final shake in 50% destain for three hours before being imaged by Biorad gel imaging dock and the file saved to an external drive.

2.2 Large Scale Protein Expression

The agar media was microwaved for roughly two minutes to liquify the stock, then allowed to cool to allow safe handling. An antibiotic of choice was added to the cooled liquid media for a final concentration of 1x. Each media bottle contains 100 mL of media, which was used to make six plates of its corresponding antibiotic media. 12 empty petri dishes were taken and placed on a workbench sterilized with 100% ethanol. An ethanol lamp was lit and used to sterilize all objects associated with the plating process by briefly running them over the flame. The mouth of the media bottle was also run over the flame before and after each pour. The tops of the dishes were run over the flame before closing the plate completely using parafilm, wrapped in groups of three with their respective antibiotic and stored in the 4^{o}C refrigerator.

From the -80^{o}C stock, a single, flame sterilized p100 pipette tip was used to collect a small sample from the surface and spread across its appropriate agar plate. Using a glass spreader stored in ethanol and placed in fire to fully remove any lingering alcohol, the cells were evenly distributed across each plate. The rod was then run through the flame again before being placed back in the ethanol; this was repeated for each plate. The resulting plates were labeled with their contents and placed into an incubator at 37^{o}C overnight. The same day, six 1 L flasks were used to make 6 L of LB broth. The samples were all autoclaved and placed on a sterilized workbench until used.

The following day, 1 mL of the appropriate antibiotic was added to each flask of broth two achieve a 1x concentration and mixed well. On a sterilized workbench, 10mL was taken from each flask into a falcon tube and cells were collected from their plates and inoculated into the 30 mL of broth using flame sterilized suspension loops. To remove any large clusters of cells remaining in

the solution, the tubes were vortexed thoroughly. The samples were then evenly distributed back into the flasks and incubated at 37°C for roughly three hours, shaking at 170 rpm until the OD_{600} reaches around 0.6, after which IPTG was added to each flask for a concentration of 500 mM to induce protein production, after which the flasks were allowed to continue incubating for an additional four hours. They were then removed, and their contents poured into preweighed and numbered 1 L centrifugation tubes, correcting their final masses with diH_2O to be equivalent. They were then centrifuged at 4,000 rpm and 4°C for 30 minutes.

During this time, 500 mL of lysis buffer was made to the following specifications: 20 mM Tris pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 0.2 mM AEBSF. The resulting supernatant solutions were then removed, and each centrifugation tube was weighed with only the pellet to calculate its mass by subtracting the previously obtained mass of the tube. Each individual pellet was resuspended using 15 mL of lysis buffer, storing all similar mutants together in falcon tubes and freezing in the -20°C freezer for later use. 1 L of Histrap buffers A and B were made this day as well, with Histrap A being 20 mM Tris pH 7.5, 300 mM NaCl, 0.2 mM AEBSF, and 30 mM imidazole, and Histrap B being identical to Histrap A except for containing 500 mM imidazole. Both Histrap buffers were filtered with a 0.22 mM Millipore filter.

2.3 Cell Lysis

The tubes were taken out of the freezer the following day and thawed alongside the protease inhibitors leupeptin, aprotinin, and pepstatin, as well as stock Benzonase with a concentration of 25 units per 1 μL. Once everything was thawed completely, the protease inhibitors were added to each sample for a final concentration of 1x. The Benzonase was added to satisfy the experimental

requirement of 250 units per gram of pellet. From this moment onward, the samples were kept on ice as much as possible to discourage protein denaturing. Once complete, the samples, lysis buffer, diH_2O, and 100% isopropanol were taken to the homogenizer along with four empty falcon tubes. The homogenizer was prepared by setting the pressure to 18,000 psi and running isopropanol through the machine, followed by diH_2O, and lysis buffer, allowing all the reagents to go to waste. One of the samples was then poured into the machine for homogenization, running the complete volume through it at least 3-4 times, collecting it each time in its original tube except for the final collection, where a new falcon tube was used. Once completely collected, lysis buffer was added to the homogenizer, and 50 mL of the lysis buffer is collected in another falcon tube. diH_2O was then run through the machine and allowed to go to waste, followed by isopropanol. This process was repeated for the other sample, ending on isopropanol and sealing the homogenizer sample reservoir.

The four tubes were then placed in a rack and slow spun for 20 minutes, allowing for the Benzonase to digest as much nucleic acid from the sample as possible. Following this, the two falcon tubes of each sample were combined and 10 μL of the sample collected as whole cell lysis and stored for later use. They were centrifuged at 16,000 rpm and 4^{o}C for 30 minutes, collecting the supernatant and throwing out the pellet, with 10μL of the soluble sample at this stage also collected. The samples were run through a previously equilibrated GE Lifesciences AKTA Start equipped with a 5 mL Histrap Fastflow column for use in Histrap purification. The sample was run through completely using Histrap A, with a final 10 μL sample collected from the resulting flow through. Once all the sample was run through the column, the buffer was changed to a gradient flow of 100% Histrap B. As Histrap B concentration increases, more protein comes off

the column, which was collected over 30 fractions, after which the column was cleaned and equilibrated, with the process repeated for the next sample.

The complete range of fractions (fraction one excluded) along with the three samples taken previously were then run on 18% gel. Chromatograms for each sample run were saved on an external drive. The fractions were then combined and transferred into Corning centrifugal concentration tubes with molecular weight cutoffs of 10,000 Da for ΔC8 and 5,000 Da for D4A4, which were then placed in a large tabletop ThermoScientific centrifuge for one hour at a time at 4,000 rpm and 4^{o}C, collecting the flow through in a falcon tube stored at 4^{o}C and refilling the tube after each spin. Once the final volumes decreased to between 3 and 6 mL, their concentrations were calculated based on the absorbance readings from a Thermofisher Nanodrop system. Different cutoffs were used due to a much higher difficulty in concentrating D4A4.

2.4 SEC

Prior to running SEC, SEC buffer was made with a composition of 20 mM Tris pH 7.5, 300 mM NaCl, and 0.5 mM TCEP, filtering the same way as Histrap A and B. The column used was a HiLoad 16/600 Superdex 200 pg which was equilibrated overnight at 4^{o}C. A 3 mL loading loop was attached and thoroughly cleaned using SEC buffer before injecting 3mL of sample, storing any leftover sample in the -80^{o}C freezer. Once the protein comes completely off the column, the appropriate fractions were then collected and run through an 18% gel. Similar to Histrap, the fractions were then collected and concentrated. This process was repeated for both SSB mutants, with the concentration for D4A4 being much lower than that of ΔC8 based on the chromatogram. The degree of concentration for the proteins at this point was dependent on the

next step that would be taken. For crystallization, the protein concentration for D4A4 and ΔC8 would need to be approximately 5 mg/mL and 16.5 mg/mL, respectively. For SEC-MALS, the protein concentrations would need to be 100mM for both mutants and for HDX the concentration would need be 1 μM for proper use in a 50 pM injection. Once the protein was concentrated appropriately, the necessary volume was put aside for the experiment of interest while the rest of the protein was stored in the -80^{o}C freezer.

2.5 SEC-MALS

Prior to running the experiment with 100 mM samples, 20 mM samples were used with below average results, leading to the use of 100 mM. With both proteins concentrated, they were then injected into a GE Healthcare Superdex S200 Increase 10/300 GL column as well as a Wyatt miniDawn TREOS and Optilab T-rEX refractive index detectors. Buffer used for equilibration was identical to SEC buffer but filtered a second time using a 0.1 mM filter. Once the instrument is ready, the sample was loaded and allowed to run to completion, with the resulting data corrected to remove any extraneous noise. With a more accurate representation of the purity of the sample, pure new samples of both SSB's were taken from the freezer and concentrated for use in crystallization and intact mass spectrometry. Because of complications with D4A4, only ΔC8 was able to be run on the LC/MS system.

2.6 Crystallization

Once the volume reached around 3 mL, a smaller Millipore concentrator tube was used with a smaller tabletop Thermo Scientific concentrator. These were spun at 16,000 rpm for five minutes at a time until the desired concentration was obtained. The fully concentrated proteins

were taken and plated in a 1:1 sitting drop setup using 10 μL drop sizes with six different crystal blocks: PEGRx HT, Crystal Screen HT, Index HT, SaltRx HT, and Peg Ion HT. The leftover protein samples were run through a gel with two different dyes: one with β-mercaptoethanol (BME) and one without. With the information given by the small crystal conditions, more focused trays could be established that are based on the successful crystallization conditions from the blocks.

Six different condition sets were established from the data retrieved from the small trays, but due to limited resources such as certain PEG concentrations not being present within the lab, a lack of uncontaminated Jeffamine, and some salts not being present in the lab, only three large trays were able to be created. These trays varied in Polyethylene Glycol (PEG) concentrations, PEG molecular weights, buffer concentrations, and buffer pH, using the most common characteristics found in the successful small trays and expand upon them to better cover possible crystallization conditions. All of the trays expanded on the already established relationships to discover the most ideal set of conditions for crystallization.

On each large tray there were 24 wells for possible combinations to make conditions, for each condition, both D4A4 and ΔC8 were tested for crystallization. In each well, 750 mL of the condition was made and 2 μL of each SSB mutant was pipetted onto a glass cover along with a third drop containing only the condition. After each protein was on the plate, 2 μL of the condition for that well was pipetted into each protein sample. The edges of the plate were then coated with a sealing oil to create an airtight seal when the plate was flipped onto the well, sealing in the mixed protein samples and the conditions. The three trays rested for a few days in a 25^{o}C refrigerator before taking them to a microscope for analysis. A few crystals had been established, but the trays

were put to rest for another week before analyzing them again. At this time, more crystals had formed, and large ones were fished out and flash frozen to be sent for the possibility to establish a crystal structure. Unfortunately, none of the crystals sent in for analysis were able to yield a good crystal structure. Being able to set up the remaining three large trays that we were unable to make would provide more information on the crystallization ability of each SSB mutant. With this information, even better conditions could be made to make more established crystals for use in structure determination.

2.7 HDX

Both samples were run through intact MS experiments using an M-Class UPLC coupled to a Waters SYNAPT G2-Si ESI-Q-TOF. 0.2 µl of samples at 50 µM were loaded on a Waters Acquity UPLC Protein BEH C4 Column, using an automated sampler and held at 80°C. For this, buffer A is 0.1% formic acid in water and buffer B is 0.1% formic acid in acetonitrile. The chromatography solvent system was flowed at 100 µL/min and maintained for one and a half minutes at 10% buffer B, altered to 90% Buffer B over four minutes, maintained at 90% buffer B for one minute, and then re-equilibrated to 10% buffer B. All mass spectrometry data was acquired using positive ion mode with all samples collected in resolution mode. The desolvation temperature was set to 175°C.

The ΔC8 sample used in HDX was quenched using Q buffer made of 1.6 M GuHCl, 0.8% Formic Acid, pH 2.3 and flash frozen with three technical replicates. Once ready, the samples were thawed and immediately used in LC/MS using a Waters HDX manager as well as the SYNAPT used for intact MS. ΔC8 was digested using a Waters Enzymate BEH pepsin column and then

trapped on a Waters Acquity UPLC Protein BEH C18 column for before separating and eluting over a 3 to 40% (v/v) acetonitrile gradient for 7 minutes at 40 μl/min and 1 °C. The resulting data for ΔC8 was processed by a Waters Protein Lynx Global Serve and saved for later analysis with respect to a potential crystal structure.

CHAPTER 3

RESULTS

3.1 Small Scale Expression Test

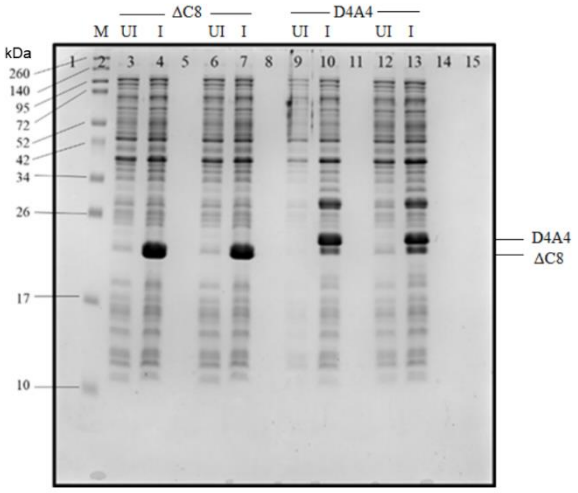

Figure 2. Small Scale Expression Test, 18% SDS-PAGE stained with Coomassie blue. Wells 1, 5, 8, 11, 14, and 15 were left blank. Well 2 contains a molecular ladder. Wells 3 and 6 contain uninduced SSB ΔC8, while wells 4 and 7 contain induced SSB ΔC8. Wells 9 and 12, like 3 and 6, contain uninduced SSB D4A4, while 10 and 13 have induced SSB D4A4.

We first tested the expression of SSB ΔC8 and D4A4 using a small-scale expression test (Figure 2). Comparison for the whole cell lysates from both induced and uninduced samples show expression of both proteins. For ΔC8, there is a strong band at approximately 20 kDa only in the induced lane. The result is similar for D4A4 with the induced band migrating slightly higher in the gel. This is expected as it has more residues and is therefore larger. D4A4, however, also has a band right underneath the dominant band suggesting it may be degraded in the cells. The interactions present in the wild type SSB may work in keeping some parts of the protein together through interactions with the acidic residues. By replacing these acidic residues with nonpolar ones, the integrity of the protein could be compromised, leading to the degradation. Other evidence

in support of this are later gels run from the histrap and SEC fractions where the two bands remain (Figures 4, 6, 8, and 10). Having confirmed expression small scale, we proceeded with large scale expression and subsequence purification.

3.2 Histrap and SEC

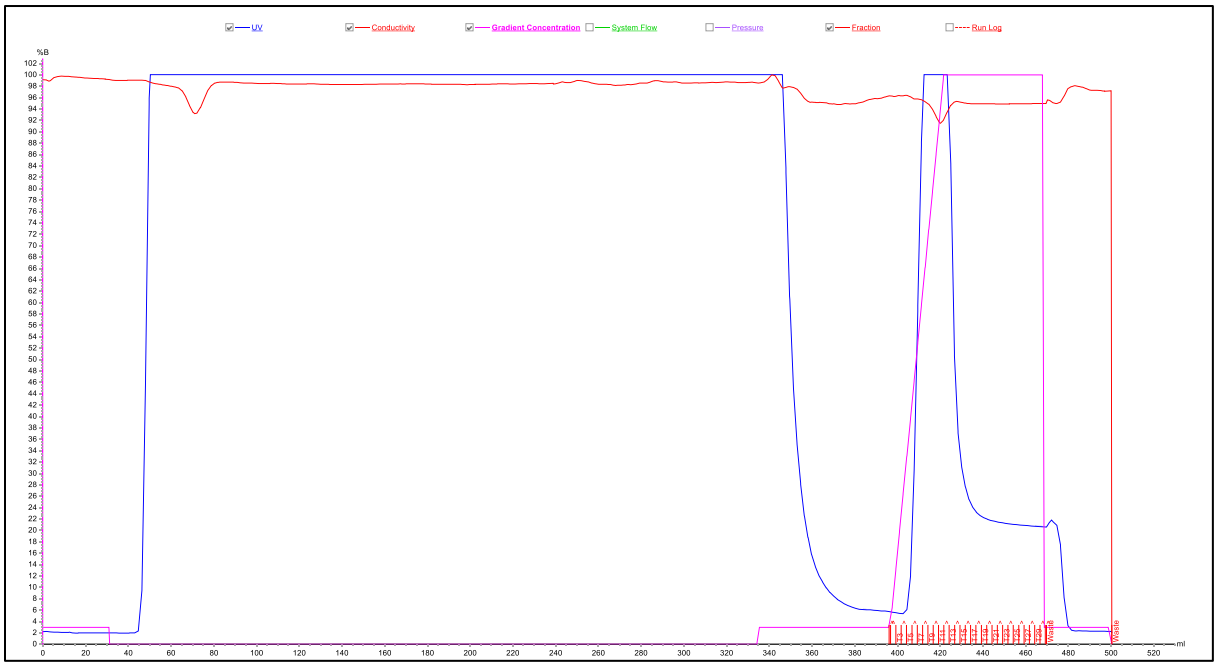

Figure 3. D4A4 Histrap Chromatogram. UV absorption at 280 nm is shown by the blue line, with conductivity shown with the red, and %B shown in pink. Fractions 1 through 30 are collected in the region marked by the lower red cells.

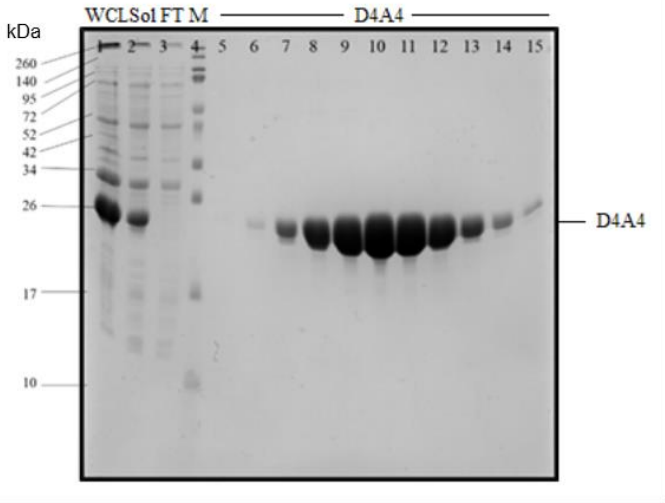

Figure 4. D4A4 Histrap, 18% SDS-PAGE stained with Coomassie blue. Wells 1, 2, 3, and 4 contain WCL, Sol, FT, and the molecular ladder, respectively. Wells 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 contain fractions 2, 5, 8, 11, 14, 16, 18, 21, 24, 27, and 30, respectively.

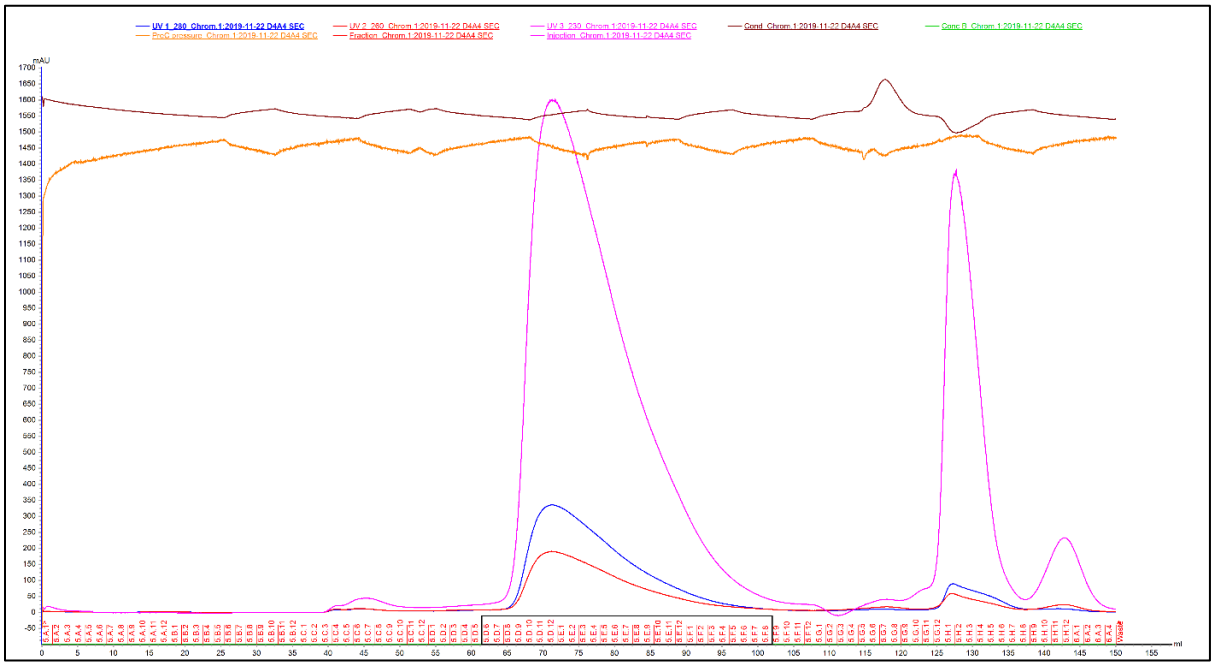

Figure 5. D4A4 SEC Chromatogram, 280, 260, and 230 nm in blue, red, and pink, respectively. Fractions taken are outlined in the black box (5D6 through 5F8, taking every other fraction). Run using column S200 and SEC buffer.

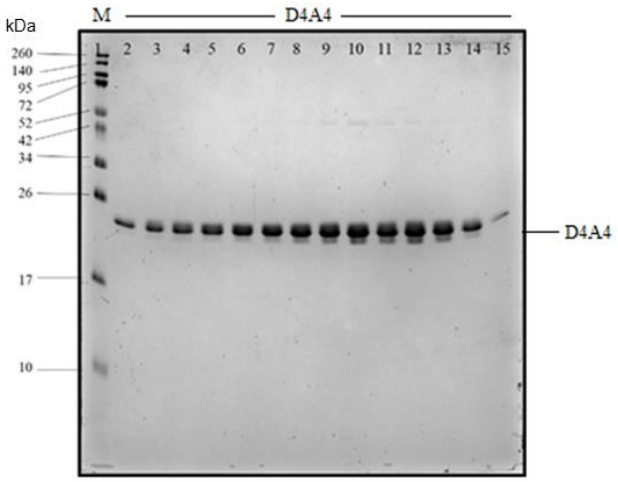

Figure 6. D4A4 SEC, 18% SDS-PAGE stained with Coomassie blue. Well 1 contains the molecular ladder, with the remaining wells containing, from 2 to 15, 5D6, 5D8, 5D10, 5D12, 5E2, 5E4, 5E6, 5E8, 5E10, 5E12, 5F2, 5F4, 5F6, and 5F8.

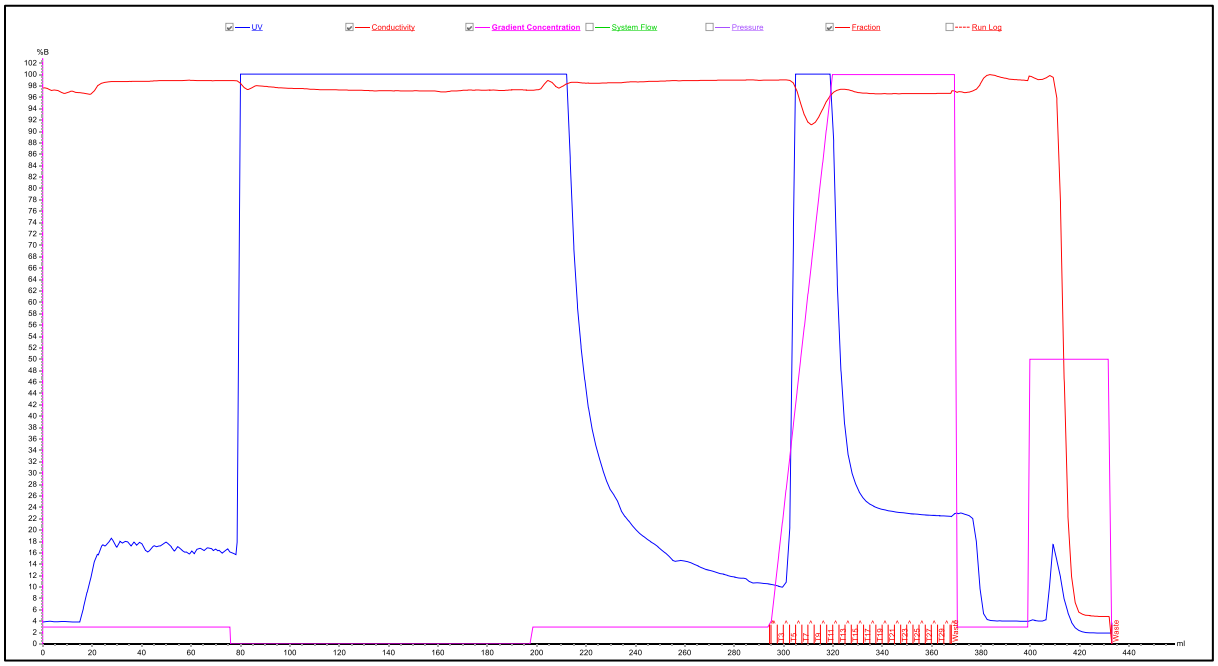

Figure 7. ΔC8 Histrap Chromatogram. UV absorption at 280 nm is shown by the blue line, with conductivity shown with the red, and %B shown in pink. Fractions 1 through 30 are collected in the region marked by the lower red cells.

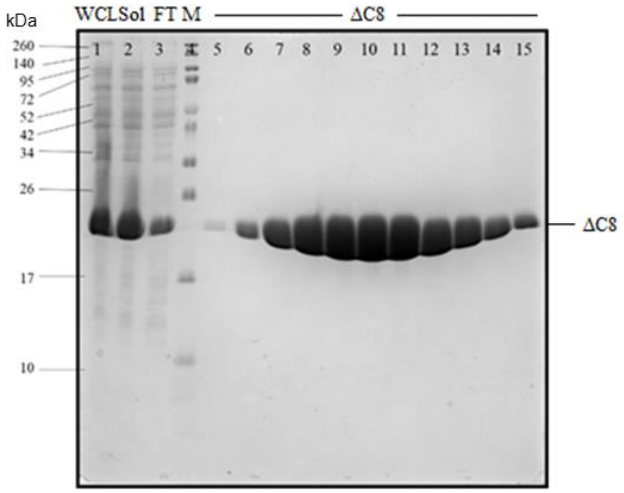

Figure 8. ΔC8 Histrap, 18% SDS-PAGE stained with Coomassie blue. Wells 1, 2, 3, and 4 contain WCL, Sol, FT, and the molecular ladder, respectively. Wells 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 contain fractions 2, 5, 8, 11, 14, 16, 18, 21, 24, 27, and 30, respectively.

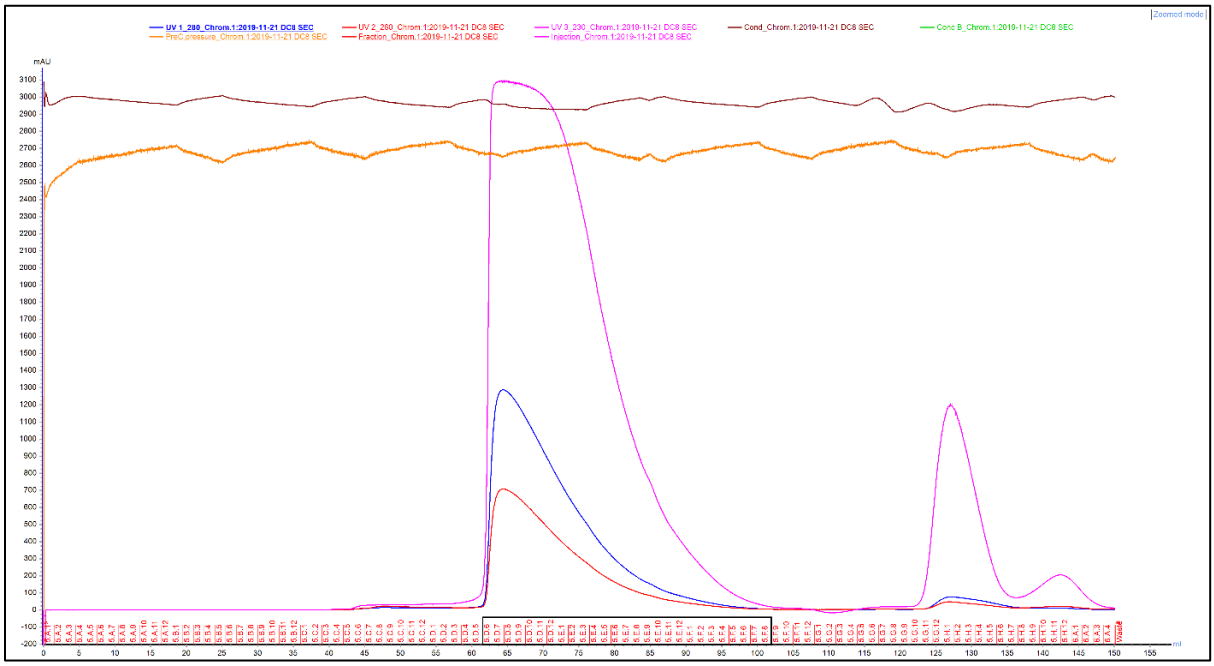

Figure 9. ΔC8 SEC Chromatogram, 280, 260, and 230 nm in blue, red, and pink, respectively. Fractions taken are outlined in the black box (5D6 through 5F8, taking every other fraction). Run using column S200 and SEC buffer.

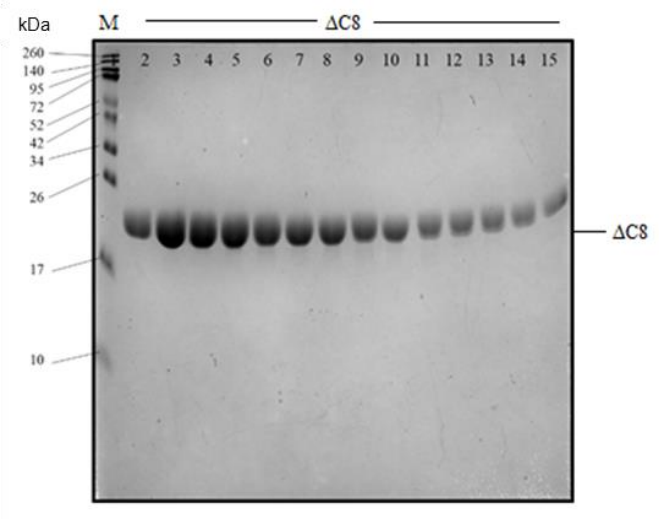

Figure 10. ΔC8 SEC, 18% SDS-PAGE stained with Coomassie blue. Well 1 contains the molecular ladder, with the remaining wells containing, from 2 to 15, 5D6, 5D8, 5D10, 5D12, 5E2, 5E4, 5E6, 5E8, 5E10, 5E12, 5F2, 5F4, 5F6, and 5F8.

Both constructs were purified with Ni-affinity chromatography (Figures 3 and 7) Followed by SEC (Figures 5 and 9). The overall yield of D4A4 was consistently lower than that of ΔC8 as seen in the SEC chromatograms. D4A4, on average, had a UV reading peaking at around 340 mAU, whereas ΔC8s reading was nearly four times as intense at around 1300 mAU (Figures 5 and 9) . This is not shown in the histrap chromatograms due to both mutants giving a higher reading than the machine is capable of interpreting, which is shown as a plateau on the chromatogram after fraction collection begins (Figures 3 and 7). The histrap gels for both mutants (Figures 4 and 8) are also in support of this, as in both of their gels showed protein expression in the WCL and Sol samples, but only in D4A4 was the protein shown to be completely lacking in the FT. These purifications were from the same volume of culture. The presence of protein in the FT would only occur if the protein lacked the histag, or if the column was completely saturated and could not hold any more protein, with the latter being true for ΔC8. But even with a much higher yield, ΔC8 was not immune to degradation, as seen in the SEC chromatogram.

SEC profiles for both mutants lacked a gaussian bell curve as they came off the column, instead being shown as a positively skewed curve. This is indicative of protein degradation, with a sharp increase in absorbance as the bulk of the protein comes off the column, and a slow decrease as smaller fragments of the protein are released. This shows that neither SSB mutant is completely stable and will degrade quickly if not stabilized by outside factors such as partner proteins or a storage solution like glycerol. This could also be a sign of a dynamic equilibrium between oligomeric states. This is not to say that partner proteins would keep it from degrading as quickly, but it is a theory that could be tested in the future. Hypothetically, it might make sense for SSB to degrade quickly when not stabilized due to the nature of its interactions. Working with ssDNA puts it into a critical situation where, if it were not controlled properly, could do more damage than harm. The presence of partner proteins could help it ensure high stability when working on ssDNA, and the lack of them could promote degradation when not in use to prevent any unneeded SSB activity.

3.3 SEC-MALS

Table 1. 20 μM D4A4, ΔC8 and 50 μM ΔC8 Molecular Weight with associated Uncertainty.

Sample	Observed MW (±kDa)	Theoretical MW (kDa)
20 μM SSB D4A4 300 mM NaCl	86.3 ±44.01	83.8
20 μM SSB ΔC8 300 mM NaCl	70.4 ±8.17	80.7
50 μM SSB ΔC8 300 mM NaCl	81.0 ±6.16	80.7

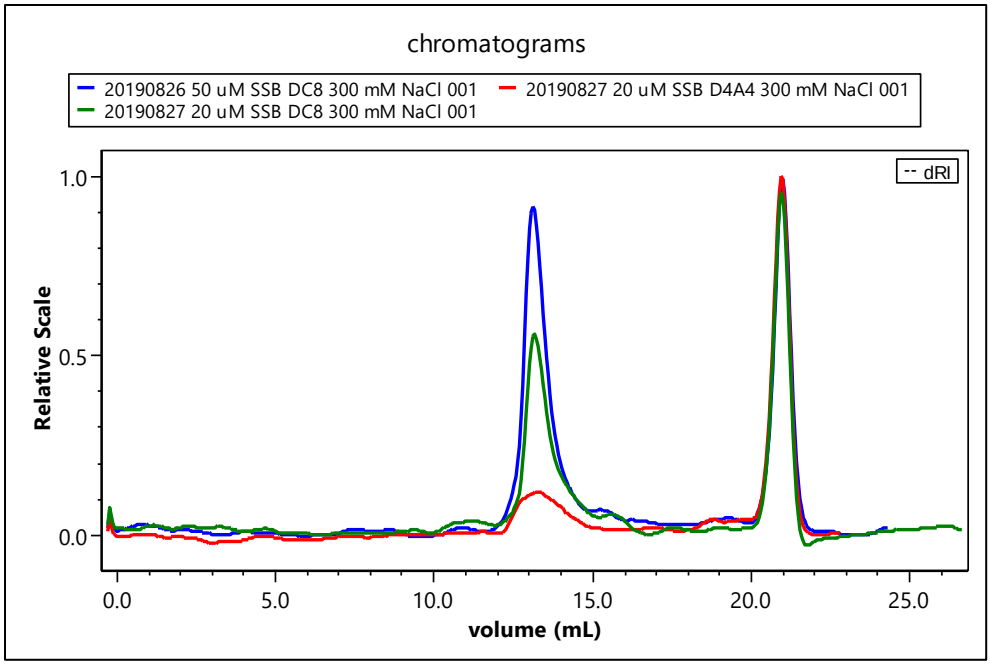

Figure 11. 20 μM D4A4, ΔC8 and 50 μM ΔC8 chromatograms run with S200 Increase 10/300 GL column and SEC-MALS buffer.

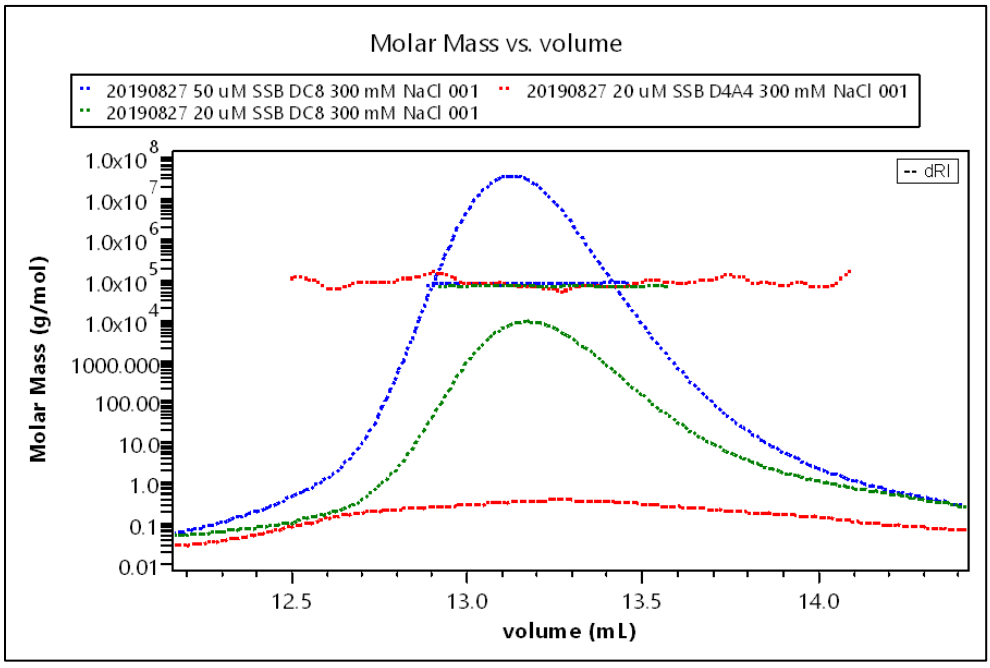

Figure 12. 20 μM D4A4, ΔC8 and 50 μM ΔC8 Molar Mass vs. Elution Volume

Table 2. 62.5 and 100 μM D4A4 and ΔC8 Molecular Weight with associated Uncertainty.

Sample	Observed MW (±kDa)	Theoretical MW (kDa)
62.5 μM SSB D4A4 300 mM NaCl	83.4 ±4.75	83.8
62.5 μM SSB ΔC8 300 mM NaCl	80.3 ±0.80	80.7
100 μM SSB D4A4 300 mM NaCl	248.2 ±236.04	83.8
100 μM SSB ΔC8 300 mM NaCl	81.7 ±0.98	80.7

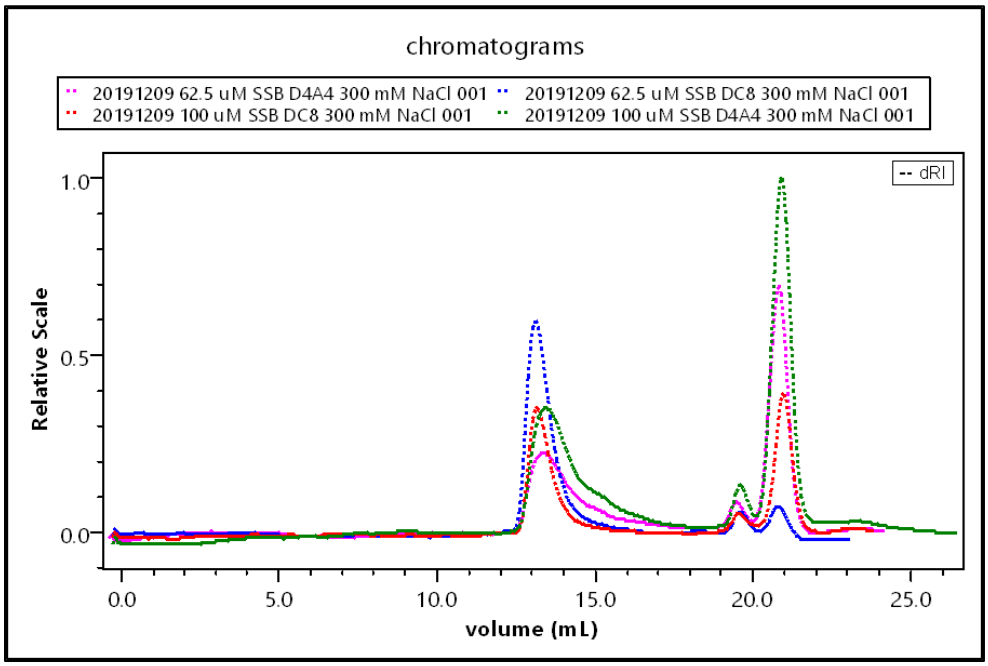

Figure 13. 62.5 and 100 μM D4A4 and ΔC8 chromatograms run with S200 Increase 10/300 GL column and SEC-MALS buffer.

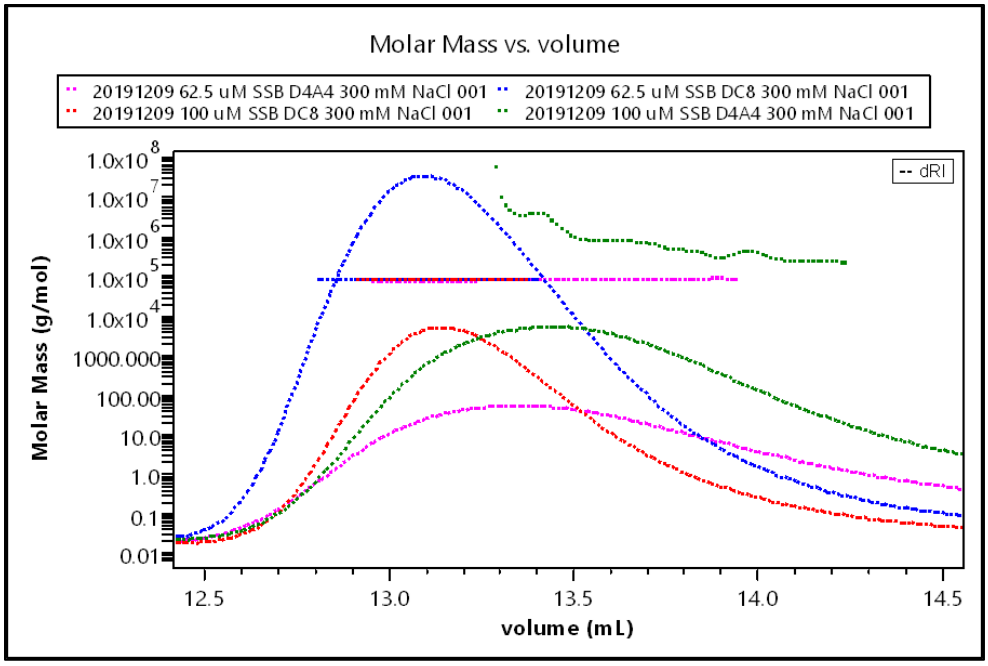

Figure 14. 62.5 and 100 µM D4A4 and ΔC8 Molar Mass vs. Elution Volume

Table 3. 62.5 µM D4A4, ΔC8 and 100 µM ΔC8 Molecular Weight with associated Uncertainty.

Sample	Observed MW (±kDa)	Theoretical MW (kDa)
62.5 µM SSB D4A4 300 mM NaCl	83.4 ±4.75	83.8
62.5 µM SSB ΔC8 300 mM NaCl	80.3 ±0.80	80.7
100 µM SSB ΔC8 300 mM NaCl	81.7 ±0.98	80.7

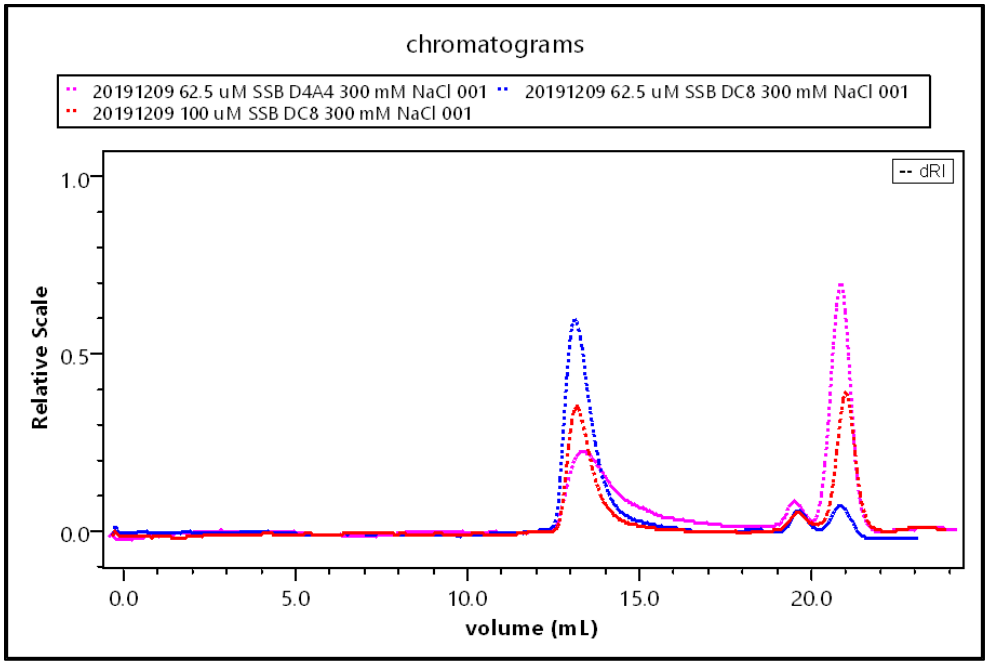

Figure 15. 62.5 µM D4A4, ΔC8 and 100 µM ΔC8 chromatograms run with S200 Increase 10/300 GL column and SEC-MALS buffer.

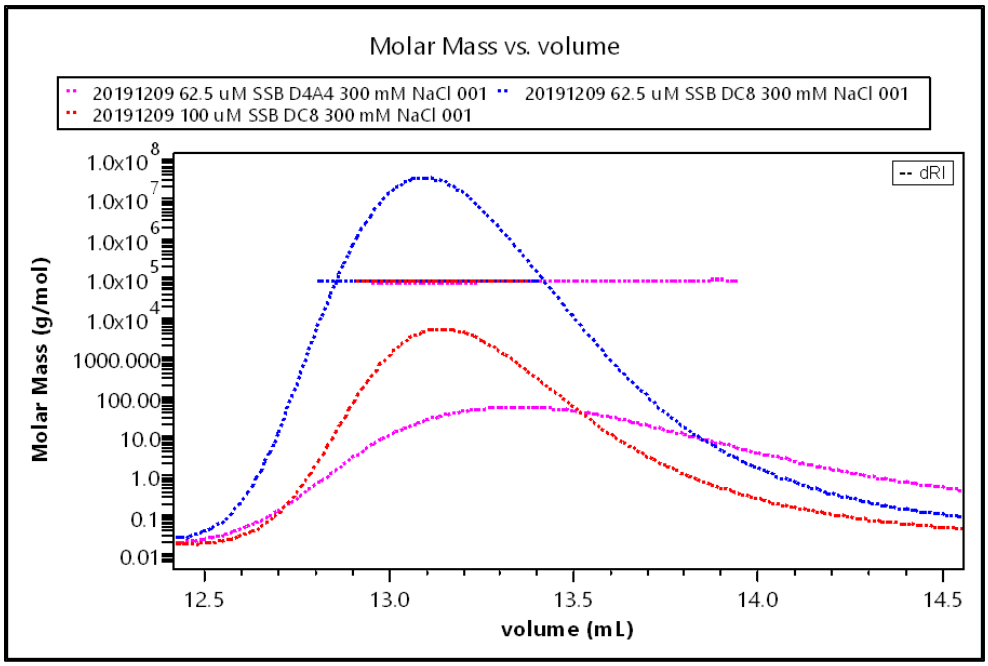

Figure 16. 62.5 µM D4A4, ΔC8 and 100 µM ΔC8 Molar Mass vs. Elution Volume

While SEC produces a pure protein sample, ensuring the sample is intact and stable while also checking if its molecular weight is accurate is important, and SEC-MALS is exactly the method of which to do this. The initial experiment was run with 20 μM samples of both SSB mutants, but due to low signal as a result of low protein concentrations, the uncertainties read 51% for D4A4 and 11.6% for ΔC8 as well as an inaccurate molecular weight for both samples with D4A4 showing 86.3 kDa and ΔC8 showing 70.4 kDa as seen in Table 1, a second sample of SSB ΔC8 was run at 50 μM. This trial was much more successful in attaining a lower level of uncertainty at 7.6% and a more accurate molecular weight of 81 kDa. As stated previously, the molecular weight for each of the mutants sit around 20 kDa, but the understanding that SSB exists as a tetramer lead to the acceptance of 81 kDa being appropriate. Although the numerical data showed an improvement, the data was still not at the desired level, as the uncertainty was still too high, and the molecular weights were still not in line with the theoretical. The uncertainty in relation to the molecular weight is exemplified in Figure 12, where the roughly horizontal lines running through the curves are representative of the uncertainty. Seeing how the increase in concentration improved the data for ΔC8, four more trials were performed with higher concentrations using both mutants.

Shown in Table 2 is the data for a 62.5 μM and 100 μM samples for both D4A4 and ΔC8. The higher concentrations proved to be fruitful as the molecular weights were more accurate for nearly every sample, and the uncertainty shared in this level of improvement. Unfortunately, the 100 μM sample for D4A4 did not resolve well, with a molecular weight nearly three times higher than what was expected as well as an uncertainty of 95.1%. In Figure 14, the horizontal lines are very flat, with the 100 μM D4A4 sample clearly being an outlier to this; because of this, a second

set of Figures and an additional Table was constructed with the exclusion of this outlier, shown as Table 3, and Figures 15 and 16. As shown in Table 3, the molecular weight was much more accurate, with all three values being much more appropriate to what would be expected from a tetramer of these proteins. The uncertainties are also much lower, with the highest of the three being at 5.7% for D4A4. The percentages for both ΔC8 samples were around 1%, with the 62.5 µM sample at 0.8%, and the 100µM sample at 1.2%. All of these can be seen in Figure 16 with the line being nearly completely flat, with some humps showing from D4A4. We conclude that both constructs are tetramers and that the non-gaussian peak likely results from both degradation and perhaps an equilibrium between different oligomers.

3.4 Intact MS and HDX

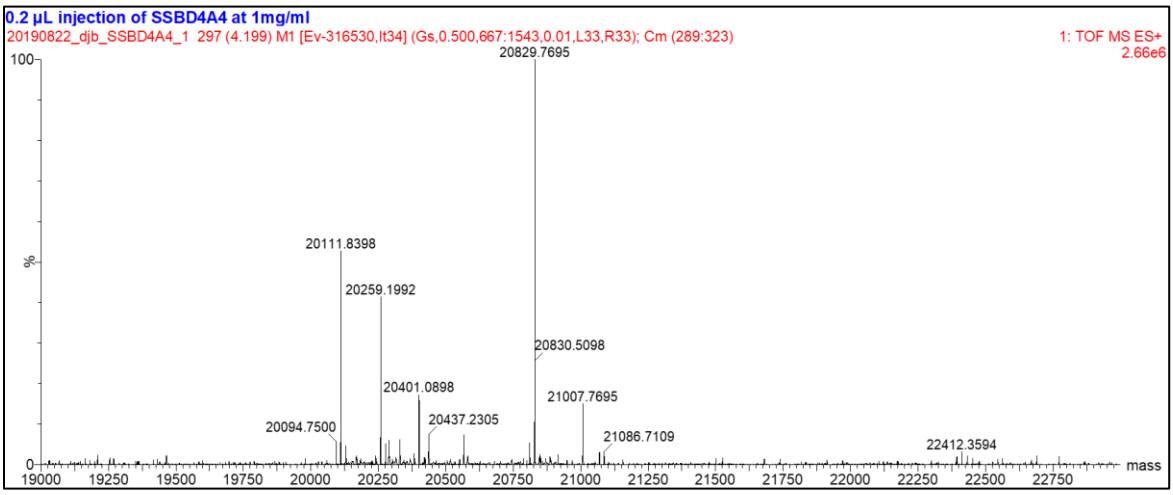

Figure 17. D4A4 Intact MS Data

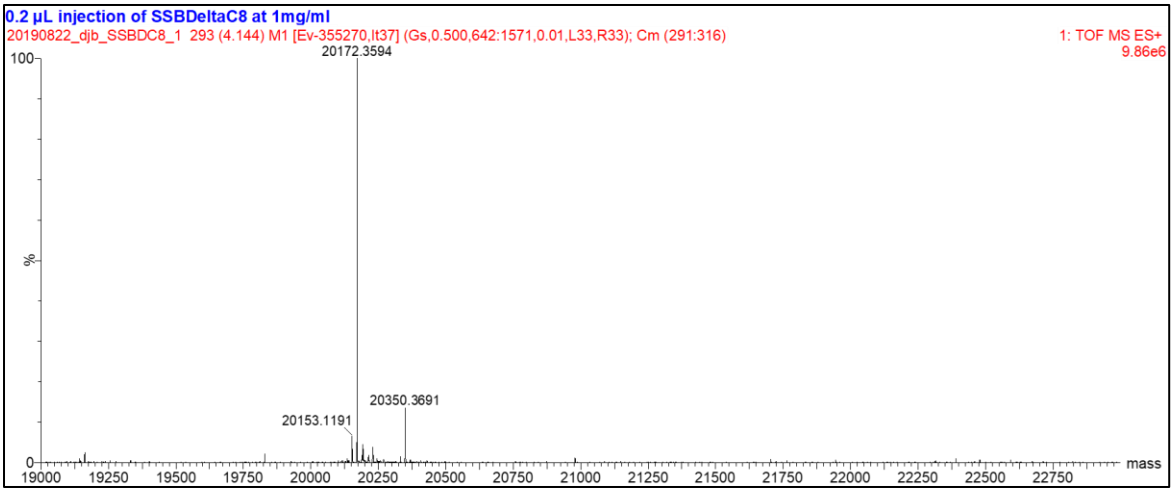

Figure 18. ΔC8 Intact MS Data

With the purity of the samples known, their weights were then tested one last time using intact MS. Seen in Figures 17 and 18, the highest peaks for both proteins were very close to the stated molecular weights of 20,962.28 Da and 20,173.33 Da for D4A4 and ΔC8, respectively. It can be seen in Figure 17 that there are many more smaller peaks in the data for D4A4 compared to ΔC8 in Figure 18, which could be due to the aforementioned possibility of D4A4 more readily degrading. Nonetheless, both mutants had accurate molecular weights aligning with what would be expected with D4A4 showing its data at 20,829.77 Da, and ΔC8 with a its trials data at 20,172.36 Da.

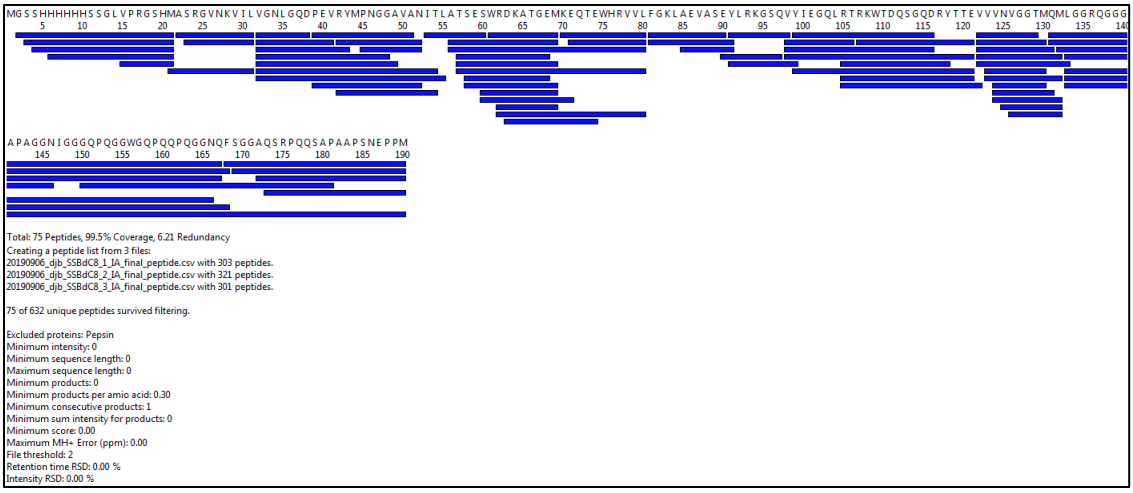

Figure 19. ΔC8 HDX Coverage Map, Coverage Percent = 99.5, Redundancy = 6.5

From a purified sample of ΔC8, an HDX coverage map was produced showing that the conditions this experiment was run at are feasible for a complete HDX experiment as there was nearly complete coverage of the protein. Because pepsin preferentially cleaves at the C-terminal ends of the aromatic amino acids phenylalanine, tyrosine, and tryptophan, as well as the C-terminal end of leucine, many of the more replicated peptides seen either begin with or end with one of these residues. The first one shown starting from residue 32, a valine residue present just after Leu31.

The next available amino acid of the four would be at residue 35 with leucine and then 43 with tyrosine. While there is no cleavage at Leu35 in any of the existing peptides, there is on existing cleavage incident at Ty+++r43. Another large peptide with high redundancy is seen starting at residue 56 with alanine, which is similarly right after a leucine residue. There is a tryptophan at residues 61 and 75, but no cleavage took place at these locations. A few of the largest peptides as well as some smaller ones after Leu55 showed cleavage after residue 80, another leucine.

The next peptide of interest is at Arg105, which like the previous peptides with high redundancy, starts right after a leucine. Dissimilar to the previous peptides however, there is not an overall common amino acid of the four that pepsin preferentially cleaves at the end of any of the peptides present here. There is a tyrosine at residue 118 that one of the peptides is cleaved at, but the other peptides cleave before after Asp116 or, more commonly, after Glu121. Under certain circumstances, pepsin will also cleave at a proline residue; however, there were not any noteworthy peptides cleaved at proline, even though ΔC8 does contain over ten proline residues, most commonly seen in the C-terminal domain.

3.5 Crystallization

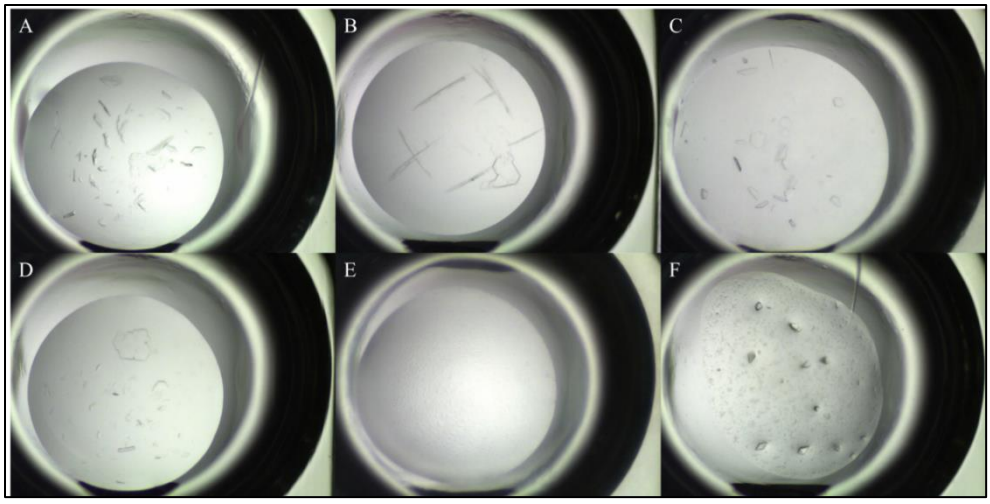

Figure 20. Small Tray SSB D4A4 Crystal Conditions. (A) Crystal B5 - 0.2 M Lithium sulfate monohydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene. (B) Index C11 - 1.0 M Ammonium sulfate, 0.1 M HEPES pH 7.0, 0.5% w/v Polyethylene glycol 8,000. (C) Index D3 - 0.1 M HEPES pH 7.0, 30% v/v Jeffamine ED-2001 pH 7.0. (D) Index F8 - 0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350. (E) PEG/Ion B2 - 0.2 M Potassium thiocyanate, 20% w/v Polyethylene glycol 3,350. (F) Salt E5 - 1.0 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 5.0.

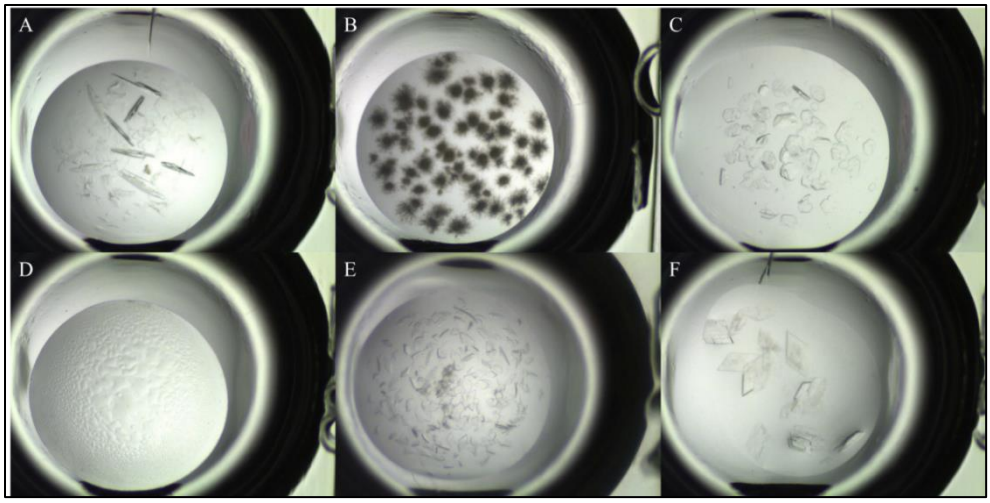

Figure 21. Small Tray SSB ΔC8 Crystal Conditions. (A) Crystal B5 - 0.2 M Lithium sulfate monohydrate, 0.1 M TRIS hydrochloride pH 8.5, 30% w/v Polyethylene. (B) Index C11 - 1.0 M Ammonium sulfate, 0.1 M HEPES pH 7.0, 0.5% w/v Polyethylene glycol 8,000. (C) Index D3 - 0.1 M HEPES pH 7.0, 30% v/v Jeffamine ED-2001 pH 7.0. (D) Index F8 - 0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350. (E) PEG/Ion B2 - 0.2 M Potassium thiocyanate, 20% w/v Polyethylene glycol 3,350. (F) Salt E5 - 1.0 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 5.0.

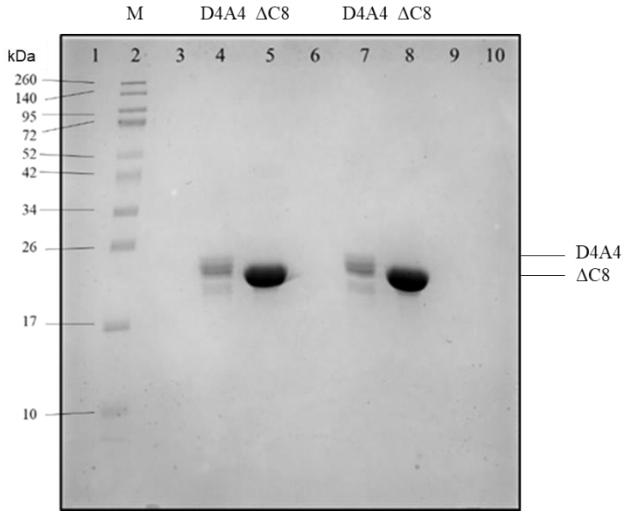

Figure 22. BME vs. non-BME SSB D4A4 and ΔC8, 18% SDS-PAGE stained with Coomassie blue. Wells 1, 3, 6, 9, and 10 were left empty. The molecular ladder was used in well 2. Wells 4 and 5 had BME dye with D4A4 and ΔC8 while wells 7 and 8 contained non-BME dye with D4A4 and ΔC8, respectively.

The purified ΔC8 and D4A4 proteins were submitted to crystallization trials. The majority of the crystal hits from the small trays were at a consistent set of conditions, but various conditions also produced successful crystals such as 0.1 M MES monohydrate buffer, PEG MME 550 and 2000, Jeffamine (which functions similarly to PEG), and many more. These conditions produced many different types of crystals ranging from long individual needles and urchin like needle clusters to prisms. Most crystals seen were needles like those seen in Figures 20-B and 21-A, with some notable plates and prisms such as the ones seen in Figures 20-D, 20-F, and 21-C. These larger crystals were the ones most commonly chosen to be sent off for diffraction analysis for both the small plates as well as the larger ones. From these samples, only the largest were harvested, with many of the chosen crystals being large prisms or plates.

Among these small tray samples, there were also many instances of aggregation and phase separation, as is seen in Figure 21-D. Many of the conditions failed to crystallize, but none more so than Salt Rx HT conditions, with only two successful crystal hits for D4A4. Both conditions were found to be very similar, as they both contained 1.0M sodium phosphate monobasic monohydrate, and potassium phosphate dibasic, with the only difference being the pH at 5.0 and 6.9 for the two hits. The most abundant crystals were found in the Index HT block, with 10 successful hits. Most small tray crystal hits occurred around 0.2 M salt concentrations, 0.1 M Tris pH 8.5 or 0.1 M HEPES pH 7.0 and 7.5. The PEG used was 3350 25% for nearly every condition, with a 20% being used in one, and different reagents such as Jeffamine being used in its place in others.

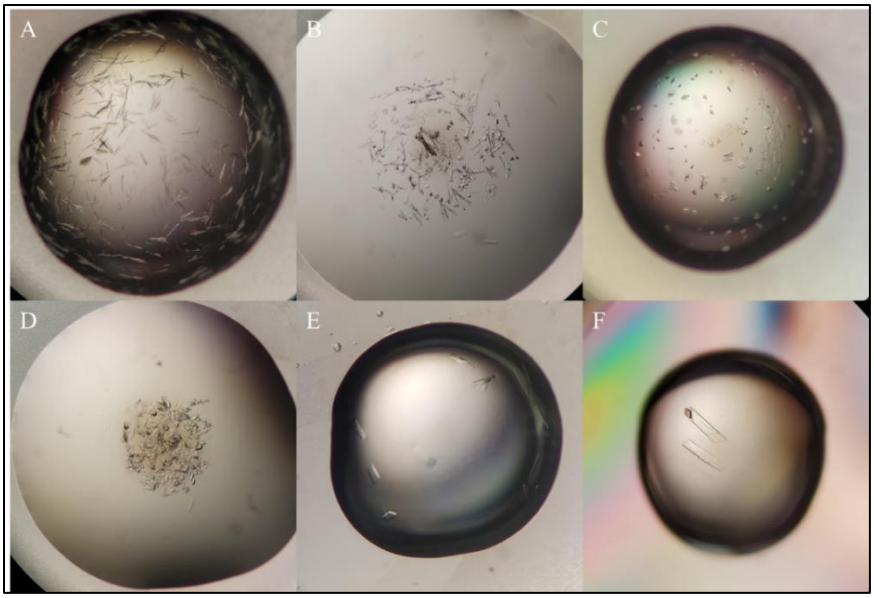

Figure 23. Large Tray SSB D4A4 Crystal Conditions. (A) Alpha A3 – 10% PEG 3350, 0.1 M HEPES pH 7. (B) Alpha B6 – 15% PEG 3350, 0.1 M HEPES pH 8.5. (C) Epsilon A6 – 15% PEG 3350, 0.1 M Tris pH 9. (D) Epsilon B4 – 20% PEG 3350, 0.1 M Tris pH 8. (E) Gamma A3 – 15% PEG 1000, 0.1 M HEPES pH 7. (F) Gamma A5 – 15% PEG 1000, 0.1 M HEPES pH 8.

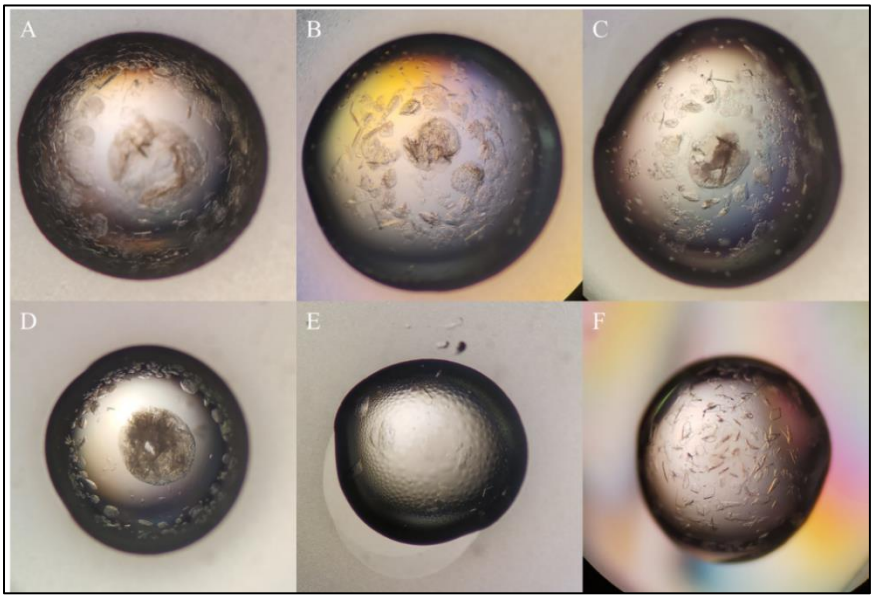

Figure 24. Large Tray SSB ΔC8 Crystal Conditions. (A) Alpha A3 – 10% PEG 3350, 0.1 M HEPES pH 7. (B) Alpha B6 – 15% PEG 3350, 0.1 M HEPES pH 8.5. (C) Epsilon A6 – 15% PEG 3350, 0.1 M Tris pH 9. (D) Epsilon B4 – 20% PEG 3350, 0.1 M Tris pH 8. (E) Gamma A3 – 15% PEG 1000, 0.1 M HEPES pH 7. (F) Gamma A5 – 15% PEG 1000, 0.1 M HEPES pH 8.

The data obtained from these trials, as well as the other crystal condition blocks, were used to put together six large trays that took advantage of the most prevalent conditions shown to be successful for crystallization. These six trays were named alpha, beta, delta, epsilon, gamma, and eta; each tray tested the most common successful conditions while experimenting with the upper and lower limits of those conditions across 24 different wells. Due to experimental limitations stated earlier, only three of these trays were able to be tested: alpha, epsilon, and gamma. Alpha ran PEG 3350 at 10, 15, 20, and 25% against 0.1 M HEPES at pH 6, 6.5, 7, 7.5, 8, and 8.5. Crystals were most commonly found at 10 and 15% PEG 3350, with slightly less success at 20%. It seemed that the crystals became less abundant and smaller as the pH and PEG percentage increased. Nearly all crystals here represent needles, with some exceptions occurring in the form of small plates and prisms as both the pH and PEG percentage rose together.

Epsilon also ran PEG 3350, but at 15, 20, 25, and 30%, which was run against 0.1 M Tris buffer at pH 6.5, 7, 7.5, 8, 8.5, and 9. Similar to alpha, the most abundant hits were at the lower PEG 3350 percentages of 15 and 20%, but with a few additional hits at the higher two as well. Often seen in these conditions was the presence of droplets and aggregates alongside crystals, showing some phase separation and protein crashing. It also seemed to appear that as the pH increased, crystals became more distinct and shifted from adopting a needlelike structure to a more prism like one, but both crystals were usually small.

Lastly, gamma, like alpha, ran 0.1 M HEPES at the same six pH's; however, it did not use the same PEG as the first two trays, instead using PEG 1000 at 15, 20, 25, and 30%. Most of the successful hits occurred at 15%, with slightly fewer found at 20% PEG 1000; 25 and 30% saw no successful crystals forming. Each of the pH's saw success at the two PEG percentages, except for

8.5 at 20% PEG 1000. While this tray saw the fewest number of hits, the crystals grown from here were among the largest. The average shape of these crystals was also different compared to the other trays, where the norm for those seemed to be needles, the norm here seemed to be plates with only a few needles. In a few of the wells, crystal formation was also accompanied with the presence of droplets and aggregation. As pH increased at 15% PEG, there seemed to be a decrease in the overall number of crystals, but the crystals that did form were on average more plate-like. As for 20% PEG, needles were nearly nonexistent, showing almost exclusively plates and some very minute prisms. Interestingly, at 20% ΔC8 seemed to crystallize more often than D4A4, which is a stark contrast to how every condition has proceeded until now.

On average D4A4 would crystallize much more often than ΔC8, who normally formed aggregates and phase separations. The number of hits for alpha and epsilon were nearly identical, with alpha having 34 out of a possible 48 and epsilon having 32. While gamma only had 18 hits, it seemed to yield the best crystals, especially for ΔC8 which, as mentioned earlier, had great difficulty crystallizing. Unfortunately, cross both the small and large trays none of the crystals sent for analysis yielded a crystal structure.

The three trays executed provide useful information for refining the conditions but using the other three (beta, delta, and eta) would be invaluable to this process. Beta tested 0.1 M HEPES at pH 6.5, 7, 7.5, and 8 against 15, 20, 25, 30, 35, and 40% Jeffamine. Delta ran 15, 20, 25, and 30% PEG 1500 with 0.1 M BICINE at pH 6.5, 7, 7.5, 8, 8.5, and 9. Perhaps most unique of the six was eta, which unlike the other five, did not test a buffer of various pH's against a crystallization assistant like PEG, instead testing salt concentration against PEG percentage. Eta ran 0.2, 0.5, 0.7, and 1 M ammonium sulfate against 15, 20, 25, 30, 35, and 40% PEG 3350. This type of tray would

provide some insight with salts and could be run with NaCl as well as the other successful salt crystallization conditions. For future crystallization experiments, different reagents will need to be used to cover as many conditions as possible to give the best chance at determining the crystal structure. Because these were the first large trays to be run with the SSB mutants, many more revisions and refinement processes are needed to find the most consistent and effective conditions.

CHAPTER 4
DISCUSSION

4.1 Future Work

Additional SSB constructs will be useful for comparison moving forward. In particular, SSB Core and WT SSB. It will also be interesting to purify the SSB binding partner RecG and analyze the complex.

SSB D4A4 and ΔC8 have much to provide in uncovering the details about SSB, but the process to retrieve that information is littered with challenges. While neither mutant was able to provide a usable crystal structure, the information uncovered by these successful trials broke down many barriers in the way of learning the details of D4A4 and ΔC8. The expression and purification conditions were determined, the molecular weight of the monomer and the existence of the tetramer in native conditions was confirmed, HDX conditions for near complete coverage were defined, and crystallization conditions were screened, providing many optimistic hits.

This data, if expanded upon and applied alongside other mutants of SSB, could reveal many hidden attributes about wild type SSB, potentially assigning structure and function to the entire C-terminal domain, allowing us to understand how it interacts with the N-terminal domain, ssDNA, and other binding partners of the SSB interactome. This information may lead to applications for eukaryotic equivalents of SSB and aid in treating potential cancers that arise from them.

REFERENCES

1. Tamaru, H. Confining euchromatin/heterochromatin territory: jumonji crosses the line. *Genes Dev.* **24**, 1465–1478 (2010).
2. Dame, R. T. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular Microbiology* (2005) doi:10.1111/j.1365-2958.2005.04598.x.
3. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251–260 (1997).
4. Li, G. & Zhu, P. Structure and organization of chromatin fiber in the nucleus. *FEBS Lett.* **589**, 2893–2904 (2015).
5. Ishikawa, F. & Naito, T. Why do we have linear chromosomes? A matter of Adam and Eve. *Mutat. Res. - DNA Repair* (1999) doi:10.1016/S0921-8777(99)00017-8.
6. Henneman, B., van Emmerik, C., van Ingen, H. & Dame, R. T. Structure and function of archaeal histones. *PLoS Genetics* (2018) doi:10.1371/journal.pgen.1007582.
7. Thanbichler, M., Wang, S. C. & Shapiro, L. The bacterial nucleoid: A highly organized and dynamic structure. *Journal of Cellular Biochemistry* (2005) doi:10.1002/jcb.20519.
8. Griswold, A. Genome Packaging in Prokaryotes : the Circular Chromosome of E coli. *Nat. Educ.* (2008).
9. Champoux, J. J. DNA Topoisomerases: Structure, Function, and Mechanism. *Annu. Rev. Biochem.* (2001) doi:10.1146/annurev.biochem.70.1.369.
10. Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C. & de la Cruz, F. Mobility of Plasmids. *Microbiol. Mol. Biol. Rev.* (2010) doi:10.1128/mmbr.00020-10.
11. Neidle, S. *Principles of Nucleic Acid Structure*. *Principles of Nucleic Acid Structure* (2008). doi:10.1016/B978-0-12-369507-9.X5001-8.
12. Rios, A. C. & Tor, Y. On the origin of the canonical nucleobases: An assessment of selection pressures across chemical and early biological evolution. *Israel Journal of Chemistry* (2013) doi:10.1002/ijch.201300009.
13. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* (1953) doi:10.1038/171737a0.
14. Meselson, M. & Stahl, F. W. The replication of DNA in Escherichia coli. *Proc. Natl. Acad. Sci.* (1958) doi:10.1073/pnas.44.7.671.
15. Lilly, J. & Camps, M. Mechanisms of Theta Plasmid Replication. *Microbiol. Spectr.* (2015) doi:10.1128/microbiolspec.plas-0029-2014.

16. Mitchison, T. J. & Salmon, E. D. Mitosis: A history of division. *Nature Cell Biology* (2001) doi:10.1038/35050656.

17. Lilley, D. M. J. Structures of helical junctions in nucleic acids. *Quarterly Reviews of Biophysics* (2000) doi:10.1017/S0033583500003590.

18. Grelon, M. Mécanismes de la recombinaison méiotique. *Comptes Rendus - Biol.* (2016) doi:10.1016/j.crvi.2016.04.003.

19. Macaisne, N., Kessler, Z. & Yanowitz, J. L. Meiotic double-strand break proteins influence repair pathway utilization. *Genetics* (2018) doi:10.1534/genetics.118.301402.

20. Mehta, A. & Haber, J. E. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* (2014) doi:10.1101/cshperspect.a016428.

21. Liang, L. *et al.* Modulation of DNA end joining by nuclear proteins. *J. Biol. Chem.* (2005) doi:10.1074/jbc.M503776200.

22. Cannan, W. J. & Pederson, D. S. Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. *Journal of Cellular Physiology* (2016) doi:10.1002/jcp.25048.

23. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* (2005) doi:10.1038/nature03443.

24. Ashton, N. W., Bolderson, E., Cubeddu, L., O'Byrne, K. J. & Richard, D. J. Human single-stranded DNA binding proteins are essential for maintaining genomic stability. *BMC Molecular Biology* (2013) doi:10.1186/1471-2199-14-9.

25. Bianco, P. R. The tale of SSB. *Progress in Biophysics and Molecular Biology* (2017) doi:10.1016/j.pbiomolbio.2016.11.001.

26. Huang, C. Y. Crystal structure of SSB complexed with inhibitor myricetin. *Biochem. Biophys. Res. Commun.* (2018) doi:10.1016/j.bbrc.2018.08.188.

27. Matsumoto, T. *et al.* Roles of Functional Loops and the CD-Terminal Segment of a Single-Stranded DNA Binding Protein Elucidated by X-Ray Structure Analysis. *J. Biochem.* **127**, 329–335 (2000).

28. Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. & Keck, J. L. SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology* (2008) doi:10.1080/10409230802341296.

29. Tan, H. Y. *et al.* The intrinsically disordered linker of E. coli SSB is critical for the release from single-stranded DNA. *Protein Sci.* (2017) doi:10.1002/pro.3115.

30. Marceau, A. H. *et al.* Structure of the SSB-DNA polymerase III interface and its role in DNA replication. *EMBO J.* (2011) doi:10.1038/emboj.2011.305.

31. Maffeo, C. & Aksimentiev, A. Molecular mechanism of DNA association with single-stranded DNA binding protein. *Nucleic Acids Res.* (2017) doi:10.1093/nar/gkx917.

32. Antony, E. & Lohman, T. M. Dynamics of E. coli single stranded DNA binding (SSB) protein-DNA complexes. *Seminars in Cell and Developmental Biology* (2019) doi:10.1016/j.semcdb.2018.03.017.

33. Caldwell, R. B. & Lattemann, C. T. Simple and Reliable Method to Precipitate Proteins from Bacterial Culture Supernatant. *Appl. Environ. Microbiol.* (2004) doi:10.1128/AEM.70.1.610-612.2004.

34. Hengen, P. N. Purification of His-Tag fusion proteins from Escherichia coli. *Trends Biochem. Sci.* (1995) doi:10.1016/S0968-0004(00)89045-3.

35. Hochuli, E., Bannwarth, W., Dobeli, H., Gentzi, R. & Stuber, D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* (1988) doi:10.1038/nbt1188-1321.

36. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* (1975) doi:10.1038/258598a0.

37. Paul-Dauphin, S. *et al.* Probing size exclusion mechanisms of complex hydrocarbon mixtures: The effect of altering eluent compositions. *Energy and Fuels* (2007) doi:10.1021/ef700410e.

38. Brooks, D. E., Haynes, C. A., Hritcu, D., Steels, B. M. & Müller, W. Size exclusion chromatography does not require pores. *Proc. Natl. Acad. Sci. U. S. A.* (2000) doi:10.1073/pnas.120129097.

39. Kumar, P. & Mina, U. *Fundamentals and Techniques of Biophysics and Molecular Biology*. *Pathfinder* (2016). doi:10.1016/B978-0-12-723140-2.50009-6.

40. Tarazona, M. P. & Saiz, E. Combination of SEC/MALS experimental procedures and theoretical analysis for studying the solution properties of macromolecules. *J. Biochem. Biophys. Methods* (2003) doi:10.1016/S0165-022X(03)00075-7.

41. Striegel, A. M. & Sinha, P. Absolute molar mass determination in mixed solvents. 1. Solving for the SEC/MALS/DRI “trivial” case. *Anal. Chim. Acta* (2019) doi:10.1016/j.aca.2018.11.051.

42. Patel, B. A. *et al.* Multi-angle light scattering as a process analytical technology measuring real-time molecular weight for downstream process control. *MAbs* (2018) doi:10.1080/19420862.2018.1505178.

43. Zimm, B. H. & Stockmayer, W. H. The dimensions of chain molecules containing branches and rings. *J. Chem. Phys.* (1949) doi:10.1063/1.1747157.

44. Podzimek, S. The use of GPC coupled with a multiangle laser light scattering photometer for the characterization of polymers. On the determination of molecular weight, size and branching. *J. Appl. Polym. Sci.* (1994) doi:10.1002/app.1994.070540110.

45. Dale, G. E., Oefner, C. & D'Arcy, A. The protein as a variable in protein crystallization. *J. Struct. Biol.* (2003) doi:10.1016/S1047-8477(03)00041-8.

46. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallographica Section F:Structural Biology Communications* (2014) doi:10.1107/S2053230X13033141.

47. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* (2019) doi:10.1038/s41592-019-0459-y.

48. Engen, J. R. Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Anal. Chem.* **81**, 7870–7875 (2009).

49. Balasubramaniam, D. & Komives, E. A. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2013) doi:10.1016/j.bbapap.2012.10.009.

50. Molnar, I. & Horvath, C. Reverse phase chromatography of polar biological substances: separation of catechol compounds by high performance liquid chromatography. *Clin. Chem.* (1976) doi:10.1093/clinchem/22.9.1497.

51. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* (1989) doi:10.1126/science.2675315.

52. Collins, D. C. & Lee, M. L. Electrospray ionization gas-phase electrophoresis under ambient conditions and it's potential or high-speed separations. *Fresenius. J. Anal. Chem.* (2001) doi:10.1007/s002160000636.

53. Wolff, M. M. & Stephens, W. E. A pulsed mass spectrometer with time dispersion. *Rev. Sci. Instrum.* (1953) doi:10.1063/1.1770801.

BIOGRAPHICAL SKETCH

Matthew Franciskovich was born in Chicago, Illinois and raised in the Houston, Texas area. He graduated from Spring High School in 2014 and was accepted into the Baylor College of Medicine summer internship program that summer. From 2014 until 2018, he attended Baylor University under the Baylor Interdisciplinary Core honours college, receiving his bachelor's degree in biochemistry with a minor in educational psychology. During his four years in undergrad, he was also accepted into the leadership program. He was accepted into The University of Texas at Dallas in 2018 for a master's degree in chemistry where he worked in Dr. Sheena D'Arcy's laboratory focusing on the purification and characterization of single stranded binding protein. He also briefly worked on the characterization and analysis of histone H2A-H2B dimers and H3H4 tetramers. He completed his master's degree in Summer 2020, after which he will attend Sam Houston State University for a second master's degree in health.

CURRICULUM VITAE

Matthew Franciskovich

Highlights

- Experienced in a variety of teaching skills and communication methods.
- Proficient understanding of biochemical, biomedical, and biophysical techniques as well as those used in structural biology.
- Established understanding of Microsoft office, Adobe, UNICORN, ASTRA, PyMOL, ChemDraw, ImageLab, and Mendeley.

Organizations and Awards

National Honour Society
- 2013 – Current

Baylor College of Medicine Saturday Morning Science Student
- 2014

Baylor University Leadership Program
- 2014 – 2018

Fellowship Team Assistant Director

Baylor Undergraduate Research in Science and Technology (BURST)
- 2014 - 2018

American Chemical Society
- 2017 – Current

Executive Team Treasurer

Baylor Animal Conservation Society
- 2017 – 2018

Executive Team Treasurer

Baylor Interdisciplinary Core Honour's College Graduate
- 2018

The Society for Collegiate Leadership & Achievement
- 2019 – Current

Education

Baylor University – Waco, TX

- BS in Biochemistry, Minor in Educational Psychology
- Graduated 05/2018 under the honors college

The University of Texas at Dallas – Dallas, TX

- Prospective master's degree in Summer of 2020

Professional Experience

The University of Texas at Dallas

- 3/2019 – 5/2020

Researcher in D'arcy Group – Dallas, TX

- Worked in protein purification using highly technical laboratory methods with sophisticated instrumentation.

The University of Texas at Dallas

- 8/2018 – 05/2019

Chemistry TA and Tutor – Dallas, TX

- Engaged students in laboratory curriculum and guided their understanding in a meaningful and productive way.

Baylor University Bookstore

- 10/2015 – 04/2018

Sales Associate – Waco, TX

- Substantially improved work environment by organizing and teaching about large supplies of electronics.

Something Special Women's Shelter

- 06/2017 – 08/2017

Store Volunteer – Shenandoah, TX

- Accelerated uptake of donations through effective teamwork.

CHI St. Lukes Medical Center

- 06/2017 – 08/2017

ER Volunteer – The Woodlands, TX

- Recorded patient vital signs.
- Reported illnesses in patients.
- Aided in IV fluid preparation.
- 06/2016 – 08/2016

ICU Volunteer – The Woodlands, TX

- Worked alongside nurses to administer medication to patients.
- Provided high level care to patients.
- Actively prepared and organized medical records.

Baylor College of Medicine

- 05/2014 – 08/2014

Project Intern – Houston, TX

- Created high quality 3D models of diseased brain proteins.
- Worked alongside medical researchers to develop treatments for Huntington's disease.