A Thesis

Entitled

A Study of DNA Replication and Repair Proteins from Bacteriophage T4 and a Related Phage

by

Anne B. Senger

Submitted as partial fulfillment of the requirements for

the Master of Science in Chemistry

Advisor: Timothy C. Mueser, Ph.D.

Graduate School

The University of Toledo

December 2004
Copyright © 2004

This document is copyright material. Under copyright law, no parts of this document may be reproduced without the expression and permission of the author.
An Abstract of

A Study of DNA Replication and Repair Proteins from Bacteriophage T4 and a Related Phage

Anne B. Senger

Submitted as partial fulfillment of the requirements for
the Master of Science in Chemistry

The University of Toledo
December 2004

In the Mueser laboratory, we study how DNA replication and repair proteins recognize DNA in a structure-specific manner. Bacteriophage T4 is used as a model system to study DNA replication as it encodes all ten proteins required for DNA replication. Much is known about how the individual proteins function in replication but not much is known about the structural aspects of the protein-protein or protein-DNA interactions at the replication fork. The goal of our research is to study how these replication proteins interact with each other and
with DNA. We work towards achieving this goal by crystallizing the protein-protein and protein-DNA complexes and then solving their structures, using macromolecular crystallography techniques. We then use the structural information gathered to analyze the interactions.

The overall goal of this master’s thesis project was to learn many of the techniques involved in protein chemistry and protein crystallization. My research was tailored to protein expression, purification and crystallization so I could learn an array of techniques and become familiar with various pieces of instrumentation. I wanted to be able to use this knowledge in future research positions.

My work was focused on two of the replication proteins from Bacteriophage T4: T4 gene 59 helicase assembly protein and T4 gene 32 single-stranded binding protein. These two proteins interact in the absence of DNA and form a complex at the replication fork. I was responsible for expressing mutated and truncated forms of the native proteins on a large scale and developing purification protocols in order to prepare pure protein for crystal screening. After my research with the T4 helicase assembly protein began, I also started working on a similar helicase assembly protein from a related system – bacteriophage KVP40 59 protein. I was also responsible for developing a purification protocol for single-stranded DNA substrates that were used to prepare forked substrates for the crystal screening.
Acknowledgements

First and foremost, I would like to thank my boss, Dr. Timothy Mueser. I feel incredibly lucky and honored to have been able to work for him over the past two years. I have learned so much while working in his laboratory and have made many friends along the way. He gave me the opportunity to do amazing research and for that, I am truly thankful. I would also like to thank my parents, sisters, grandparents and all of my extended family. They have always been incredibly supportive of me and I know they are very proud of my accomplishments. I am eternally grateful to my boyfriend, Pete, and all of my girlfriends across the country for their love, support and comic relief! Thank you to Dr. Ronald Viola and Dr. Max Funk for sitting on my thesis committee. I will be forever indebted to the Department of Chemistry faculty and staff for their knowledge and assistance over the past two and a half years. I would also like to thank the University of Toledo and the National Science Foundation for funding. Thank you to Dr. Nancy Nossal at NIH for materials and collaborations. Last but certainly not least, I would like to thank the people I have spent an enormous amount of time with in the last two years - all of my labmates, past and present: Aude, Brandon, Jen, Juliette, Kelly, Laurence, Pooja, Steve, Vinu and Wilawan. I have never had the opportunity to work with a better group of people. I wish you all the best of luck in your endeavors. Wherever life may take you, I know you will all be very successful.
# Table of Contents

Abstract ........................................................................................................................ iii

Acknowledgements ........................................................................................................ v

Table of Contents .......................................................................................................... vi

List of Figures ................................................................................................................ ix

List of Tables .................................................................................................................. xiii

1 Chapter 1 ..................................................................................................................... 1

Introduction .................................................................................................................... 1

1.1 DNA Replication ................................................................................................... 1

1.2 Research summary ................................................................................................. 4

2 Chapter 2 .................................................................................................................... 7

Methodology .................................................................................................................. 7

2.1 Protein expression ................................................................................................. 7

2.2 Protein purification ............................................................................................... 9

2.3 Nuclease activity .................................................................................................. 13

2.4 Protein crystallization .......................................................................................... 13

2.5 Protein crystal manipulation ................................................................................ 18

2.6 X-ray diffraction data collection ......................................................................... 20

2.7 DNA purification ................................................................................................... 23

2.8 DNA annealing ...................................................................................................... 26

3 Chapter 3 .................................................................................................................... 27

DNA Purification .......................................................................................................... 27
5.3 (32 minus A) Experiments ................................................................................ 80

5.3.1 Protein expression ..................................................................................... 80

5.3.2 Protein purification ................................................................................... 81

6 Chapter 6 ........................................................................................................ 89

KVP40 59 protein ................................................................................................ 89

6.1 Background .................................................................................................. 89

6.2 Experiments .................................................................................................. 91

6.2.1 Protein expression ..................................................................................... 92

6.2.2 Protein purification ................................................................................... 92

6.2.3 Nuclease activity ........................................................................................ 97

6.2.4 Solubility screen ........................................................................................ 99

6.2.5 Initial crystal screening ............................................................................ 102

6.2.6 KVP40 59 + fork DNA ............................................................................ 104

6.2.7 KVP40 59 + KVP40 41 ........................................................................... 106

7 Chapter 7 ........................................................................................................ 108

Conclusions and Future Work .......................................................................... 108

References ......................................................................................................... 111

Appendices ......................................................................................................... 113

Appendix I: List of Abbreviations .................................................................... 114

Appendix II ......................................................................................................... 116
List of Figures

Figure 1.1: Bacteriophage T4 replication proteins at the fork.

Figure 1.2: Flow diagram outlining protein chemistry and crystallography methods used in my program of study.

Figure 2.1: New Brunswick Scientific Innova 4000 Incubator Shaker used for protein preparation.

Figure 2.2: Beckman Coulter™ TJ-25 centrifuge.

Figure 2.3: Branson™ Sonifier 250.

Figure 2.4: BIO-RAD BioLogic™ Duoflow™ HPLC system.

Figure 2.5: Initial crystal screening trays.

Figure 2.6: Honeybee sitting-drop crystallization robot.

Figure 2.7: Nikon SMZ1500 microscope and Nikon COOLPIX™ 990 digital camera.

Figure 2.8: 24-well format crystal trays.

Figure 2.9: Rainin EDP 10 ml programmable pipette.

Figure 2.10: In-house x-ray data collection sources.

Figure 2.11: BioCARS 14-BMC. a. Quantum 315 CCD detector. b. Close-up of Quantum 315 CCD detector.

Figure 2.12: BioCAD/SPRINT Perfusion chromatography system.

Figure 2.13: Speed-vac concentrator and refrigerated condensation trap.

Figure 3.1: Fork DNA substrates used in protein-DNA complexes.
Figure 3.2: Sequences of two fork DNA substrates used in protein-DNA complexes.

Figure 3.3: Poros HQ chromatograms from the purification of single-stranded DNA oligos.

Figure 3.4: TBE gel showing Poros HQ purified oligos.

Figure 3.5: TBE gel showing annealed DNA samples.

Figure 4.1: Ribbon diagram of T4 gene 59 protein.

Figure 4.2: Ribbon diagram of T4 gene 59 with positions of two cysteine atoms labeled and starred.

Figure 4.3: Native T4 gene 59 protein at room temperature.

Figure 4.4: Low-salt lysis protocol for 59C215S.

Figure 4.5: SDS-PAGE gel for low-salt lysis of 59C215S mutant.

Figure 4.6: SDS-PAGE gel of high-salt extraction of 59C215S.

Figure 4.7: Protocol for large-scale cell grow-up for 59C42S.

Figure 4.8: SDS-PAGE gel showing expression of 59C42S.

Figure 4.9: SDS-PAGE gel from low-salt lysis of 59C42S.

Figure 4.10: SDS-PAGE gel from high-salt extraction of 59C42S.

Figure 4.11: Purification protocol for native gene 59 protein.

Figure 4.12: SP Sepharose chromatogram for 59C42S.

Figure 4.13: SDS-PAGE gel from SP Sepharose run of 59C42S.

Figure 4.14: Poros HS chromatogram for 59C42S.

Figure 4.15: SDS-PAGE gel from Poros HS run of 59C42S.
Figure 4.16: Flow-chart showing sample preparation for nuclease activity gel samples of 59C42S.

Figure 4.17: Nuclease activity gel of 59C42S.

Figure 4.18: Fork DNA substrate used with binary complex.

Figure 4.19: SDS-PAGE gel showing stabilization of native T4 gene 59 protein by 20% glucose and TCEP.

Figure 4.20: Crystals of T4 gene 59 + glucose and TCEP.

Figure 5.1: Full-length T4 gene 32 protein displaying the three structurally and functionally distinct domains.

Figure 5.2: Cooperative binding of T4 32 protein on the lagging strand template arm.

Figure 5.3: Ribbon diagram of T4 gene 32 protein core.

Figure 5.4: SDS-PAGE gel showing protein expression of (32 minus B).

Figure 5.5: SDS-PAGE gel showing low-salt lysis of (32 minus B).

Figure 5.6: Purification protocol for (32 minus B).

Figure 5.7: Q Sepharose chromatogram from (32 minus B).

Figure 5.8: SDS-PAGE gel from Q Sepharose run of (32 minus B).

Figure 5.9: Poros PE real-time chromatogram for T4 (32 minus B) protein.

Figure 5.10: Poros HQ chromatogram for T4 (32 minus B) protein.

Figure 5.11: SDS-PAGE gel showing pure (32 minus B) protein fractions from Poros HQ.

Figure 5.12: Formation of a complex between T4 gene 32 protein and T4 gene 59 protein at the replication fork.
Figure 5.13a: Fork DNA substrate used in ternary complex screens with 59C42S and (32 minus B).

Figure 5.13b: Schematic diagram showing both T4 59C42S and (32 minus B) bound at the replication fork.

Figure 5.14: SDS-PAGE gel showing expression of (32 minus A).

Figure 5.15: SDS-PAGE gel showing results from low-salt lysis of (32 minus A).

Figure 5.16: Q Sepharose chromatogram for (32 minus A).

Figure 5.17: SDS-PAGE gel showing (32 minus A) fractions from Q Sepharose.

Figure 5.18: Poros HS chromatogram for (32 minus A).

Figure 5.19: SDS-PAGE gel showing results of Poros HS run of (32 minus A).

Figure 6.1: Sequence alignment of T4 gene 59 protein and KVP40 59 protein.

Figure 6.2: SDS-PAGE gel showing expression of KVP40 59 protein.

Figure 6.3: SDS-PAGE gel showing results of low-salt lysis for KVP40 59 protein.

Figure 6.4: SP Sepharose chromatogram for KVP40 59 protein.

Figure 6.5: SDS-PAGE gel showing KVP40 59 protein samples from SP Sepharose.

Figure 6.6: Poros HS chromatogram for KVP40 59 protein.

Figure 6.7: SDS-PAGE gel showing purified KVP40 59 protein.

Figure 6.8: Nuclease activity gel for KVP40 59.

Figure 6.9: Flow-chart displaying protocol for protein solubility screen.

Figure 6.10: Results from protein solubility screen done on KVP40 59 protein.

Figure 6.11: Hits from initial crystal screens of KVP 40 59 protein.
List of Tables

Table 1.1: DNA Replication and Repair Proteins from Bacteriophage T4.

Table 1.2: Instrumentation and techniques associated with protein chemistry and crystallography methods.

Table 2.1: Crystal Screening Kits.

Table 5.1: Key properties of the three domains and truncations of T4 gene 32 protein.

Table 6.1: 100 mM salts and buffers used for 16-condition solubility screen.
Chapter 1

Introduction

The overall objective for this master’s thesis project was to learn the techniques associated with protein crystallography in order to study proteins, namely DNA replication and repair proteins from Bacteriophage T4. This goal was to be achieved by working on research projects that would give me the opportunity to explore the characteristics of these proteins and begin to understand how they interact with other proteins and also DNA.

In this thesis, I will discuss the projects I have worked on since joining the Mueser laboratory in January of 2003. I will share the information that has been collected on these proteins in the past, the discoveries I made while studying these proteins and the direction in which these projects will now be heading.

1.1 DNA Replication

In the Mueser lab, we study how DNA replication and repair proteins recognize DNA in a structure-specific manner, not a sequence-specific manner. Bacteriophage T4 is used as a model system for DNA replication and encodes all proteins required for DNA replication (Cha and Alberts 1989). The T4 system has been studied for years and it is now known that many T4 proteins play a significant or essential role in DNA replication. In a reconstituted system, these
proteins catalyze the synthesis of leading and lagging strands and the speed of replication is comparable to that in vivo (Nossal, 1994).

The bacteriophage T4 genome encodes ten proteins that are involved in DNA replication and repair. The replication fork is shown below in Figure 1.1. The DNA polymerase (gene 43 protein) synthesizes the leading and lagging strands of the DNA by extending the pre-existing pentaribonucleotide primers and has 3'-5' editing exonuclease activity (Cha and Alberts 1989). There are three polymerase accessory proteins that increase the processivity of the polymerase on single-stranded templates and are required for nicked or forked duplex templates (Nossal, 1994). The circular clamp (gene 45 protein) holds the polymerase on the lagging strand template while the clamp loaders (gene 44/62 proteins) load the clamp onto the DNA (Jones, Green et al. 2004). The complex of the accessory proteins and the polymerase is referred to as the polymerase holoenzyme (Cha and Alberts 1989). The hexameric helicase (gene 41 protein) unwinds the duplex DNA ahead of the polymerase and interacts with the primase (gene 61 protein) to synthesize RNA primers (Burke, Munn et al. 1985). The single-stranded binding protein (gene 32 protein) cooperatively binds to the lagging strand of the DNA while stimulating lagging strand synthesis and increasing the rate of primer synthesis (Jones, Green et al. 2004). The helicase assembly protein (gene 59 protein) loads the helicase onto the DNA and binds at the replication fork (Mueser, Jones et al. 2000). While the helicase can move onto the DNA unassisted, gene 59 protein increases the rate at which the helicase is loaded onto the DNA, therefore increasing the rate of primer synthesis (Jones, Mueser et al. 2000).
RNaseH is then able to remove the pentamer primers from the 5' end of the Okazaki fragments (Bhagwat, Hobbs et al. 1997). The DNA polymerase again moves along the lagging strand, filling in the gaps left by the removal of the primers. DNA ligase repairs the nicks in the final step of the lagging strand synthesis. There is one more helicase, the dda helicase, which stimulates leading strand synthesis in vitro but is non-essential in vivo (Ma, Wang et al. 2004).

![Fig. 1.1: Bacteriophage T4 replication proteins at the fork.](image)

Clearly, a lot is known about the individual proteins that operate at the bacteriophage T4 replication fork. However, less is known about the specific interactions between these proteins. The goal of our research is to examine these interactions and learn about the complexes the proteins form between one another and how they interact with DNA. I have worked on two of the replication proteins, shown below in Table 1.1: single-stranded binding protein (gene 32) and the helicase assembly protein (gene 59).
Table 1.1: DNA Replication and Repair Proteins from Bacteriophage T4.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-stranded binding</td>
<td>32</td>
<td>Stimulates lagging strand synthesis.</td>
</tr>
<tr>
<td>protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicase assembly protein</td>
<td>59</td>
<td>Stimulates loading of helicase onto DNA.</td>
</tr>
</tbody>
</table>

(adapted from Nossal, 1994)

1.2 Research summary

As was mentioned above, the goal of this master’s research project was to learn the techniques associated with protein expression, purification and crystallization. Shown below in Figure 1.2 is an overview of the main methods discussed in this thesis.

![Flow diagram outlining the order of protein chemistry and crystallography methods used in my program of study.](image)

The projects I worked on began at the protein expression level. I was responsible for expressing T4 gene 59 protein and T4 gene 32 protein on a large scale and then using protein purification methods to get both of the proteins into a form that was suitable for crystallization screening. I was also responsible for
developing a purification protocol for single-stranded DNA substrates that were to be used with both T4 gene 59 and T4 gene 32 proteins for co-crystallization studies.

In my opinion, the most important aspect of the protein chemistry and crystallography methods are the techniques and instrumentation used with each method. As a master’s student, I was given the opportunity to learn how to use the instrumentation associated with the various methods used to study proteins. Shown below in Table 1.2 are the pieces of instrumentation used with each method shown in Figure 1.2.

Table 1.2: Instrumentation and techniques associated with protein chemistry and crystallography methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Technique/Instrumentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein expression</td>
<td>USB Luria broth, Castle™ gravity steam sterilizer, New Brunswick™ Scientific Innova 4000 Incubator Shaker, Beckman Coulter™ TJ-25 centrifuge, Invitrogen NuPage 4-12% SDS-PAGE gels</td>
</tr>
<tr>
<td>Protein purification</td>
<td>Branson™ Sonifier 250, BIO-RAD™ Biologic Duo-flow high-performance liquid chromatography system, BIO-RAD™ Biologic Quad-tec UV-visible detector, Invitrogen NuPage 4-12% SDS-PAGE gels, Agilent Technologies™ UV-visible spectrophotometer</td>
</tr>
<tr>
<td>DNA Purification</td>
<td>BioCAD/SPRINT perfusion chromatography system, Invitrogen 20% TBE gels, Agilent Technologies™ UV-visible spectrophotometer</td>
</tr>
<tr>
<td>Crystallization and X-ray diffraction data collection</td>
<td>Genomic Solutions Honeybee sitting-drop crystallization robot, A/B gradients, DCA crystal imaging system, Rigaku R-axis IV image plate, Rigaku Saturn 92 CCD detector</td>
</tr>
</tbody>
</table>
Once I became familiar with how each instrument worked, I was able to work independently and tailor the protocols associated with each instrument to a specific protein. All of these methods and many of the techniques and instruments will be discussed in detail in Chapter 2.
Chapter 2

Methodology

The methods used in this thesis research project helped me understand how proteins interact with each other. In this chapter, I will discuss the various methods and instrumentation that were used to examine the properties of DNA replication and repair proteins from bacteriophage T4.

2.1 Protein expression

After successfully cloning the protein into an expression host, the protein is expressed on a larger scale. The proteins I worked on were expressed in *Escherichia coli*. We grow the cells in a media called luria broth, or LB. This media is composed of casein peptone, yeast extract and sodium chloride. Six liters of LB is used in six, 2 L shaker flasks at 25 g/L. Cells inoculated with 25 mL of an overnight culture are allowed to grow for 3-4 hours while continuously being shaken in an incubator (New Brunswick) at 37 °C and 250 rmp, shown below in Figure 2.1.
The expression plasmids carry a resistance to a certain antibiotic so these antibiotics are added to the shaker flasks before the incubation period. Common antibiotics used in our laboratory are kanamycin, ampicillin and chloramphenicol.

At the end of the 3-4 hours, the cells are harvested by centrifugation (20 minutes, 5525xg). The centrifuge used in our laboratory is shown below in Figure 2.2.

The cell pellets, which contain the protein of interest, are stored at -20 ºC until needed for protein purification. The LB supernatant broth is autoclaved and then discarded.
2.2 Protein purification

Although the protein has successfully been expressed in its host, it is not yet in a form that is suitable for crystallization. There are other proteins and cell debris that need to be removed before the protein can be studied. If not removed, these additional proteins and cellular matter can contaminate further experiments and hinder the crystallization process.

The first step to remove the extraneous cellular matter is cell lysis. Hen egg-white lysozyme is added to lysis buffer to break open the cells and release all contents from the cell, including the protein that has been produced by the bacteria. Once the cells have been sufficiently degraded by lysozyme, the suspension is subjected to ultrasonic sound waves, using a sonicator, shown below in Figure 2.3, which bursts the bacterial protoplasts and shears any DNA that is present in the sample.

![Figure 2.3: Branson™ Sonifier 250.](image)

The solution is then centrifuged at a high speed (20,000xg) to separate the cell pellet from the supernatant, called the cell free extract (CFE), which now contains
the soluble protein. The CFE is decanted and stored at -80 °C and the cell pellet is discarded. If the protein is not found in the CFE at this point then further means of isolation need to be taken, such as adding various salts to the cell pellet to extract the protein.

High-performance liquid chromatography (HPLC) is the next step we use in the protein purification process. The HPLC we use for purification is shown below in Figure 2.4.

![Figure 2.4: BIO-RAD BioLogic™ Duoflow™ HPLC system.](image)

This process involves a mobile phase (protein solution + buffer and salt) to be passed over a stationary phase (solid media) with the protein eluted in a pure state. Solution is forced over the media using high flow rate and pressure. There are many characteristics of the protein that need to be taken into account before column purification can begin, such as protein pI, molecular weight and thermal stability. These characteristics are manipulated in such a way that allows for the protein to be separated from everything else in the CFE and this is done through the use of various types of solid media. Typically, a liquid salt gradient is used to
elute the protein and fractions are collected throughout the run. A UV-Vis
detector is also utilized so it is visible where the protein is eluting from the
column.

The native, folded structure of a protein is essential for biological activity
so the chromatographic method and conditions must be chosen carefully so as to
retain the three-dimensional structure of the protein throughout the separation
process (Huber, 2000). This is especially important if the protein will be used for
further experiments in which the biological activity is important. There are many
types of chromatography media one can use with HPLC but here I will focus on
the two types I utilized: ion-exchange chromatography and hydrophobic
interaction chromatography. Ion-exchange chromatography separates proteins
based on electrostatic interactions (pI) and hydrophobic interaction
chromatography separates proteins based on hydrophobic interactions.

There are two types of ion-exchange chromatography: cation exchange
and anion exchange. Cation exchange stationary phases are coated with
negatively charged groups while the stationary phases of anion exchange columns
are coated with positively charged groups. I utilized two cation exchange
chromatography columns: SP Sepharose is a low-resolution cation exchange and
Poros HS is a high-resolution cation exchange. The protein in the mobile phase
interacts with the groups coating the resin (stationary phase) based on ionic
strength. SP Sepharose is a derivatized cross-linked agarose with bound dextran.
Poros HS is a perfusion medium which allows for a much higher pressure limit,
flow-rate and higher resolution and is composed of poly(styrene-divinylbenzene),
which is an organic polymer. Both of these cation exchange columns have resins coated with sulfopropyl groups (-CH\textsubscript{2}CH\textsubscript{2}CH\textsubscript{2}SO\textsuperscript{3}\textsuperscript{−}), which allows for the positively charged protein to electrostatically interact with the groups coating the resin while the negatively charged groups pass through in the flow-through. The protein is then eluted from the column using a linear salt gradient, usually NaCl or NH\textsubscript{4}Cl.

I utilized two anion exchange chromatography columns: Q Sepharose and Poros HQ. Just like SP Sepharose and Poros HS, Q Sepharose is low-resolution anion exchange while Poros HQ is high-resolution anion exchange. The only difference between cation exchange and anion exchange is the charge of the groups that coat the resin. Positively charged quaternary polyethyleneimine groups coat the resin of Q Sepharose and Poros HQ, allowing for the negatively charged protein to electrostatically interact with these groups on the resin and adhere to the stationary phase while the positively charged groups pass through in the flow-through. The protein is then eluted from the column using a salt gradient similar to that of SP Sepharose and Poros HS. Q Sepharose resin is composed of 6% agarose with bound dextran while the Poros HQ resin is composed of poly(styrene-divinylbenzene).

Hydrophobic interaction chromatography (HIC) is governed by interactions between the mobile phase and the stationary phase however the interactions are very different from those of ion-exchange chromatography. HIC separates proteins based on their hydrophobic properties. The stationary phase has a very hydrophobic, non-polar surface and the mobile phase is polar, usually
water (Huber, 2000). The hydrophobic interaction column we use is Poros PE, which will remove any contaminating *E. coli* nuclease from the protein. As the protein solution is loaded onto the column, the endogenous nuclease in the protein sample interacts with the stationary phase through hydrophobic interactions and then protein is eluted in the flow-through. A high-salt buffer containing 600 mM NH$_4$Cl is utilized, which allows for the nuclease to interact with the non-polar stationary phase, removing it from the sample.

2.3 Nuclease activity

Contaminating nuclease in a protein sample can prevent crystal formation. If you attempt to crystallize a protein in the presence of DNA, the nuclease in the protein sample will degrade the DNA and prohibit the complex from forming.

Poros PE removes endogenous nucleases but an agarose gel can be used to confirm that all of the nuclease has been removed from the sample. Pet vector (Pet 28) and λ DNA are run in the presence and absence of the protein sample being studied. This gel is analyzed upon completion. By analysis of the gel, we are able to determine whether or not the contaminating nuclease has been removed from the sample. If it has not been removed, other means of removal need to be taken before the protein is used for initial crystal screens.

2.4 Protein crystallization

Once the protein has been isolated, purified and the removal of endogenous nuclease has been confirmed, the next step is to attempt to crystallize the protein. The ultimate goal of our research is to produce diffraction quality
crystals. As many know, this is not an easy task and requires a lot of patience. Once diffraction quality crystals have been obtained, we can start to gather structural information about the protein. Structural information is used by professionals in the pharmaceutical, medical and biochemical fields to develop drugs, vaccines and different treatments for various diseases plaguing those in this country and around the world.

Protein crystals are highly-ordered molecules and are quite different from those of minerals and other substances. They are not packed nearly as tight and are consequently, much less dense. This loose-packing allows for solvent channels to pass between the molecule. Solvent can be a very high percentage of the chemical make-up of a protein crystal. Consequently, protein crystals are oftentimes not very stable and can be very temperamental. They will frequently not grow in anything other than the solution there were purified in.

There are a number of crystal screening kits that can be used when setting up initial crystallization screens. Sparse matrix screening has proved to be useful in the past (Jancarik and Kim, 1991) and this type of screen, among others, are commonly used in our lab. The kits used frequently in our lab are shown below in Table 2.1.
Table 2.1: Crystal Screening Kits

<table>
<thead>
<tr>
<th>Name of Screen</th>
<th>Type of Screen</th>
<th>Distributor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystal Screen I™</td>
<td>Sparse matrix</td>
<td>Hampton Research</td>
</tr>
<tr>
<td>Crystal Screen II™</td>
<td>Sparse Matrix</td>
<td>Hampton Research</td>
</tr>
<tr>
<td>Natrix™</td>
<td>Sparse Matrix</td>
<td>Hampton Research</td>
</tr>
<tr>
<td>Wizard I™</td>
<td>Random sparse matrix</td>
<td>Emerald Biostructures</td>
</tr>
<tr>
<td>Wizard II™</td>
<td>Random sparse matrix</td>
<td>Emerald Biostructures</td>
</tr>
<tr>
<td>Cryo I™</td>
<td>Cryo-protectant screening kit</td>
<td>Emerald Biostructures</td>
</tr>
<tr>
<td>Cryo II™</td>
<td>Cryo-protectant screening kit</td>
<td>Emerald Biostructures</td>
</tr>
<tr>
<td>Ion Screen</td>
<td>Anion, cation, and pH matrix screen</td>
<td>Prepared in-house</td>
</tr>
</tbody>
</table>

Screens are set up at room temperature and at 4 °C in order to determine the optimal growth condition for the protein. Most of these screens consist of 48 different conditions so two screens can be set-up in one, 96-well sitting-drop plate. Single-drop and multi-drop trays (3 drops per well) are utilized for these screens, shown below in Figure 2.5.

![Figure 2.5: Initial crystal screening trays. a. Corning 96-well tray. b. Greiner 96-well round bottom tray.](image)

There are two tray set-up methods that can be utilized. Trays can be set up manually, using a multi-channel pipette to pour the trays and a programmable pipette to add the protein and well solutions, which is very time consuming and tedious. The other method is to set up the tray using a robotic system. The Honeybee sitting-drop crystallization robot we are able to use is shown below in
Figure 2.6. This piece of instrumentation is a new addition to our facility and is part of the Ohio Macromolecular Crystallography Consortium. This robot greatly shortens the amount of time it takes to set up trays, as the tray set-up is automated.

Figure 2.6: Honeybee sitting-drop crystallization robot (part of the Ohio Macromolecular Crystallography Consortium).

Once these screens have been set up, they must be checked to determine whether or not the protein crystallized. Protein crystals can be very small and are oftentimes not visible to the naked eye so the screens must be viewed under a microscope, shown below in Figure 2.7.

Figure 2.7: Nikon SMZ1500 microscope and Nikon COOLPIX™ 990 digital camera.

When a protein appears to have crystallized in the screen, this is deemed a “hit”. All of the conditions are then given a score based on size, shape and
appearance, ranging from 0 (clear) to 10 (3-D crystals > 0.2 mm). Protein growth is also confirmed by staining the potential hits with a blue stain (Izit stain, Hampton Research). If the crystal absorbs the color by turning a dark blue, we can be confident that it is a protein crystal.

Once the positive hits have been identified, they are then optimized in expansion trays with the aim of producing single, diffraction quality crystals. Expansion trays can be set up using 24-well, 48-well and 96-well trays. The 24-well trays we use in our laboratory are shown below in Figure 2.8.

The method we use to set-up these trays was developed in our lab and is called A/B gradients. In this method, shallow and coarse gradients are used to vary a single chemical parameter around the initial crystallization conditions (Senger and Mueser, 2004, in press). Parameters that can be varied include, but are not limited to, buffer, PEG and salt concentration. With this method, two solutions are prepared, an initial solution (A) and final solution (B), using standardized pipetting maps. Step gradients are prepared by adding decreasing amounts of A to consecutive wells followed by an increasing volume of B. Coarse gradients are used to further optimize the growth conditions and shallow gradients are used to
produce diffraction quality crystals. A Rainin-EDP programmable pipette is used to prepare the trays, shown below in Figure 2.9.

![Rainin EDP 10 ml programmable pipette](image)

This pipette has a programmable variable feature that can hold up to 12 steps. An entire 24-well tray can be pipetted within a matter of minutes using one program and one pipette tip.

2.5 Protein crystal manipulation

Now that the conditions for crystal growth have been optimized and the crystals have grown in these optimized conditions, the crystals need to be manipulated in a way that will allow for them to be put on an x-ray diffractometer. This is done so data on the individual crystals can be collected and eventually, the structure of the protein can be solved. As stated above, protein crystals are fragile so much care must be taken when manipulating them for data collection. The crystal packing is also rather loose so it is oftentimes difficult to transfer the crystal from one liquid medium to the next. The protein crystals must also be placed in a cryoprotectant that will freeze clear as glass
when the crystal is placed in a stream of liquid nitrogen for data collection. Some solutions that have proved to be good cryoprotectants include, but are not limited to, glucose, glycerol, sucrose, 1,2-propanediol, 2-methyl-2,4-pentanediol (MPD) and ethylene glycol. The cryoprotectant will stabilize the crystal so it will not crack, split or completely break apart when placed in the liquid nitrogen. There is not a universal cryoprotectant (McPherson 2004) so it oftentimes takes a while to determine the best cryoprotectant for a protein.

When the cryoprotectant conditions have been optimized the protein crystal(s) can be mounted for data collection. There are two mounting methods we use: capillary, for room temperature collection and loop, for data collection in the cold. The capillaries used for room temperature data collection are either glass or quartz and filled with a liquid solution similar to the protein crystal solution. Most of the liquid surrounding is then very carefully removed from the capillary, leaving a small amount behind so the crystal will not dry out. Both ends of the capillary are then sealed with beeswax and the capillary can then be mounted on a diffractometer suited for room temperature data collection. Loop mounting is used for data collection at very low temperatures. A nylon loop is attached to a small, metal tube which is then inserted into a steel base, collectively called a crystal mounting pin. The pin can then be mounted onto a goniometer head for data collection. There are loops of various sizes, since protein crystals normally range in size from 20 µm to 1 mm. The crystals are picked out of the solution with the loop and then either frozen in liquid nitrogen or placed directly into the cryostream of the diffractometer.
2.6 X-ray diffraction data collection

X-ray diffraction data collection can be used once the protein crystals have been properly mounted. An X-ray diffractometer is used to generate a diffraction pattern of the proteins we study. Crystals are mounted on the diffractometer and the crystal is exposed to X-rays. The crystal will diffract some of the X-rays, forming a unique diffraction pattern. This pattern is then collected and processed to obtain data unique to the protein being studied. This data is then phased and used in molecular modeling to solve the structure of the protein.

We are fortunate enough to be able to use two in-house X-ray detectors: a Rigaku R-Axis IV image plate detector and a Rigaku Saturn 92 CCD detector, both shown below in Figures 2.10a and b. We also have a new Rigaku FR-E X-ray source.

![Figure 2.10: In-house x-ray data collection sources. a. Rigaku Saturn 92 CCD detector. b. Rigaku R-Axis IV image plate detector.](image)

While the R-axis IV image plate has been a part of our facility for quite some time, the CCD, or charge-coupled device, detector is a new addition. These pieces of instrumentation give us two different X-ray diffraction options. The CCD has a much faster read-out time and does not take as long as the R-axis IV to
screen crystals and collect data. The image plate can be used for both room temperature and low temperature data collection while the CCD can only be used for data collection at low temperatures.

We are also fortunate to be able to utilize the facilities at the Advanced Photon Source at Argonne National Laboratory. This facility is a national synchrotron-radiation light source funded by the Department of Energy. APS is in the shape of a ring and its circumference is one mile and there is one very powerful source of X-rays for the different beamlines to use. Electrons are accelerated to near the speed of light and their directionality is then changed using bending magnets to form a ring. The electrons are decelerated when bent by the magnet and then emit high-brilliance radiation. There are thirty-five beamlines currently in operation at APS and there is one bending magnet per beamline. We used beamline 14-BMC at BioCARS, shown below in Figure 2.11.
This beamline is part of the Consortium for Advanced Radiation Sources and is used for structural biology research. I made three trips to APS while a member of the Mueser laboratory and although I was not able to take my own crystals on any of the trips, I was able to assist others with mounting and screening crystals and collecting data.
2.7 DNA purification

Macromolecular crystallographic techniques are used to study both protein-protein and protein-DNA interactions. By studying these interactions, we gain insight into the nucleic acid binding site of the protein(s) as well as other DNA-binding activities. Various DNA substrates can be used to optimize the crystal growth conditions of these complexes by using different types of DNA, such as single-stranded, double-stranded and fork DNA substrates. Altering the sequence of the substrate, the location of the binding site and the length of the oligo can also assist in optimizing the growth conditions (Ellenberger, 2001).

Perfusion chromatography, using a BioCAD/Sprint workstation, was employed for the rapid purification of the oligos and this method had been studied previously (Ellenberger, 1996). The BioCAD/SPRINT perfusion chromatography system (PerSeptive Biosystems, Inc.) used for purification is shown below in Figure 2.12.

Anion exchange chromatography has been discussed as it pertains to proteins and it is similar with nucleic acids. Oligonucleotides are small DNA molecules that
have both their hydrophilic sugar-phosphate backbone and the hydrophobic nucleobases exposed to the solution (Huber, 2000) so anion exchange is the most obvious method. A Poros HQ anion exchange column is used for the purification of the single-stranded DNA substrates, which would be annealed after purification. During purification, the negatively charged DNA will interact with the positively charged column resin and are then eluted with a salt gradient (Huber, 2000).

In our purification protocol, a linear gradient of NH₄OAc is run and DNA fractions are collected. The absorbance at 296 nm is measured so the elution of the DNA from the column can be viewed throughout the experiment. The wavelength that is commonly used to measure DNA is 260 nm however if we used this wavelength, the high absorbance of the DNA would saturate the measurement, making it difficult to obtain an accurate absorbance reading. The wavelength we use (296 nm) is at the threshold of measuring DNA and while the signal will be weaker, the absorbance will be in the range of sensitivity and allow for analysis of the chromatogram. The whole process of loading the sample onto the column and then running the sample across the linear gradient takes all of approximately 20 minutes, which is relatively quick.

The DNA samples are then concentrated overnight in a Speed-vac, shown below in Figure 2.13.
The NH₄OAc and NH₄OH in the chromatography buffers will evaporate off of the samples as they are concentrated so 100x TE buffer (1 µL buffer/100 µL of solution) is added to the samples before concentrating to avoid depurination of the oligos. TE is used in the hybridization buffer for DNA annealing so it is very convenient to add it at this point so the sample is ready to be annealed after being concentrated in this manner. The Speed-vac is attached to a refrigerated condensation trap, shown above in Figure 2.3, which collects all of the condensation as the samples are dried.

The purity of the samples is confirmed through the use of TBE (Tris, boric acid, EDTA) gels. These are non-denaturing gels that are run at 4 ºC. The gel is then viewed by UV shadowing to confirm the purity of the samples after being run over the Poros column and then concentrated. The concentrations of each oligo are then determined by measuring the absorbance at 260 nm, taking into account the dilution factor and extinction coefficient.
2.8 DNA annealing

Single-stranded DNA samples can be annealed in order to form double-stranded and fork DNA substrates. Each oligo has a given sequence and in order to anneal that oligo to another single-stranded oligo, their sequences must be complementary, either entirely (to form dsDNA) or partially (to form fork DNA).

The purified single-stranded oligos must be placed in a hybridization buffer of 1x TE and 100 mM NaCl before the annealing process can begin. Once the hybridization buffer has been added, the oligo(s) are heated to 100 °C (boiling) in water and then cooled slowly overnight by surrounding it with styrofoam, which acts as an insulator. Annealing is confirmed by using TBE gels. The single-stranded starting material and annealed material are both loaded onto the gel. Upon completion, the gel is viewed under a UV lamp and annealing is confirmed if the single-stranded material has migrated farther down the gel than the annealed sample. The annealed sample will not travel as far because it is larger than the single-stranded oligos.
Chapter 3

DNA Purification

Various protein-DNA complexes will be discussed in depth in this thesis and the method to purify the DNA substrates used in these complexes is of great importance and needs to be explained. We were fortunate enough to have a DNA purification system in our lab, which enabled us to purify DNA in-house for all of the protein-DNA complexes discussed in the previous chapters.

In this chapter, I will explain why we used the different DNA substrates in each complex. I will also discuss the protocol that I developed for DNA purification.

3.1 Background

It will be discussed in the next chapter that T4 gene 59 protein has a binding preference for fork DNA substrates. It will also be discussed that T4 gene 32 protein binds to the lagging strand template arm of fork DNA. In knowing these two pieces of information, binary and ternary complexes of these proteins in the presence of DNA were studied: T4 59C42S + fork DNA and T4 59C42S + T4 (32 minus B) + fork DNA and KVP40 59 + fork DNA. All of these complexes will be discussed in depth in the next chapters.
Two fork DNA substrates were used for the initial screens and they are shown below in Figure 3.1. The sequences of these two fork substrates are shown below in Figure 3.2. The complimentary duplexes are identical between the two forks and the leading strand does not vary between the two substrates. The leading and lagging strand template arms are non-complimentary, as is shown in the figure.

![Figure 3.1: Fork DNA substrates used in protein-DNA complexes. a. lead9+9/lag9+9 used with T4 59C42S and KVP40 59. b. lead9+9/lag9+15 used with T4 59C42S and (32 minus B).](image)

<table>
<thead>
<tr>
<th>a. lead9+9/lag9+9</th>
<th>b. lead9+9/lag9+15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ GTCCTAACT</td>
<td>5’ CAAGCAGTCTAACT</td>
</tr>
<tr>
<td>TTGAGGCAG 3’</td>
<td>TTGAGGCAG 3’</td>
</tr>
<tr>
<td>AACTCCGTC 5’</td>
<td>AACTCCGTC 5’</td>
</tr>
<tr>
<td>3’ TGATGGCAG</td>
<td>3’ TGATGGCAG</td>
</tr>
</tbody>
</table>

![Figure 3.2: Sequences of two fork DNA substrates used in protein-DNA complexes. a. lead9+9/lag9+9 used with T4 59C42S and KVP40 59. b. lead9+9/lag9+15 used with T4 59C42S and (32 minus B).](image)

The fork substrate shown in Figures 3.1a and 3.2a, with leading and lagging strands of identical length at 18 bases, was used with the T4 59C42S + fork DNA and the KVP40 59 + fork DNA screens. The 59 protein has been shown to load the helicase onto the fork with a gap of at least 6 but no more than 12 bases on the lagging strand (Mueser, Jones et al. 2000) so this substrate should have been well-
suited for the complex. This substrate was also used with the KVP40 59 + fork DNA screens, since the KVP and T4 59 proteins are of the same size. The substrate shown in Figures 3.1b and 3.2b, with a leading strand that is 18 bases long and a lagging strand template arm of 15 bases, was used with the T4 59C42S + T4 (32 minus B) + fork DNA ternary complex screens. A single 32 protein molecule covers 6-7 nucleotides on the lagging strand (Jones, Green et al. 2004) so with the extension, this substrate should have also been well-suited for the screens.

3.2 Purification protocol

The standard desalted oligos were ordered in a 1 μmole synthesis from Integrated DNA Technologies. The desalting means the protecting groups have been removed and the oligo released from the resin using ammonia. The samples arrived lyophilized. The DNA samples were then diluted to 2 mL with deionized water to prepare the samples for purification.

Anion exchange chromatography was used to purify the single-stranded DNA substrates and Poros HQ was the anion exchange column used in this purification protocol. We chose to use volatile chromatography buffers, which contained ammonia, because the use of the volatile buffers would allow the ammonia to evaporate off of the samples during the concentration process, leaving the pure DNA sample behind. The initial (A) buffer used was 10 mM NH₄OH and 1 M NH₄OAc while buffer B was also composed of 10 mM NH₄OH but the NH₄OAc concentration was increased to 3 M. The Poros HQ column was equilibrated with these buffers and the individual samples were then loaded onto
the column and run across the gradient. The absorbance at 296 nm was monitored throughout the run (refer to Chapter 2). Shown below in Figure 3.3a, b and c are the chromatograms from Poros HQ for all three substrates that were purified.
Figure 3.3: Poros HQ chromatograms from the purification of single-stranded DNA oligos. a. Chromatogram for the lead18 oligo. b. Chromatogram for the lag18 oligo. c. Chromatogram for the lag24 oligo. The blue peak monitors the absorbance at 296 nm.
Figure 3.3a shows the chromatogram for the lead18 (lead9+9) oligo and Figure 3.3b displays the chromatogram for the lag18 (lag9+9) oligo. Interestingly enough, these two oligos did not stick to the Poros HQ column but passed through and were collected in the flow-through, which is why the gradient was not run with these two substrates. It is unclear why this happened but it is possible that the column was overloaded. Although they did not stick to the column, they were pure when they eluted in the flow-through. Figure 3.3c displays the chromatogram from the lag24 (lag9+15) oligo and this substrate did stick to the column and the gradient was run. The DNA started to elute from the column in fraction 7, peaking in fraction 8.

The two flow-through samples collected from the first two runs and fractions 7 and 8 from the last run were then run on a TBE gel to check the purity of the samples. This gel is shown below in Figure 3.4.

After analysis of this gel, it was determined that all three of the substrates were pure because a single band was visible on the gel for each oligo. You will also notice that the lag24 substrate did not migrate as far down the gel as the other two substrates, which is what was expected since it was slightly larger than the lead18 and lag18 substrates.
The samples were aliquoted into 1.5 mL Eppendorf tubes and 1 µL of 100x TE was added to each tube before putting the samples in the Speed-vac to concentrate. The heat setting on the Speed-vac was turned to medium and the samples were spun overnight. In the morning, the substrates had been sufficiently concentrated, as a viscous, gel-like DNA sample remained in each tube. A volume of 50 µL of autoclaved milli-Q water was then added to each tube to dissolve the viscous sample and slightly dilute the very concentrated DNA. The fractions from each run were then pooled together into one-1.5 mL Eppendorf tube and the absorbance at 260 nm was taken. The absorbance was then used to determine the concentrations of each oligo, taking into account the extinction coefficient unique to each oligo and the dilution factor.

Once the single-stranded samples had been purified and the concentrations determined, the DNA could be annealed to form the fork DNA substrates needed for crystal screening. A 100 nmol amount of each sample was used, annealing lead18 to lag 18 and lead18 to lag24, forming the fork substrates shown in Figure 3.1. The samples were placed in an annealing buffer, or hybridization buffer, before the annealing process started. The annealing buffer used was 1x TE and 100 mM NaCl. Since the TE had been added before the samples were concentrated, the NaCl was the only component of the buffer that needed to be added. Once the volumes of each oligo to be used had been determined and the samples were added together, 100 mM NaCl was added to each microtube. The samples were then placed in a beaker filled with 500 mL of water on a hot plate. Once the water boiled, the samples were then placed in a styrofoam container,
still sitting in the beaker, surrounded by pieces of styrofoam. The styrofoam acts as an insulator and cools the DNA down at a slow rate, giving the samples a chance to properly anneal.

The samples were left in the styrofoam container overnight and another TBE gel was run in the morning to determine whether or not the samples had annealed. A sample of the single-stranded substrates (starting material) and the annealed DNA were run on the gel. The gel run after the annealing of lead18 + lag 18 and lead18 + lag 24 is shown below in Figure 3.5.

![Image of gel](image)

**Figure 3.5:** TBE gel showing annealed DNA samples. Lanes 1 and 15 are DNA gel loading solution. Lane 2 is lead18, lane 4 is lag24 and lane 6 is the annealed sample (lead18 + lag24). Lane 9 is lead18, lane 11 is lag18 and lane 13 is the annealed sample (lead18 + lag18).

Analysis of the gel proved that the samples had annealed properly. The annealed samples were heavier than the single-stranded oligos and in both cases (lanes 6 and 13) the annealed samples did not migrate as far down the gel as the starting material, which is what we would expect of a larger substrate. Also, a single band was seen for both annealed samples and we concluded that the entire DNA sample had been annealed. The purified and annealed fork DNA substrates were then ready for crystal screening.
Chapter 4

T4 gene 59 protein

The native structure of T4 gene 59 protein has been solved and the work published but there are still structural interactions of this protein that need to be studied further. It has been shown that this protein binds preferentially to fork DNA, however crystal studies of 59 protein bound to fork DNA have been unsuccessful, to date. Gene 59 protein also binds to gene 32 protein so this protein-protein interaction needs to be studied further.

In this chapter, I will discuss the work that I have been doing with T4 gene 59 protein. I will give a detailed background of two protein mutants and discuss the work that I have completed with these proteins. I will also discuss the more recent work that I have been doing with a new crystal form of native T4 gene 59 protein.

4.1 Background

The helicase assembly protein is 26 kDa in size, 217 amino acids in length and is required for recombination-dependent DNA replication (Mueser, Jones et al. 2000). This protein accelerates the loading of the hexameric helicase onto the fork, which consequently accelerates the formation of the primase-helicase complex and the synthesis of the pentamer primers (Jones, Mueser et al. 2004).
Gene 59 protein is monomeric and is composed of two domains of similar size called the N-domain (residues 1-108) and the C-domain (residues 109-217). The theoretical pI of this protein is 9.37 so, as expected, the surface of the protein is coated with basic residues (Mueser, Jones et al. 2000). Mueser, et al. solved the native structure of this protein to 1.45 Å (Figure 4.1) and found that it reveals an α-helical bundle fold. The N-terminus is shown in blue and the C-terminus is shown in red. In studying the structure of this protein, it was discovered that there is not an obvious DNA binding groove, however surface-exposed hydrophobic residues suggest sites for potential contact with fork DNA substrates (Mueser, Jones et al. 2000).

The helicase assembly protein interacts with both 32 and 41 proteins, while also having the ability to bind to single and double-stranded DNA (Ishmael, Alley et al. 2001). However, it has been shown in gel shift assays that gene 59 protein has an affinity for fork DNA substrates with single and double-stranded arms of at least six but no more than 12 nucleotides long (Mueser, Jones et al. 2000). Gene 59 protein has also been shown to accelerate the binding of gene 32
protein onto lagging strand template arms that are too short for cooperative binding (Jones, Green et al. 2004). The complex formed between these two proteins is unusually strong and they are able to bind to each other even in the absence of DNA (Jones, Green et al. 2004). The same is true for the complex between the 41 helicase and 59 protein.

As stated above, native T4 gene 59 protein has been shown to preferentially bind to fork DNA substrates (Jones, Green et al. 2004). While this protein does also bind ssDNA and dsDNA, it has a preference for fork DNA substrates. However, the results of initial crystals screens of native T4 gene 59 in the presence of fork DNA substrate were not favorable. The protein precipitated in the presence of DNA. This protein also precipitates in the presence of sulfate and phosphate, which might explain why it would precipitate in the presence of the DNA sugar-phosphate backbone. So, why would a proven DNA binding protein precipitate in the presence of DNA, especially fork DNA, which it has been shown to preferentially bind to?

One possible answer to this question became more evident when the amino acid sequence of the native protein was further examined. This protein is composed of 217 amino acids, two of them being cysteines located at position 42 and 215, shown below in Figure 4.2.
In the native protein, these two cysteine molecules will form internal disulfides, which form a ladder when the protein is at room temperature. This ladder is shown below in Figure 4.3.

The cross-linking of the cysteines and consequent formation of this ladder could contribute to the precipitation of this protein in the presence of DNA. After the formation of this ladder was discovered, site-directed mutagenesis was used at locations 42 and 215 and the cysteines were mutated to serine and alanine,
forming four protein mutants: C42A, C42S, C215A and C215S. The immediate
goal of my project was to develop a purification protocol for the mutants with the
aim of preparing the proteins for crystallization.

Work had already been done on the alanine mutants before I started
working in the Mueser lab so I focused my work on the two serine mutants: C42S
and C215S. I was given *E.coli* glycerol stocks of the proteins; the cloning had
already been done when I joined the lab.

4.2 59C215S Experiments

Pooja Talaty had been working with this mutant of the 59 native protein
and had successfully expressed the protein on a large scale, growing the protein in
LB. However, she had not developed a purification protocol for the protein so
that is where I started my work.

4.2.1 Cell lysis

A purification protocol had previously been developed for the native T4
59 protein so that protocol was followed for the mutant. The first step was to do a
low-salt lysis on the cells collected in the large scale grow-up. I started with
7.70 g of C215S cells and after the cells had been thawed, I added 70.3 mL of
low-salt buffer to the cells (146 mL of buffer per 16 g of cells). The low-salt
buffer was composed of 50 mM BisTris HCl, pH 6.5, 100 mM NH₄Cl, 1 mM
DTT and 1 mM AEBSF. The protocol for low-salt lysis is shown below in Figure
4.4.
To prepare the 100 µL lysis sample for the SDS-PAGE gel, the sample was spun at 10,000xg for 2 min in a tabletop centrifuge. The cell lysate was then separated from the pellet and 25 µL of 2x SDS was added to each tube. These samples were then run on an SDS-PAGE gel. The gel was run at room temperature for 35 min at 200 V. The results of the gel are shown below in Figure 4.5.
While the protein was found in the lysate there was also protein that remained in the pellet. When a low-salt lysis was done on the native protein, all of the protein was found in the pellet while other proteins and impurities were found in the lysate. As a consequence, a high-salt extraction could then be done on the low-salt pellet containing the native protein and the protein could be extracted out of the pellet. The protein in the lysate was fairly pure at that point. Since the mutated protein extracted into the low-salt pellet, we decided to do a high-salt extraction on the remaining low-salt pellet.

The cells from the low-salt lysis were thawed and a high-salt buffer composed of 50 mM BisTris HCl, pH 6.5, 750 mM NH₄Cl, 10 mM MgCl₂, 1 mM DTT and 1 mM AEBSF was added. A very small amount of hen egg white lysozyme was also added. The same protocol for stirring, sonicating and centrifuging the cell solution was used. The SDS-PAGE gel from the high-salt extraction is shown below in Figure 4.6.
The protein that was in the low-salt pellet was still not completely extracted into the lysate after the high-salt extraction. The next option was to do a 1 M (NH$_4$)$_2$SO$_4$ extraction on the high-salt pellet. This extraction was never done, as the work being done on the other mutant (59C42S) was proceeding with greater success.

4.3 59C42S Experiments

I began my work with this protein at the expression level. As the cloning had already been done, I needed to express the protein on a large scale, using LB as the medium, and develop a purification protocol with the aim of preparing pure protein for crystallization screening. I was given a glycerol stock of 59C42S that had been prepared by Pooja Talaty and that glycerol stock was the source of 59C42S protein.

4.3.1 Protein expression

The 59C42S protein had been cloned into BL21 (DE3) Gold pLysS cells and a glycerol stock of 59C42S had been made. The first experiment was to do a large scale preparation of the protein. A swab of the glycerol stock was added to
3-250 mL Erlenmeyer flasks and put in a shaker overnight at 37 °C, each flask containing 100 mL of LB. The 59C42S mutated protein is resistant to ampicillin and the cell line is resistant to chloramphenicol so these antibiotics were added to the overnights at 100 mM (total volume), which amounted to a volume of 100 µL. The entire protocol for the large scale cell grow-up is shown below in Figure 4.7.

Figure 4.7: Protocol for large-scale cell grow-up for 59C42S.

1. Glycerol stock
   ↓
2. 100 mM AMP and CAM added to 100 mL of LB in 3-250 mL Erlenmeyer flasks.
   ↓
3. Shake overnight culture at 37°C.
   ↓
4. 50 mL of solution from each flask added to 6-2000 mL flasks containing 1000 mL of LB.
   ↓
5. 100 mM AMP and CAM added to each flask.
   ↓
6. Shake at 37°C until O.D.₆₀₀ = 0.3-0.6.
   ↓
7. 1 mL 0-hr sample taken; Cell growth induced by adding 1 mM IPTG (total volume) to each flask.
   ↓
8. Shake for 3 hrs. at 37°C.
   ↓
9. 1 mL 3-hr sample taken; cells collected by centrifugation at 5525xg for 10 min.
   ↓
10. Cells weighed and stored at -20°C; LB autoclaved and discarded.

To prepare the 0-hr and 3-hr samples for the SDS-PAGE gel, the tubes were spun at 10,000g for 2 min in a tabletop centrifuge. The supernatant was then taken out of each tube and discarded. A 25 µL volume of bug buster was added to each tube to dissolve the pellet and 25 µL of 2x SDS was added to each tube.
The gel was run at 200 V for 35 min at room temperature. The results of the gel are shown below in Figure 4.8.

![Figure 4.8: SDS-PAGE gel showing expression of 59C42S. Lane 1 is the 0-hr sample, lane 2 is the molecular weight marker and lanes 3-7 are the 3-hr samples.](image)

The 25 kDa expressed well and the cells collected from this grow-up could now be used to purify the protein for crystal screens.

4.3.2 Protein purification

Now that 59C42S had successfully been expressed on a large-scale and the cells collected, a protocol for protein purification was needed. A protocol had previously been developed for the native protein so we started by attempting to adapt this protocol to the mutated protein. The first step in the purification protocol was low-salt cell lysis. This is the same low-salt lysis that was done with the 59C215S mutant (refer to Figure 3.3). The SDS-PAGE gel from the low-salt lysis for 59C42S is shown below in Figure 4.9.
The results of the low-salt lysis were very similar to those for 59C215S. The protein was found in both the cell pellet and lysate. However, instead of doing a high-salt extraction on the low-salt pellet, another grow-up was done and the high-salt extraction was done on fresh cells. The same high-salt lysis buffer was used and the protocol used for the high-salt extraction is identical to that used for the low-salt lysis. The SDS-PAGE gel from the high-salt extraction is shown below in Figure 4.10.

The protein was primarily found in the lysate, which means that the high-salt had extracted the protein from the pellet. The lysate was also partially pure at this point, as some of the impurities and other proteins had been removed from the cell pellet. Since the high-salt extraction worked, we can now omit the low-salt
lysis step in the protocol and just do a high-salt extraction of the protein from the
pellet as the first and only step of the cell lysis.

The protein had now been extracted from the cell but there were still
impurities that needed to be removed. Ion-exchange chromatography was
naturally the next step in the purification process. We started by using the same
protocol that had been developed for use with the native protein using SP
Sepharose followed by Poros HS, which is shown below in Figure 4.11.
1. Equilibrate **SP Sepharose** with buffer B (high-salt), followed by buffer A (low-salt).

2. Adjust conductivity of protein sample to be at or slightly below that of SP Sepharose Buffer A using 25 mM BisTris HCl pH 6.5.

3. Load protein lysate sample onto SP Sepharose column at 5 ml/min.

4. Collect flow-through in case protein does not stick to the column.

5. Run SP Sepharose method, which will elute the protein using a salt gradient and collects 5 ml fractions using a fraction collector.

6. Collect samples from fractions thought to contain protein and run on an SDS-PAGE gel to make sure protein is contained in those fractions.

7. Pool protein fractions and adjust conductivity to be at or slightly below that of **Poros HS** buffer A using 25 mM BisTris HCl pH 6.5.

8. Equilibrate Poros HS with buffer B (high-salt) followed by buffer A (low-salt).

9. Load sample onto column at 10 ml/min, collecting flow-through.

10. Run Poros HS method, which will elute the protein using a salt gradient and collects 3 ml fractions.

11. Run SDS-PAGE gel to check purity of protein sample.

12. Pool pure protein fractions, concentrate and store at -80°C until needed.

Before loading the sample onto SP Sepharose, the protein lysates from the high-salt extraction were thawed and the chromatography buffers were prepared. SP Sepharose buffer A contained 25 mM BisTris HCl, pH 6.5, 100 mM NH₄Cl and...
10 mM MgCl$_2$. Buffer B contained the same concentrations of MgCl$_2$ and BisTris HCl, pH 6.5 but the salt concentration was increased to 750 mM. The protein was then run over the SP Sepharose column using a salt gradient of 0-100% Buffer B. The chromatogram from this run is shown below in Figure 4.12.
Figure 4.12: SP Sepharose chromatogram for 59C42S. The green line monitors the absorbance at 280 nm, the purple monitors the absorbance at 260 nm, the red monitors the conductivity and the black monitors the salt gradient. Fractions labeled with an “x” were run on an SDS-PAGE gel.
The protein started eluting from the column in fraction 13, peaking in fraction 19. An SDS-PAGE gel was then run, stepping across the 280 nm peak, to check for protein and purity of the sample. Samples from fractions 14, 17, 19, 21, 24 and 27 were prepared for an SDS-PAGE gel. Twenty-five microliters of each fraction was collected and 25 µL of 2x SDS was added to each sample. The SDS-PAGE gel is shown below in Figure 4.13.

![Figure 4.13: SDS-PAGE gel from SP Sepharose run of 59C42S. Lane 1 is the molecular weight marker. Lanes 2-7 are fractions 14, 17, 19, 21, 24, and 27, respectively.]

We concluded from the gel that the protein did elute from the column and was partially pure at this point. Fractions 14 thru 27 were pooled and this sample was then loaded on the Poros HS column to further purify the protein. The chromatography buffers used for Poros HS are the same as those used for SP Sepharose. The conductivity of the sample was again adjusted to be at or below that of buffer A. Once the conductivity had been adjusted, the protein sample was run over the Poros HS column. The chromatogram from this run is shown below in Figure 4.14.
Figure 4.14: Poros HS chromatogram for 59C42S. Fractions labeled with an “x” were run on an SDS-PAGE gel to check protein purity.
The protein began eluting from the column in fraction 43, peaking in fraction 48. An SDS-PAGE gel was run after this column to check the purity of the sample. Fractions 48 and 50 were run on the gel, which is shown below in Figure 4.15.

![Figure 4.15: SDS-PAGE gel from Poros HS run of 59C42S. Lane 1 is the molecular weight marker, lane 2 is fraction 48 and lane 3 is fraction 50.](image)

After examining the SDS-PAGE gel, the protein was deemed pure enough for further experiments and fractions 43-53 were pooled and concentrated.

### 4.3.3 Nuclease activity

The protein had been expressed, extracted and purified. One more experiment needed to be done before the protein would be ready for co-crystallization screening with DNA. As was discussed in the Chapter 2, any nuclease that may be present in the sample needs to be removed before the crystal screens can be set-up. If present, endogenous nuclease can ruin a crystallization experiment set-up in the presence of a DNA substrate. We wanted to attempt to crystallize T4 59C42S in the presence of fork DNA, so we needed to make sure that nuclease had been removed from the sample by the purification method.

In order to check for the nuclease, an agarose gel was run. First, a DNA sample is incubated in the presence and absence of 59C42S and then the gel is
We used λ DNA ladder and Pet 28 as nuclease substrates. These two samples were chosen because they are used in our lab in molecular biology experiments so they are readily available.

Concentrated protein was used for this nuclease activity study. When the protein was concentrated, a flow-through solution was collected and used as the diluent for the nuclease activity experiment. A flow-chart showing the sample preparation is shown below in Figure 4.16.

1. 1 µL λ DNA added to tube 1 and 1 µL of Pet 28 added to tube 2.

2. Add 9 µL of concentration flow-through added to each sample.

3. 5 µL of each sample then added to 0.5 µL of 59C42S.

4. Now have four tubes total.

5. Pet 28, Pet 28 + 59C42S, λ DNA, λ DNA + 59C42S

6. Incubate samples for 1 hr. at 37 ºC.

7. Add 2 µL of gel loading solution to each sample and run gel.

The gel is run at 90 V for 45 min at room temperature. The gel was stained with ethidium bromide and a UV transilluminator was used to examine the gel. The results from running the gel are shown below in Figure 4.17.
Stop buffer was added to the samples before running them on the gel. The stop buffer contained 23% glycerol, 0.075 M EDTA, 5.6% SDS and bromphenol blue. The binding of 59C42S would normally retard the DNA and prevent it from migrating into the gel but the stop buffer used would denature 59C42S protein and allow DNA to migrate into the gel. This would allow us to get a clear picture of the status of the DNA and if any nuclease was present in the sample. When this gel was examined, it was determined that there may have been some background nuclease activity present and the possible degradation product is shown with a yellow arrow in the above figure. If there was a large amount of nuclease present in the protein sample, the bands towards the top of the gel in lanes 2 and 4 would not be present in lanes 1 and 3 although the degradation product band would still be present towards the bottom of the gel in lanes 1 and 3. With a large amount of nuclease present, the λ DNA and Pet 28 would have been completely degraded. However, the DNA was not degraded, as shown by the identical bands seen in the presence and absence of protein. I decided to go ahead with crystallization experiments since the DNA bands were present. After reviewing the results of the gel and crystal screens, a better strategy would have been to add EDTA to the protein dialysis buffer, which would have chelated the
metals in the small amount of nuclease that may have been present in the protein sample, destroying the nuclease activity and increasing the possibility of crystal growth.

4.3.4 Crystal screening with fork DNA

Initial crystal screens of T4 59C42S + fork DNA were set-up in September of 2003. Pure DNA substrates (Oligos etc.) were provided by Dr. Nancy Nossal, our collaborator on this project from NIH. The substrate used is shown below in Figure 4.18.

![Figure 4.18: Fork DNA substrate used with binary complex.](image)

The two single-stranded oligos, one the leading strand and one the lagging strand, were of identical length at 18 bases long. The two oligos were annealed before the screens were set up. The protein concentration was 0.578 mM (15 mg/ml) and the DNA concentration was 3 mM. Although the concentrations differed, the same number of nanomoles of each component was used to form the 1:1 complex. The DNA was added to the protein sample and once the two were combined, the sample was incubated at room temperature for 10 min, giving the components of the complex a chance to bind.

Four Corning 96 wells trays were then set up, two at room temperature and two at 4 °C. The four screens used were the Natrix, Ion Screen, Wizard I and
Cryo I (refer to Table 2.1). Well solution (1 µL) was pipetted into each drop, followed by 1 µL of the protein/DNA sample. The trays were then taped and stored until examined. The trays were scored 6 and 10 days after being set up. Unfortunately, there were not any crystal hits in these initial screens. Over 70% of the drops precipitated, which suggested that the protein and DNA concentrations were not too low. Had the samples not been concentrated enough, many of the wells would have been clear.

There were a few options that could be explored for the project at this point. The first option was to design a new fork DNA substrate. However, T4 gene 59 protein had been shown to bind to fork DNA substrates with template arms of at least six but no more than 12 bases long (Mueser, Jones et al. 2000). Based on the current mode, the substrate we used would have been the correct length.

The next option was to set up another set of trays using different screens. However, when the second set of crystal screens was attempted, the protein precipitated in the presence of the fork DNA. An attempt was made to stabilize the complex by aliquoting out the sample and adding various cryoprotectants (glycerol, glucose, sucrose, MPD, ethylene glycol) but this was not successful. Interestingly enough, there were two aspects of this second set-up that were different from the first set-up: the protein was from a different prep and the DNA had been purified in-house. Although the protein had been prepared in the same manner as the first sample there is no guarantee that the DNA was prepared the same way in-house as it had by a technician at Oligos, etc. The second DNA
sample was also purchased from a different company (Integrated DNA Technologies) and they were purchased as desalted oligos. These two aspects of the second attempt to set up the screens could have made a difference in the stability of the complex, resulting in precipitation of the sample.

A third attempt was made at the crystallization of this co-complex. The complex was not formed before the screens were set-up but as the screen was being set-up. The DNA would be added to the drop after the well solution and protein had been added. The DNA was diluted with water to try and prevent precipitation from happening upon complex formation with the 59 mutant. Although it would take the drops a longer time to equilibrate, since the DNA was very dilute, crystal formation may occur, instead of immediate precipitation.

Two trays were set-up at room temperature and two were set-up at 4 °C. The screens used were Natrix, PEG/Ion screen and crystal screens I and II. The protein concentration was 8 mg/mL (0.32 mM) and the initial DNA concentration (before addition of water) was 0.53 mM. All trays were set-up using the Honeybee sitting-drop crystallization robot and the DNA was then pipetted by hand. The DNA was pipetted at room temperature for two trays and at 4 °C for the cold trays. After the DNA had been pipetted, the trays were left at their respective temperatures to equilibrate.

The trays were checked three times within two weeks of set-up and there was not any crystal growth seen. In the room temperature trays, 97% of the drops precipitated and in the cold trays, 90.7% of the drops precipitated, which averages to 94% overall. Nearly all of the drops had precipitated within the first three days
and there was not much change, if any, in the next two weeks. The dilution of the DNA with distilled water did not appear to help stabilize the complex. Pipetting the DNA into the drop after the protein and well solution did not appear to help either.

4.4 Crystallization of native T4 gene 59 + glucose

In addition to the work I have been doing with the T4 gene 59 protein mutant, I have also been working on growing a new crystal form of the native 59 protein. As stated in Chapter 2, cryoprotectants are used to stabilize the protein so it will not break, split or crack when frozen in liquid nitrogen for data collection. It was discovered that glucose, a common cryoprotectant, stabilizes one of the crystal forms of the native 59 protein. It was also discovered that TCEP (tris [2-carboxyethyl] phosphine), which like DTT reduces disulfide bonds, also stabilizes this protein. A gel showing this is shown below in Figure 4.19.
It was discussed earlier in this chapter how the two cysteines in the native protein form disulfide bonds, cross-linking, and form a ladder at room temperature, shown above in lane 2 of the SDS-PAGE gel. DTT was added to this protein at room temperature to try and reduce the disulfide bonds and dissolve the ladder but it did not work (lane 3 above) as the ladder was still present. When TCEP was added to the protein, it did a much better job of reducing the ladder (lane 4 above) and when the protein was in the presence of both TCEP and 20% glucose (lane 5) the monomer was observed. This work was done by Dr. Mueser before I joined the lab and although he was able to grow crystals in the presence of glucose and TCEP and data was collected, he had not pursued modeling the glucose into the structure. Three past members of the Mueser lab also worked with this crystal form of the protein but were not successful in growing crystals. I started working on this project with the goal of growing diffraction quality crystals and then after collecting data, being able to model the glucose into the structure.

4.4.1 Expansion trays

Since Dr. Mueser had previously grown diffraction quality crystals of this protein, the well conditions for production crystals had been determined when I started working on this project. The optimized crystal growth conditions were 12-18% PEG 6000, 100 mM TAPS pH 8.5, 100 mM NaOAc, 10 mM TCEP and 20% glucose. The native protein had been modified, using selenomethionine in place of methionine residues in the protein. I started by attempting to grow crystals of the selenomethionine 59 protein. Seven trays were set up over a time period of four months, starting in October of 2003, with protein concentrations ranging
from 11 – 21 mg/mL. Crystals had previously been grown at 15 mg/mL and usually took seven to ten days to grow. Although a few crystals grew, none of them were single crystals or diffraction quality.

A second attempt was made to grow crystals of 59 native in the presence of glucose and TCEP, beginning in June of 2004. This time, unmodified native protein was used. Two diffraction quality crystals were grown at a protein concentration of 16.5 mg/mL. The crystals were rather small, at 200 and 100 µm respectively, but looked decent enough to mount and put on the diffractometer. These two crystals are shown below in Figure 4.20.

<table>
<thead>
<tr>
<th>Figure 4.20: Crystals of T4 gene 59 protein + glucose and TCEP. a. 100 µm crystal. b. 200 µm crystal.</th>
</tr>
</thead>
</table>

There was not a need to optimize the cryoprotectant conditions, since the crystals were grown in the presence of glucose. Once mounted on a pin, the crystals would be ready to go onto the diffractometer.

4.4.2 Data collection

We were able to utilize the new instrumentation for the Ohio Macromolecular Crystallography Consortium here at UT. We decided to use the new CCD and the 200 µm crystal was the first of these two crystals put on the machine, diffracting to 2.4 Å. The data was integrated and scaled to 2.8 Å, with a mosaicity of approximately 0.7. This mosaicity is high, but it is considerably lower than the mosaicity of the crystals that had been grown previously (> 1.0).
After discussing this new data set with Dr. Mueser, he decided to look back through the data sets he had collected on this crystal form. We ended up re-integrating one of these older synchrotron data sets and were able to integrate and scale the data to 2.0 Å with a mosaicity of 0.2. It was this data set that was used to build the structure and model in the glucose.
Chapter 5

T4 gene 32 protein

T4 gene 32 protein is a single-stranded binding protein that has three structurally and functionally distinct domains. This attribute allows the protein to be modified, forming three distinct truncations.

In this chapter, I will discuss the work I have been doing with T4 gene 32 protein. I will give a detailed background on each truncation and discuss the work I have been doing with two of the three protein truncations.

5.1 Background

T4 gene 32 protein is a single-stranded binding protein that stimulates lagging strand synthesis and cooperatively binds to the lagging strand template arm of the replication fork. This 34 kDa protein is 301 amino acids in length. It is a monomer and contains three structurally and functionally specific domains (Villemaim, Ma et al. 2000). Figure 5.1 shown below displays the three domains of T4 gene 32 protein.
The A domain is the acidic domain, the B domain is the basic domain and the core domain is the protein core. The A domain (residues 254-301) is responsible for the interactions with other T4 proteins (Waidner, Flynn et al. 2001). The B domain (residues 1-21) is responsible for the cooperative binding to the lagging strand template arm of the replication fork (Ma, Wang et al. 2004), as seen below in Figure 5.2. If the B domain were to be removed, the cooperativity would be lost.

The core domain (residues 22-253) contains a structural Zn (II) molecule and is also the location of the single-stranded DNA binding site (Villemain, Ma et al. 2000). The theoretical pI at the C-terminus is 3.9 while the theoretical pI at the N-terminus is much more basic at 8.7.
The structure of the core protein bound to ssDNA was solved to 2.2 Å (Shamoo, Friedman et al. 1995). A zinc (II) molecule was found to be bound in the protein core. The ribbon diagram of the protein core is shown below in Figure 5.3. The zinc is bound to three cysteine molecules and one histidine, which is similar to a zinc-finger protein. This protein also has a typical OB-fold, or oligonucleotide/oligosaccharide fold. Proteins with this fold contain five β-strands connected by an α-helix (Agrawal and Kishan 2001).

Figure 5.3: Ribbon diagram of gene 32 protein core. The Zn (II) molecule is shown as a grey sphere.

T4 gene 32 protein stimulates lagging strand synthesis at the replication fork and is able to form a complex with gene 59 protein. It is also been found that 32 protein is required for helicase-dependent leading strand DNA synthesis when the helicase is loaded by 59 protein (Jones, Green et al. 2004). The 32 protein is not required when the helicase is not loaded by the helicase assembly protein, suggesting that 32 and 59 protein have a strong affinity for each other.

T4 gene 32 protein can be modified to form three structurally and functionally distinct truncations: (32 minus A), (32 minus B) and the protein core. The removal of the C or N-terminus does not eliminate the DNA binding
properties of the protein (Waidner, Flynn et al. 2001). Table 5.1 shown below displays the key properties of each of the three T4 gp 32 protein domains and truncations.

Table 5.1: Key properties of the three domains and truncations of T4 gene 32 protein.

<table>
<thead>
<tr>
<th>Protein truncation/domain</th>
<th>Amino acids</th>
<th>Molecular Weight</th>
<th>pI</th>
<th>Major function</th>
</tr>
</thead>
<tbody>
<tr>
<td>B domain</td>
<td>1-21</td>
<td>~3 kDa</td>
<td>8.70</td>
<td>Responsible for gp32-gp32 interactions.</td>
</tr>
<tr>
<td>A domain</td>
<td>254-301</td>
<td>~6 kDa</td>
<td>3.90</td>
<td>Key role in interactions with other T4 replication proteins</td>
</tr>
<tr>
<td>Core domain</td>
<td>22-253</td>
<td>26 kDa</td>
<td>5.40</td>
<td>Contains single-stranded DNA binding site and structural Zn$^{2+}$ ion.</td>
</tr>
<tr>
<td>32 minus A</td>
<td>1-253</td>
<td>28 kDa</td>
<td>8.12</td>
<td>Binds to gene 32 protein but not other T4 replication proteins.</td>
</tr>
<tr>
<td>32 minus B</td>
<td>22-301</td>
<td>31 kDa</td>
<td>4.76</td>
<td>Binds to other T4 replication proteins but not gene 32 protein.</td>
</tr>
</tbody>
</table>

The (32 minus A) truncation, which removes the acidic domain of the protein, has a molecular weight of 28 kDa and consists of residues 1-253 (Waidner, Flynn et al. 2001). It is the acidic (A) domain of the protein that is responsible for interactions with other T4 replication proteins (Ishmael, Alley et al. 2001) so with the removal of this domain, this truncation will not interact with those proteins, such as the gene 43 polymerase, gene 61 primase and gene 59 helicase assembly protein. The (32 minus B) truncation, which does not contain the basic domain of the protein, has a molecular weight of 31 kDa and consists of residues 22-301. The basic domain of the protein is responsible for the cooperativity of the gene 32 protein, a function which is absent in the (32 minus B) protein (Ma, Wang et al. 2004). The (32 minus B) truncation contains the A domain so it is this truncation that will interact with other T4 replication proteins. The third truncation, the
protein core, has a molecular weight of 26 kDa and consists of residues 22-253. This truncation contains the intrinsic nucleic acid binding site (Waidner, Flynn et al. 2001).

We were very interested in the functions of the (32 minus B) protein because it is this truncation, along with the full-length protein, that would interact with other T4 replication proteins. Extensive gel shift assays have been done, exploring the DNA binding and protein binding properties of this protein and these recent assays have shown that 32 protein is essential for leading strand synthesis when the helicase has been loaded by T4 gene 59 protein (Jones, Green et al. 2004). Although this study did not discuss (32 minus B), extensive work was done with the full-length protein. In knowing that the acidic domain is responsible for interactions with other T4 proteins, we wanted to explore the possible interactions between (32 minus B) and T4 gene 59 protein.

5.2 (32 minus B) Experiments

We were fortunate enough to receive this truncation from Richard Karpel at the University of Maryland, Baltimore so the cloning of (32 minus B) was not done in the Mueser lab. The goal of this project was to express the protein on a large scale and develop a purification protocol with the aim of using the purified protein for crystal screening. Since (32 minus B) was the main focus, we wanted to co-crystallize T4 59C42S in the presence of (32 minus B) and also crystallize both of these proteins in the presence of fork DNA.
5.2.1 Protein expression

We started our work with this truncation with protein expression on the large scale. The same expression protocol was used with this protein as was used with T4 gene 59 protein (refer to Figure 4.6). The expression construct of (32 minus B) carries a resistance to ampicillin so this antibiotic was added to the overnight cultures (100 mM) and then also added to the shaker flasks before inoculation. The SDS-PAGE gel for protein expression is shown below in Figure 5.4. Since the B domain has been removed to form this truncation, the apparent molecular weight is 31 kDa, slightly less than that of the full-length protein (34 kDa).

![Figure 5.4: SDS-PAGE gel showing protein expression of (32 minus B). Lane 1 is the molecular weight marker, lane 2 the 0 hr. sample and lanes 3 and 4 the three hours samples.](image)

The protein expressed well at the apparent molecular weight that corresponded to this truncation (31 kDa) so the cells collected from this large-scale grow-up were used to develop the purification protocol.

5.2.2 Protein purification

The first step of purification is cell lysis and we started with a low-salt lysis. The low-salt lysis protocol that was used with the 59C215S and 59C42S
protein mutations (refer to Figure 4.3) was used with (32 minus B) as well. The low-salt lysis buffer used with this protein contained 100 mM NaCl, 40 mM Tris pH 8.0, 10 mM MgCl$_2$, 2 mM CaCl$_2$ and 1 mM EDTA. The SDS-PAGE gel from the low-salt lysis is shown below in Figure 5.5.

![Figure 5.5: SDS-PAGE gel showing low-salt lysis of (32 minus B). Lane 1 is the molecular weight marker, lane 2 the low-salt lysis pellet and lane 3 is the low-salt lysis lysate.](image)

The protein was extracted from the cell pellet in the low-salt lysis, so we then used the lysates to further purify and isolate the protein.

Affinity chromatography was used by other research groups to isolate (32 minus B). A purification protocol had previously been formulated for the full-length protein starting with dialysis after lysis, followed by DEAE Sephacel and ssDNA cellulose column chromatography (Waidner, Flynn et al. 2001). Although this protocol worked, there were several limitations that we wanted to avoid, such as low flow rate (0.5 mL/min) which consequently made this protocol very time consuming. Also, the resin used with the ssDNA cellulose is a low-capacity resin, which was not amenable to large scale purification.

This pI of the C-terminus of T4 gene 32 protein is 4.76 so we decided to use anion exchange affinity chromatography with (32 minus B), which contained
the C-terminus but not the N-terminus, making the truncation acidic. The proposed chromatography protocol is shown below in Figure 5.6.

<table>
<thead>
<tr>
<th>Step</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Equilibrate <strong>Q Sepharose</strong> column with buffer B (high-salt) followed by buffer A (low-salt).</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>2.</td>
<td>Adjust conductivity of protein lysates to be at or slightly below that of Q Sepharose buffer A.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>3.</td>
<td>Load sample onto column at 5 mL/min., collecting flow-through in case protein does not stick to the column.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>4.</td>
<td>Run Q Sepharose method, which elutes the protein using a salt gradient and collects 5 mL fractions.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>5.</td>
<td>Collect samples from fractions thought to contain protein and run on an SDS-PAGE gel to make sure protein is contained in those fractions.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>6.</td>
<td>Equilibrate <strong>Poros PE</strong> with high-salt PE buffer.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>6.</td>
<td>Pool Q Sepharose fractions containing protein and load sample onto Poros PE reverse-phase column.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>7.</td>
<td>Collect protein in flow-through and once protein has been collected, adjust conductivity to be at or slightly below that of <strong>Poros HQ</strong> buffer A.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>8.</td>
<td>Equilibrate Poros HQ column with buffer B (high-salt) followed by buffer A (low-salt).</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>9.</td>
<td>Load sample onto Poros HQ column, collecting flow-through.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>10.</td>
<td>Run Poros HQ method, which elutes protein using a salt gradient and collects 3 mL fractions.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>11.</td>
<td>Run SDS-PAGE gel to check purity of protein sample.</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>12.</td>
<td>Pool fractions containing pure protein, concentrate and store at -80°C until needed.</td>
</tr>
</tbody>
</table>

**Figure 5.6:** Purification protocol for (32 minus B).
Before loading the (32 minus B) protein sample onto Q Sepharose, the lysates containing the protein from the low-salt lysis were thawed and the chromatography buffers to be used for Q Sepharose were prepared. Buffer A contained 25 mM BisTris HCl, pH 6.5, 50 mM NaCl and 10% glycerol (total volume) and Buffer B contained the same concentration of BisTris HCl, pH 6.5 and glycerol but the NaCl concentration was increased to 500 mM. The protein sample was loaded onto the column and Q Sepharose was then run using a salt gradient of 0-100% Buffer B. The chromatogram from Q Sepharose is shown below in Figure 5.7.
Figure 5.7: Q Sepharose chromatogram from (32 minus B). Red, green, purple and black traces are the same as for 59C42S SP Sepharose chromatogram (refer to Figure 3.12). Fractions labeled with an “x” were run on an SDS-PAGE gel.
After analyzing the above chromatogram, we came to the conclusion that the protein started eluting from the column in fraction 61, peaking in fraction 68. We stepped across the peak and ran a number of fractions between 61 and 78 on an SDS-PAGE gel. The results of that gel are shown below in Figure 5.8.

The (32 minus B) truncation eluted at the predicted molecular weight and was partially pure at this point. Fractions 63-78 were pooled and directly loaded on Poros PE.

Poros PE is governed by hydrophobic interactions between the stationary phase and the mobile phase, as was discussed in Chapter 2. It was found in the purification of the full-length protein that the contaminating nuclease in the sample interacts with the groups coating the resin and the protein is eluted in the flow-through (Jones, Green et al. 2004). The column was equilibrated with a buffer composed of 25 mM Tris, pH 7.5, 50 mM NaCl, 600 mM (NH₄)₂Cl, 2 mM β-mercaptoethanol and 1 mM EDTA. The fractions pooled from the Q Sepharose were directly loaded onto the column. Once loaded, (32 minus B) eluted in the flow-through, which was seen on the real-time chromatogram, shown below in Figure 5.9.
Figure 5.9: Poros PE real-time chromatogram for T4 (32 minus B) protein. The green line is the 280 nm trace, the purple line is the 260 nm trace, and the red line monitors the conductivity.
The (32 minus B) sample was collected manually until the 280 nm trace reached zero again. Once all of the protein had been collected, the column was equilibrated with buffer and the sample was then loaded onto the Poros HQ column.

As was discussed in Chapter 2, Poros HQ is a high-resolution anion exchange column. When the negatively charged protein sample is loaded onto the column, it interacts with the positively charged quaternary amino groups coating the resin and “sticks” to the column. Protein is then eluted across a salt gradient of NH₄Cl and/or NaCl. When this (32 minus B) sample came off of the Poros PE column, the conductivity was rather high so before loading onto Poros HQ, the conductivity of the sample was adjusted with 25 mM BisTris HCl, pH 6.5 to be at or below that of Buffer A. The same buffers used with Q Sepharose are used with Poros HQ.

Once the conductivity of the protein sample had been adjusted, the column was equilibrated with buffers A and B and the (32 minus B) sample was loaded onto the column. The chromatogram from Poros HQ is shown below in Figure 5.10.
Figure 5.10: Poros HQ chromatogram for T4 (32 minus B) protein. The colored lines are the same for this column as for other affinity chromatography columns (refer to Figure 4.12). Fractions labeled with an “x” were run on an SDS-PAGE gel.
The (32 minus B) sample started eluting from the column in Fraction 30, peaking in Fraction 34. An SDS-PAGE gel was run after Poros HQ to check the purity of the sample. The protein peak on the chromatogram was not a nice, sharp peak so we were concerned that the protein may have had reduced solubility on the column. The SDS-PAGE gel would be able to show us whether or not all of the protein fractions were pure. This gel is shown below in Figure 5.11.

Fractions 34, 35, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, and 55 were run on this gel and the sample was pure and expressed at the expected molecular weight. As stated above, we were concerned that there was a solubility problem during the run but as seen in the gel, all of the fractions were pure. Fractions 30 thru 55 were pooled and concentrated and the concentrated (32 minus B) would then be used for initial crystal screens.

5.2.3 Crystal screening with 59C42S

The pure protein from both the 59C42S and (32 minus B) protein preps were subjected to crystal screens as a binary complex. The C-terminal A-domain of 32 protein binds to T4 gene 59 protein, forming a 1:1 complex. These two
proteins even bind in the absence of DNA, so we wanted to set up initial crystal screens as a binary complex.

Before making the complex that was to be used for the crystal screens, we wanted to do an experiment on the analytical scale to check the solubility of the binary complex. One microliter of T4 gene 59C42S was added to 1 µL of T4 (32 minus B) protein on a microscope slide. When this was done, the proteins did not precipitate so the initial crystal screens were then set up. The screens were set up at room temperature, using the Honeybee sitting-drop crystallization robot, with each protein at 8 mg/mL, making the concentration of the binary complex between 8 and 10 mg/mL. The first set of screens was set up in early June of 2004 and Crystal screens I and II and Index screens were used. Most of the drops precipitated and not a single crystal hit was seen in any of these initial screens. A second set of screens was set up in late June of 2004, using the PEG/Ion screen and Wizard I and II screens. Again, most of the drops precipitated and no crystal growth was seen. These screens were checked at least once a week for two months and no crystal growth was seen in any of the screens.

5.2.4 Crystal screening with 59C42S and fork DNA

At the same time that the initial crystal screens of 59C42S + (32 minus B) were set up, screens of the two proteins in the presence of a fork DNA substrate were set up as well. As was discussed in Chapter 1, T4 59 protein and T4 32 protein form a complex at the replication fork, shown below in Figure 5.12.
The DNA substrate used for the screens was designed based on the results from the DNA binding studies that had been done previously with T4 gene 59 and T4 gp 32 (Jones, Green et al. 2004). T4 gene 59 protein preferentially binds to fork DNA with a lagging strand template arm of at least six but not more than 12 nucleotides long (Mueser, Jones et al. 2000). A single molecule of T4 gene 32 protein covers six nucleotides on the lagging strand template arm so the substrate used needed to be long enough to bind both gene 59 protein and gene 32 protein. The substrate used in these screens in shown below in Figure 5.13a.

This substrate was purified in-house (Chapter 3). The 15 nucleotide length of the lagging strand template arm should allow for both T4 59C42S to bind at the fork and T4 (32 minus B) to bind to the template. A schematic diagram of 59C42S and (32 minus B) bound at the replication fork is shown below in Figure 5.13b.
The (32 minus B) truncation binds non-cooperatively on the lagging strand template arm so one molecule of (32 minus B) will bind to one molecule of 59C42S, as shown above in the diagram.

The same experiment on the analytical scale was done before formation of the ternary complex. A 1 µL volume of 59C42S and 1 µL of (32 minus B) were added to 1 µL of DNA on a microscope slide. The complex did not precipitate so the initial crystal screens were then set up. The screens were set up at room temperature with each protein at 8 mg/mL while the DNA concentration was 0.843 mM, which made the concentration of the complex between 8 and 10 mg/mL. The Honeybee sitting-drop crystallization robot was used to set up the first set of screens in early June of 2004. The Index screen and Crystal screens I and II were used and most of the drops precipitated. A few fibers were seen in the screens but no crystal growth was seen. A second set of screens was set up in late June of 2004, using the PEG/Ion and Wizard I and II screens, and again, most of the drops precipitated and no crystal growth was seen. These screens were also checked at least once a week for two months and no crystal growth was seen. Since neither binary nor ternary complex produced any crystals in the initial sets of screens, this project is at a stand-still at the moment. In the future, more work will be done with this project.
5.3 (32 minus A) Experiments

As was discussed earlier in this chapter, we obtained two of the three protein truncations of the full-length T4 32 protein. The (32 minus A) truncation does not contain the C-terminal acidic domain that is responsible for interactions with other T4 replication proteins. Theoretically, this truncation would not interact with T4 gene 59, T4 gene 43 or T4 gene 61 proteins. However, the crystal structure of this truncation has not been solved, so a purification protocol needed to be developed for this protein truncation in order to set up initial crystal screens of (32 minus A) in the presence and absence of fork DNA.

Just like (32 minus B), we were fortunate enough to receive glycerol stocks of (32 minus A) from Richard Karpel at UM-Baltimore, so we did not do the cloning for this truncation. The goal of this project was to express (32 minus A) on a large scale and develop a purification protocol with the aim of using the purified protein for initial crystal screens.

5.3.1 Protein expression

The first experiment done with this protein truncation was expression on the large scale. Again, the same protocol used with 59C42S was used with this truncation (refer to Figure 4.6) the only difference was the antibiotic used. The plasmid for (32 minus A), like (32 minus B), carries a resistance to ampicillin so this was added to the overnight cultures and the shaker flasks at 100 mM (total volume). An SDS-PAGE gel showing protein expression is shown below in Figure 5.14.
The (32 minus A) truncation is approximately 28 kDa and we saw good expression at this molecular weight, as seen in the above figure. The cells collected from this grow-up were used to develop a purification protocol.

5.3.2 Protein purification

The first step of the purification is cell lysis. The same low-salt lysis protocol used with 59C215S was used with this truncation (refer to Figure 4.3) except the lysis buffer used with this protein contained 100 mM NaCl, 40 mM Tris HCl, pH 8.0, 10 mM MgCl₂, 2mM CaCl₂ and 1 mM EDTA, which is the same low-salt lysis buffer used with (32 minus B). The SDS-PAGE gel run after the lysis is shown below in Figure 5.15.
Ion-exchange chromatography was again used to purify (32 minus A) that was now in the lysate. A purification protocol had been developed for (32 minus B) so this protocol was used with (32 minus A) as well – Q Sepharose followed by Poros PE and then Poros HQ. The lysate was run across Q Sepharose using a salt gradient of 0-100% Buffer B and the chromatogram is shown below in Figure 5.16.

Although the lanes were slightly smeared, it appeared that the low-salt lysis had extracted the protein from the cell pellet into the lysate. At this point, the protein had been extracted in the lysate and could be purified further.

Ion-exchange chromatography was again used to purify (32 minus A) that was now in the lysate. A purification protocol had been developed for (32 minus B) so this protocol was used with (32 minus A) as well – Q Sepharose followed by Poros PE and then Poros HQ. The lysate was run across Q Sepharose using a salt gradient of 0-100% Buffer B and the chromatogram is shown below in Figure 5.16.
Figure 5.16: Q Sepharose chromatogram for (32 minus A). Fractions labeled with an “x” were run on an SDS-PAGE gel.
The 280 nm peak on the above chromatogram is rather broad and the protein appeared to begin eluting in fraction 25 however, only the fractions around the peak were considered. The early protein fractions (25-43) could still contain a lot of other protein debris, even after being run across the salt gradient of this column. By not including these early fractions, some of the protein would be discarded. However, there is more protein found in the later fractions, shown by the increase in absorbance at 280 nm. There was not a real risk of losing very much of the protein by discarding the early fractions. An SDS-PAGE gel was run with (32 minus A) samples from Q Sepharose, starting with Fraction 44. The gel is shown below in Figure 5.17.

![Figure 5.17: SDS-PAGE gel showing (32 minus A) fractions from Q Sepharose. Lane 1 is the molecular weight marker and lanes 2-7 are fractions 47, 50, 53, 56, 59 and 62, respectively.](image)

The protein expressed at the expected molecular weight and Fractions 44-65 were then pooled and loaded directly onto Poros PE.

The same protocol and buffer were used for Poros PE with (32 minus A) as had been used previously with (32 minus B) (refer to section 5.2.2). The protein again eluted in the flow-through while the contaminating nuclease stuck to the column. The real-time chromatogram of this truncation looked exactly like that of (32 minus B) (refer to Figure 5.9). Once the (32 minus A) sample was...
collected, the conductivity was adjusted to be at or below that of Poros HQ buffer A and the sample was then loaded onto Poros HQ.

When the (32 minus A) sample was loaded onto Poros HQ, it did not stick – it eluted in the flow-through. The fact that the protein did not interact with the quaternary amine groups of the Poros HQ resin was troubling. The (32 minus B) truncation had stuck so why wouldn’t this truncation stick as well? We then took a few steps back and examined the individual truncations. The (32 minus B) truncation contains the C-terminal acidic domain and has a pI of 4.76. This acidic pI made (32 minus B) a perfect candidate for anion exchange chromatography. However, (32 minus A) does not contain the C-terminal acidic domain but it does contain the N-terminal basic domain of this protein. This would make the pI more basic, which it does. We found that the pI of this truncation is 8.12 which would make it a better candidate for cation exchange chromatography. This does not explain why (32 minus A) stuck to Q Sepharose but we wanted to load it onto Poros HS to check our theory.

The flow-through containing (32 minus A) collected from Poros HQ was loaded onto Poros HS and a salt gradient of 0-100% Buffer B was used. The chromatogram from Poros HS is shown below in Figure 5.18.
Figure 5.18: Poros HS chromatogram for (32 minus A). Fractions labeled with an “x” were run on an SDS-PAGE gel.
The (32 minus A) protein sample did stick to this column, as we had hypothesized, and we saw a nice, sharp 280 nm peak. The (32 minus A) sample started eluting from the column in Fraction 16, peaking in Fraction 19. An SDS-PAGE gel was then run to check the purity of the sample. This gel is shown below in Figure 5.19.

![Figure 5.19: SDS-PAGE gel showing results of Poros HS run of (32 minus A). Lanes 1-6 are Fractions 16, 17, 18, 19 and 21, respectively.](image)

Although a large amount of (32 minus A) was recovered, it was not pure enough for crystal screening. At this point, the work being done with T4 59C42S and T4 (32 minus B), as well as DNA purification work, was progressing so the (32 minus A) fractions 16-23 from Poros HS were pooled, concentrated and stored at -80°C.

In the future, it would be a good idea to do a solubility screen on (32 minus A) to optimize the chromatography buffers used in the purification process. Tris buffer was used in the lysis buffer but BisTris buffer was used in the chromatography buffers. Since the buffer used was not consistent throughout the entire purification protocol, this could have had an effect on the level of protein purity at the end of the protocol. Once the solubility screen has been done and the chromatography buffers optimized, perhaps the (32 minus A) purity would
increase after purification. Once concentrated, that pure protein could be used for initial crystal screens.
Chapter 6

KVP40 59 protein

Most of the work discussed in this thesis has been focused on replication and repair proteins from the Bacteriophage T4 system. However, this is not the only phage that we study. We recently began studying similar replication and repair proteins from another phage, Bacteriophage KVP40. KVP40 is a T-even phage most distance from bacteriophage T4 which encodes similar replication and repair proteins. By studying this new system, we are able to learn about the proteins encoded by this new phage and gain insight into how proteins from T4 operate in comparison to those from KVP40.

In this chapter, I will introduce the helicase assembly protein from KVP40, which is called KVP40 59 protein. I will discuss the work that has been done to develop an expression and purification protocol as well as the more recent work that has been done with initial crystal screening of this protein. Finally, I will briefly discuss where this project will be headed in the near future.

6.1 Background

As stated earlier, we mainly study DNA replication and repair proteins from bacteriophage T4. However, in order to learn more about the DNA
replication proteins, we need to explore similar protein(s) from different phage since some T4 proteins do not crystallize. Bacteriophage KVP40 is a distantly-related T4-like phage (Miller, Heidelberg et al. 2003). It infects eight Vibrio species, including *Vibrio cholerae*. The KVP40 genome was just recently published and it was discovered that there are regions of the genome that show “extensive conservation” to bacteriophage T4 (Miller, Heidelberg et al. 2003). This genome is considerably larger than that of T4 at 244,835 base pairs and all ten replication and repair proteins encoded by T4 are also encoded by bacteriophage KVP40.

We are interested in learning more about the 59 helicase assembly protein from bacteriophage KVP40. Not much is known about this protein but it does share similarities with T4 gene 59 protein. Figure 6.1 below shows the sequence alignment between T4 gene 59 protein and KVP40 59 protein. There are many areas of high conservation (highlighted in yellow) and other areas that show great similarity (shown with “+” signs).

```
| T4     | 1 MIKLRMPAGGERY--IDGKSYKLYLMIKQHRNGKYDVIKYNWCRVSDAAYQKR 55 | ++G SV+++Y +K HM+GKYD+ KY M+ +A++KR |
| KVP40  | 1 MKRLLEVDVLPLRTWNGYSFRVYCAVKAHSGKYDISKYKLMKPTPSAFERK 55 |
| T4     | 56 RDKYFQKLSEKYKELALIFISNLVANQDAWIGIDSADALVFEYEGYSTRKQ 110 | ++D KF++L+++ L+ ++SNL AN +A +I+ ADA F+ ++ G L+ |
| KVP40  | 56 SDKVFFERLAKRLETNDCYQLLVSNLAAENGPSFYAEAGADAHEFNLKHTGFEI 110 |
| T4     | 111 IKFKFEIDRNTYFFSKVEVSAFEKIF--EYNPKQOSYIFKLLQNIISPETF 165 | ++ ++I+ K E FK++I E +F ++ I L+ N IS ETF |
| KVP40  | 111 YSQQYKNEISIPTL--ENKKFKDLERGEGHPL-----VIMQLILRTISIETF 165 |
| T4     | 166 ILLD5INIDKHE-QtDIKLNWNYSIKLKLAYRKLHNIDSQKAKNVPIETVKSCKY 222 | ++L+ LN + + D + +D++ W+ + + AY K+ID + AK ++III+ K |
| KVP40  | 166 VIENRLMNFPVIDKTSPIFWEFKTRALAYDKLKLIDEELAKELFQVQKNQSKI 222 |
```

Figure 6.1: Sequence alignment of T4 gp 59 protein and KVP40 59 protein. Conserved amino acid residues are highlighted in yellow.
KVP40 59 is a 25 kDa helicase assembly protein that is 216 amino acids in length. The theoretical pI is 9.33, which is very similar to that of T4 gene 59 protein (9.37). Hopefully by studying this protein and how it interacts with DNA, we will not only learn about KVP40 59 protein but can gain some insights into the T4 native protein and its ability to bind to forked DNA substrates.

The work that will be discussed here was focused mainly on the helicase assembly protein (KVP40 59 protein) with some work being done on a complex between the helicase assembly protein and the helicase (KVP40 41 protein). The ultimate goal of this project is the same as the goal for all of our projects: crystallization and structure determination. In chapter 4, I mentioned that T4 gene 59 protein precipitates in the presence of DNA so we wanted to explore the interaction between KVP40 gene 59 protein and fork DNA. Perhaps it will remain soluble, since it is such a distant relative. We also know that T4 gene 59 and T4 gene 41 form a complex and we are interested in co-crystallizing the same two proteins from KVP40.

6.2 Experiments

The work outlined in this chapter was focused on developing a purification protocol for the 59 protein. Initial crystal screens of KVP40 59, KVP40 59 + fork DNA and KVP40 59 + KVP40 41 were also attempted. We were fortunate enough to receive *E.coli* glycerol stocks of KVP40 59 from Nancy Nossal at NIH. Our work on this protein started with large scale protein expression.
6.2.1 Protein expression

The first step to take with this protein was to do a large scale expression. This protein was expressed in the same host as the T4 gene 59 protein (BL32 (DE3) pLysS cells) so the same antibiotics were used with the KVP40 59 protein as were used with T4 gene 59 (ampicillin and chloramphenicol). The entire protein expression protocol used with KVP40 59 is the same as that used for T4 gene 59 (refer to Figure 4.7). Shown below in Figure 6.2 is an SDS-PAGE gel showing protein expression of KVP40 59.

![SDS-PAGE gel showing expression of KVP40 59 protein. Lane 1 is the molecular weight marker, lane 2 the 0-hr sample and lanes 3-5 are the 3-hr samples.](image)

Good expression was seen at the predicted molecular weight. The cells were harvested from this large scale expression and then used for protein purification.

6.2.2 Protein purification

A protocol had been developed for both native T4 gene 59 and the protein mutant, 59C42S. Since KVP40 59 protein is the native form of the same protein from T4, we started with the protocol for the T4 gene 59 protein and the first step of that is a low-salt lysis. The low-salt lysis buffer and protocol that were used with the C215S mutant (refer to Figure 4.4) were used with KVP40 59. The SDS-PAGE gel, shown below in Figure 6.3, shows the results of the low-salt lysis.
Now that the protein had been extracted, further purification experiments were needed. Again, ion-exchange chromatography was used to further purify KVP40 59 protein. Since a protocol for the T4 gene 59 protein had previously been developed, we followed that protocol for this protein as well (refer to Figure 4.11). The conductivity of the sample was adjusted to be at or slightly below that of Buffer A and the sample was then loaded onto SP Sepharose. The SP Sepharose chromatogram is shown below in Figure 6.4.

Although lane 3 (the cell lysate) is slightly smeared, we were able to determine that the protein had been extracted from the pellet. A small amount of protein does remain in the pellet, but nearly all of the protein was found in the lysate.

Figure 6.3: SDS-PAGE gel showing results of low-salt lysis for KVP40 59 protein. Lane 1 is the molecular weight marker, lane 2 the cell pellet and lane 3 is the cell lysate.
Figure 6.4: SP Sepharose chromatogram for KVP40 59 protein. Fractions labeled with an “x” were run on an SDS-PAGE gel.
The protein began to elute from the column using a salt gradient in Fraction 14, peaking in Fraction 20. KVP40 59 protein was almost completely pure after this first column and this is shown below in Figure 6.5.

Figure 6.5: SDS-PAGE gel showing KVP40 59 protein samples from SP Sepharose. Lane 1 is the molecular weight marker and lanes 2-4 and Fractions 14, 16 and 20, respectively.

KVP40 59 protein was found at the predicted molecular weight. Fractions 17-25 were pooled, the conductivity adjusted and the sample was loaded onto Poros HS. The chromatogram from Poros HS is shown below in Figure 6.6.
Figure 6.6: Poros HS chromatogram for KVP40 59 protein.
A nice, sharp peak was seen when this protein eluted from the column, starting in fraction 39 and peaking in fraction 43. An SDS-PAGE gel was not run immediately to check the purity of the protein. However, the purity of the sample was checked after the protein had been concentrated. An SDS-PAGE gel displaying the pure sample is shown below in Figure 6.7.

By immediate glance at this gel, one would think that the protein was not pure when compared to Figure 6.4. However, this gel was overloaded with concentrated protein, which is why it appears to be less pure. If the gel had been run immediately following Poros HS with more dilute samples, the bands that are seen above would not have been seen. The protein would have appeared to be more pure than it does in the gel shown above.

6.2.3 Nuclease activity

Once the KVP40 59 protein fractions from Poros HS had been pooled and concentrated, a nuclease activity gel was run. This gel was run for the T4 59C42S protein as well to determine whether or not the contaminating nuclease had been removed from the sample in the purification process. The same needed to be
done with this protein. The samples were prepared in the same fashion and the gel was run at the same voltage and for the allotted time. The agarose gel is shown below in Figure 6.8.

![Figure 6.8: Nuclease activity gel for KVP40 59. Lane 1 is λ DNA, lane 2 is λ DNA + KVP40 59, lane 3 is Pet 28 and lane 4 is Pet 28 + KVP40 59. The arrows indicate possible degradation products.](image)

The same stop buffer was used in this experiment as had been used with T4 59C42S. It was determined from this gel that a trace amount of nuclease activity may have been present in this protein sample. There are two suspicious bands towards the bottom of the gel in both lanes containing protein that could be degradation products. If there was a large amount of nuclease in the sample, the DNA would be completely degraded and the DNA bands present in lanes 1 and 3 would not be present in lanes 2 and 4. The degradation product band would be the only band present in lanes 2 and 4. However, since the DNA bands were present, I decided to go ahead with the crystallization studies. In hindsight, the same strategy that should have been used with T4 59C42S protein should have been used here as well. EDTA should have been added to the dialysis buffer, which would have chelated the metals in the nuclease that may have been present.
and greatly diminished the nuclease activity. If the nuclease activity had been diminished, the initial crystal screens may have been more of a success.

### 6.2.4 Solubility screen

KVP40 59 was dialyzed overnight in the buffer that had been used for the T4 gene 59 protein, which was 50 mM NH$_4$Cl, 50 mM BisTris HCl, pH 6.5 and 10 mM MgCl$_2$. When the protein precipitated and came out of solution in the overnight dialysis, a solubility screen was run on the precipitated protein. The solubility screen was developed in the Mueser lab as a way to increase the solubility of a protein under study (Collins, Tomanicek et al. 2004). Once optimized, the protein is in an environment that satisfies the parameters required for crystal formation.

KVP40 59 protein was subjected to a 16-condition solubility screen, which consists of various salts and buffers. Twelve chloride and sodium salts are used to test various cations and anions while four buffers are used to test the pH dependency of the protein. The sixteen solutions used in the screen are shown below in Table 6.1.

<table>
<thead>
<tr>
<th>Chloride salts</th>
<th>Sodium Salts</th>
<th>Buffers</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH$_4$Cl</td>
<td>NaFormate</td>
<td>MES pH 5.6</td>
</tr>
<tr>
<td>NaCl</td>
<td>NaAcetate</td>
<td>PIPES pH 6.5</td>
</tr>
<tr>
<td>KCl</td>
<td>NaCacodylate</td>
<td>HEPES pH 7.5</td>
</tr>
<tr>
<td>LiCl</td>
<td>Na$_2$SO$_4$</td>
<td>TAPS pH 8.5</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>Na$_3$PO$_4$</td>
<td>Distilled H$_2$O</td>
</tr>
<tr>
<td>CaCl$_2$</td>
<td>Na$_2$Citrate</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: 100 mM salts and buffers used for 16-condition solubility screen (Collins, Tomanicek et al. 2004).
The procedure for the solubility screen is shown below in Figure 6.9. Usually, the protein is dialyzed overnight against distilled water but since KVP40 59 had precipitated overnight in the first dialysis buffer, dialyzing against water was not necessary. Water is used as a dialysis solution because most proteins will precipitate if dialyzed against water. This allows for the precipitated protein to be used in the screen.

1. Dialyze 5-10 mg of protein overnight against distilled water.
2. Protein will have precipitated; suspend the precipitated protein and aliquot into 17-1.5 mL Eppendorf tubes.
3. Centrifuge at max speed for 2-5 min to pellet the protein. Remove the supernatant from each tube.
4. Add 5 µL of 100 mM salt and buffer to each respective tube and resuspend the pellet. Let tubes sit at room temperature for 10 min.
5. After ten minutes, centrifuge the tubes to pellet the non-dissolved protein.
6. Use 2 µL of the contents of each tube and add 1 mL of Bradford reagent to each tube and immediately measure the absorbance at 595 nm.

The BIO-RAD reagent is used to determine protein concentration. When the absorbance is taken at 595 nm, an accurate reading of the amount of protein that has redissolved in each buffer or salt is obtained. The buffers and/or salts with the highest amounts of resolubilized protein can then be used to optimize the solvent.
The amounts of resolubilized protein are graphed and the results of the screen for KVP40 59 are shown below in Figure 6.10.

Figure 6.10: Results from protein solubility screen done on KVP40 59 protein. The absorbance is shown on the x-axis and the salts and buffers and shown on the y-axis.

KVP40 59 protein appeared to have a high solubility in a few of the chloride salts, namely MgCl\(_2\), NH\(_4\)Cl and CaCl\(_2\). The sodium salts that gave a high amount of redissolved protein were NaCitrate and Na\(_2\)SO\(_4\). The best way to put these results to use is to do a cross-coupled experiment and try multiple conditions to find an optimized solvent.

After the solubility screen results had been analyzed, the protein was dialyzed in 100 mM NH\(_4\)Cl, 50 mM BisTris HCl, pH 6.5 and 10 mM MgCl\(_2\). Although BisTris was not used in the solubility screen, it was used successfully in
the chromatography buffers. The MgCl\textsubscript{2} is used at a low concentration (10 mM) to decrease the chance of getting magnesium crystals in the initial crystal screens. KVP40 59 precipitated again in this dialysis buffer but 50 mM NaCitrate was added once precipitation occurred and nearly all of protein went back into solution. The final optimized buffer was 50 mM NaCitrate, 50 mM BisTris HCl, pH 6.5, 100 mM NH\textsubscript{4}Cl and 10 mM MgCl\textsubscript{2}.

6.2.5 Initial crystal screening

Although the KVP40 59 protein was not completely in solution after dialysis, the initial crystal screens were still set up. The screens used were the Natrix screen, PEG/Ion screen, Wizard I and Cryo I. One set of screens was set up at room temperature and one set was set up at 4 °C with the protein at 8.53 mg/mL. A number of hits from these initial screens were recorded and a few of the better ones are shown below in Figure 6.11.

![Figure 6.11: Hits from initial crystal screens of KVP 40 59 protein. a. Ion screen (4 °C) b. Natrix screen (RT) c. Natrix screen (4 °C).](image)

The hits were dyed with Izit stain (Hampton Research), which is a blue dye that is absorbed by the crystal if it is protein. These crystals absorbed the Izit stain, so it was assumed they were protein.
The crystal hit shown in Figure 6.11a appeared to be the most promising of all the hits recorded in the trays so the well conditions were then expanded. The well conditions from the screen were 20% PEG 4000, 100 mM Na MES pH 5.6, and 200 mM LiCl. There were three components of the tray that could be varied across a gradient in an expansion tray but we chose to vary the concentration of PEG and use four different buffers with a pH close to 7. An expansion tray was set up at 4 °C, using a 4x6 configuration (4 experiments of 6 gradients in one tray) with 14-24% PEG 4000 across each gradient and 100 mM buffer in each experiment, using four buffers. The buffers used were Na MES pH 5.6, Na Cacodylate pH 6.5, Na PIPES pH 6.5 and Na HEPES pH 7.5. The protein concentration was 19.3 mg/mL, which is more than twice that of the concentration in the initial screens (8.53 mg/mL). Unfortunately, all of the drops precipitated in this initial expansion tray and no crystal formation was seen. Another expansion tray was set up with the same conditions but the protein concentration was the same as that used in the screens (8.53 mg/mL) and again, all of the drops in the tray precipitated.

When the first expansion tray did not display and crystal growth, the conditions of a few of the other hits were examined. When the buffer was varied in the first expansion tray, no crystal growth was seen. For the next expansion tray, we wanted to vary a different component, like PEG or salt concentration. Upon examination of a few other hits from the screens, we chose to expand on the conditions of three hits and would vary the PEG concentration for two hits and the (NH₄)₂SO₄ concentration for a third hit. Three experiments were set up in one
tray and the conditions for the first experiment were 0-15% PEG 4000, 50 mM NaMES, pH 6.0, 50 mM MgSO₄. The conditions for the second experiment were 5-20% PEG 8000, 50 mM NaCac, pH 6.5, 200 mM KCl and 100 mM Mg(OAc)₂. Finally, the conditions for the third experiment were 0.5-2.5 M (NH₄)₂SO₄, 50 mM NaCac, pH 6.5 and 10 mM MgSO₄. The tray was set up at room temperature at a protein concentration of 9.0 mg/mL. Again, no crystal growth was seen in this expansion tray and most of the drops precipitated.

At this point, the results from the initial crystal screens were examined again to determine whether or not the protein concentration in the screens was too low. It was found that 34% of all of the drops in the initial screens did not precipitate and were clear. This means that 66% of all of the drops either precipitated or had some sort of crystal formation. If less than 70% of the drops either precipitate or crystallize in a screen, then it can be said that the protein concentration was too low. However, 66% is not too far below 70% so instead of re-screening the protein, KVP40 59 was screened in the presence of fork DNA and also in the presence of KVP40 41 protein, which is the helicase.

6.2.6 KVP40 59 + fork DNA

The binary complexes of native T4 gene 59 protein and a mutated form of the protein, 59C42S, in the presence of a fork DNA substrate have been studied and it was found that both forms of the T4 protein precipitate in the presence of DNA. Since KVP40 59 is a distant relative of the T4 form of the protein, perhaps it will stay in solution in the presence of DNA.
The initial screens were set up at a protein concentration of 8 mg/mL (0.32 mM) and the same fork DNA substrate that had been used with the T4 59 protein (lead18/lag18) was used with the KVP40 59 protein. The DNA substrate was purified in-house (Chapter 3) and annealed. The final DNA concentration was 0.45 mM, which is slightly more concentrated than the protein. Two room temperature screens were set up in Greiner 96-well trays, using the Natrix and PEG/Ion screens in one tray and Crystal screens I and II in the second tray. The trays were set up using the Honeybee sitting-drop crystallization robot.

The trays were checked and scored three times within two weeks of set-up and there was not any crystal growth seen. Within the first three days, 99.5% of the drops precipitated. The protein/DNA complex solution used to set up the trays was slightly cloudy before the trays were even set up so this probably did contribute to the precipitation of the drops. The cloudiness suggests the complex was not entirely stable. This was the first time crystal screens of this complex had been set up so more trays could be set up using different screening kits. The co-complex should be in solution before setting up the trays and the complex needs to be stabilized. This could be accomplished by adding different cryoprotectants to the solution, such as glucose, glycerol, sucrose or MPD. Crystallizing proteins in the presence of DNA is not easy and chances are it will take a lot of troubleshooting before optimization of the growth conditions is achieved.
6.2.7 KVP40 59 + KVP40 41

The complex that forms between the helicase (41 protein) and helicase assembly protein (59 protein) is a key complex in the T4 replication and repair protein system. This complex also forms between the two proteins from bacteriophage KVP40 and since we had access to both proteins, we wanted to study this complex as well.

In our studies, we found that KVP40 41 helicase is not a very stable protein and precipitates when concentrated. Knowing this, we decided to dialyze the co-complex together in the same dialysis cassette, in the hopes that the helicase would concentrate and stabilize. The KVP40 41 helicase was partially precipitated to start but upon addition of the 59 protein and subsequent dialysis, it did not precipitate more. Both of these proteins had been run through the solubility screen so it was rather difficult to find many commonalities in the respective results of the screen. The dialysis buffer that was used was 100 mM NaOAc, 10 mM Mg(OAc)\(_2\) and 50 mM BisTris HCl, pH 6.5. The proteins were dialyzed overnight in this buffer and the trays were set up the following day.

Two trays were set up at room temperature and two at 4 °C. The same screens were used here as were used with the KVP40 59 + fork DNA complex. All four trays were set-up at room temperature using the Honeybee sitting-drop crystallization robot and the two cold trays were then placed at 4 °C. The trays were checked and scored three times within two weeks of set-up and no crystal growth was seen. In the room temperature trays, 33.9% of the drops precipitated and in the cold trays, only 9.9% of the drops precipitated, which averages to
21.9%. The 41 protein was slightly cloudy before it was added to the 59 protein, which tells us that the protein was not completely in solution. Although the solution did not precipitate more upon addition to 59 protein, the sample was still not completely clear after dialysis. The concentration of the 41 protein was very low prior to dialysis (1.5 mg/mL) and the complex probably did not concentrate enough during dialysis. This may explain why so many clear drops were seen in the screens. If the concentrations of the proteins are slightly higher to begin with, the complex may be more stable.
Chapter 7

Conclusions and Future Work

The goal of this master’s thesis was to learn protein chemistry and macromolecular crystallography techniques and use the techniques learned and the skills gained to work towards the characterization of DNA replication and repair proteins from Bacteriophage T4 and related phage. In my mind, this goal was achieved, as I was given the opportunity to work on multiple projects, not only expressing, purifying and crystallizing the protein(s) being studied but gaining information about the protein itself and the effect of this research in a larger context. I was able to master laboratory techniques that will be of use in future research positions.

It can be said that there was more success at the protein chemistry level with these projects than there was at the crystallization level. The expression and purification protocols for four proteins (T4 59C42S, T4 (32 minus) A, T4 (32 minus B) and KVP40 59) were developed as well as a purification protocol for single-stranded DNA, which was not an easy task and took many weeks to perfect. Although crystals of the various binary and ternary complexes were not obtained, this was the first time anyone in the Mueser lab had attempted to
crystallize these very large complexes. These are complexes that will take tremendous effort to successfully crystallize, which was known when the project began.

Crystals were successfully grown of T4 gene 59 in the presence of glucose and TCEP, which were determined to stabilize this protein. The structure of this new crystal form is currently being built and once refinements are complete, the process of modeling the glucose into the structure will begin. This will take some time but we are confident this will be successful. As a Master’s student I did not expect to have the opportunity to assist with the structure building process so I was very fortunate to have this opportunity to learn something that is typically reserved for PhD students.

In the future, the binary complexes of T4 59C42A + fork DNA, T4 59C42S + (32 minus B), KVP40 59 + fork DNA and KVP40 41 + KVP40 59 and the ternary complex of T4 59C42S + T4 (32 minus B) + fork DNA will be explored further. These are not easy complexes to crystallize and once crystallized will not be easy to solve. T4 gene 59 protein, native and mutant, is still not stable in the presence of fork DNA. If we can solubilize this protein in the presence of DNA, perhaps by using different DNA substrates and varying the protein concentration, crystallization of this complex may be easier. The same goes for the KVP40 41 + KVP40 59 binary complex. The concentration of the KVP40 41 protein needs to be higher when in co-complex and once the protein concentration process has been optimized, the co-complex may crystallize.
For the co-complexes with DNA, the substrates used can be varied by using fork DNA substrates with lagging strand template arms of varying lengths. We know T4 gene 59 protein and T4 gene 32 protein bind a certain number of bases and only two DNA substrates were used in the binary and ternary complex screens. In the future, the screens could be set up using more than one substrate. This may help in optimization of the complex.
References


Appendices

A. Appendix I: List of Abbreviations

B. Appendix II: Rapid preparation of custom grid screens for crystal growth optimization
Appendix I: List of Abbreviations

1. AEBSF – 4-(2-Aminoethyl)-benzenesulfonylfluoride
2. AMP - ampicillin
3. BIO-CAD –
4. CAM - chloramphenicol
5. CCD – charge-coupled device
6. CFE – cell free extract
7. DB – dialysis buffer
8. DNA – deoxyribonucleic acid
9. DTT – dithiothreitol
10. EDTA – Ethylene diamine tetra-acetic acid
11. GP – gene product
12. HEL – hen-egg white lysosome
13. HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
14. HPLC – high performance liquid chromatography
15. IPTG – Isopropyl-beta-D-thiogalactoside
16. KAN - kanamycin
17. LB – luria broth
18. MES – 2-(4-morpholino)ethanesulfonic acid
19. MR – molecular replacement
20. OD$_{260}$ – optical density at 260 nm
21. OD$_{280}$ – optical density at 280 nm
22. OD$_{600}$ – optical density at 600 nm
23. PEG – polyethylene glycol
24. PIPES – 1,4 – piperazinebis(ethanesulfonic) acid
25. RNA – ribonucleic acid
26. SDS-PAGE – sodium dodecyl sulphate-polyacrylamide gel electrophoresis
27. TAPS – N-(tris(hydroxymethyl)methyl)-3-aminopropanesulfonic acid
28. TBE – Tris, Boric acid, EDTA
29. TCEP – Tris [2-carboxyethyl] phosphine
30. TE – Tris, EDTA
31. UV-Vis – ultraviolet-visible
Appendix II

Rapid preparation of custom grid screens for crystal growth optimization.

Anne B. Senger and Timothy C. Mueser

Macromolecular Crystallography Laboratory, 4211 Wolfe Hall, Department of Chemistry, University of Toledo, Toledo, Ohio 43606, USA. E-mail: Timothy.Mueser@utoledo.edu

Synopsis

A method for the rapid preparation of custom grid screen crystallization trays using standardized pipetting maps is presented.

Abstract

Initial crystallization conditions, typically discovered using sparse matrix screens, are refined using expansion trays - crystal setups of coarse and shallow gradients which vary single chemical parameters around the initial conditions. These custom grid screens are used to optimize crystallization conditions for the production of diffraction quality crystals. The trays are typically formulated by pipetting stock solutions into individual wells which requires numerous volume calculations. This tedious process is plagued by pipetting errors due differences in viscosity and small volumes with large dilutions as well as other problems such as evaporation and poor mixing. Preparation of two solutions, the initial and final conditions, allows for the use of standardized pipetting maps, a process we refer to as A/B gradients. Step-gradients can be prepared rapidly by adding decreasing amounts of the initial (A) solution to consecutive wells followed by the addition of the final (B) solution with increasing volumes. This simple idea can be applied to both coarse and shallow grids where the pipetting errors are confined within the boundaries defined by the A and B solutions. Programmable electronic pipettes can be used to rapidly prepare A/B gradients and are inexpensive alternatives to robotic liquid handlers.

Keywords: Crystallization, custom grid screen, pipetting maps.

1. Introduction

As we all are well aware, protein crystallography has seen major advances in recent years. Data collection and processing is now completed in hours rather than days, and model building is completed in days rather than months. Genes are now commercially available and kit based
cloning technology with affinity-labelling has simplified the preparation of crystallization targets. The last labor-intensive step remains in the preparation of single diffraction quality crystals.

A tremendous effort has also been focused on the creation of initial crystallization screens and the use of robotics to amplify total throughput. The modern protein crystallography laboratory typically has a set of favorite crystallization kits based on a Monte Carlo style sparse matrix screen (Jancarik and Kim 1991), an incomplete factorial approach (Carter and Carter 1979), or perhaps a maximum-likelihood method (Rupp 2003) to investigate the chemical parameters of crystallization. There has also been a significant amount of focus on the plates used for the initial crystallization trials. The higher density SBS plates, such as the Corning-96 and Greiner plates, have become the new standards. Scoring these kit-based crystallization trials requires many hours sitting at the microscope, pen in hand, identifying potential hits with the omnipresent concern of mineral crystals. Although perhaps not the most cost-effective choice for a small laboratory, crystallization robotic systems, such as a sitting-drop crystallization robot and automated crystal imaging system, alleviate many of the labor-intensive steps involved in setting up initial crystal screens.

A final step remains before diffraction studies can begin - the promising crystallization hits need to be optimized. One needs to be able to reproduce the conditions that invoke crystallization and then produce the large single crystals. In our experience, hanging-drop vapor diffusion is still the best method for preparation of expansion trays and production crystals. Expansions involve a grid screen search around the initial conditions, typically a step-gradient of the chemical components, perhaps varying the concentration of the precipitating agent in a row while varying the salt or buffer components in the column. The preparation of these "expansion" trays involves calculating dilutions from stock solutions and the pipetting of individual wells with the required components. Accumulation of pipetting errors is fairly severe especially with highly viscous stock solutions.

One of the first robotic systems used to prepare hanging-drop crystal trays was the Cyberlab C200, (Cyberlab, Brookfield, CT, now a subsidiary of Gilson). While beta-testing one of the first Cyberlab robots, it was clear that the systematic addition of individual stock solutions was very time consuming, involving many steps, a multitude of pipette tips, and was complicated by differences in viscosity. In addition, a separate mixing step was required to prepare homogeneous well mixtures. With concentrated PEG solutions that settled to the bottom of the well, the mixing
became even more problematic. The liquid handler improved accuracy and endurance but the speed did not increase substantially over the rate of manual preparation. To address the speed of preparation and viscosity problems, a simple answer became obvious. Rather than prepare individual wells for each different crystal trial, a standardized pipetting map could be applied to all trays with the preparation of premixed initial and final conditions. A similar method has been described previously using Excel spreadsheets to calculate individual well volumes (Eisele, 1993). We describe here a simple protocol for the rapid preparation of expansion and production crystallization trays.

2. Experimental Methods.

The chemical composition of expansion trays is dependent on the conditions discerned from crystal trials. A coarse gradient grid screen can be used to observe the behavior of the protein with each component and a shallow gradient grid screen can then be used to optimize the growth of production crystals for diffraction experiments. We prepare expansion tray gradients by varying the amount of initial and final solutions added to the wells. We refer to this method as A/B gradients, with A as the initial condition and B as the final. Gradients are established by decreasing the amount of solution A added to subsequent wells while increasing B to compensate thus trivializing the effort required for the preparation of the well solutions. The viscosities of A and B are similar and mixing times are minimal.

2.1. Preparation of 24-well hanging drop expansion trays.

Standardized pipetting maps are presented in Table 1. The 24-well trays are easily subdivided into four rows with six steps in a gradient, the 4x6, where each row is a step gradient of one component. Each consecutive row can then be used to test alternate conditions, perhaps increasing salt or changing buffer composition. A 4x6 row can be prepared by decreasing the volume of A by 200 µL each step. Then, using the same pipette tip, B can be added in increasing volume. A 2x12 has steps of 100 µL and a 1x24 tray has steps of 40 µL. Mixing the entire tray momentarily on a rotary shaker provides homogenous well solutions ready for expansion tray setup. Production crystals can be prepared in a similar manner using shallow, multi-step gradients. The 24-well trays can be subdivided into two, 12-step gradients (2x12) or used in its entirety as a 1x24 (Table 1). This 1x24 format is also very useful for the preparation of pH gradients as some crystallization conditions are affected dramatically by small perturbations in pH. The volumes of solutions added do not change with different experiments. The well
solutions can be prepared quickly without the need for cumbersome calculations. A and B solutions can be prepared in advance and stored until needed.

2.2. Preparation of 48-well hanging drop and 96-well sitting drop expansion trays.

Standardized pipetting maps for the 48-well hanging-drop and 96-well sitting-drop format are presented in Table 2. The well volumes for the 48-well and 96-well formats are 0.5 ml and 0.1 ml, respectively. The 48-well format can obviously be subdivided a multitude of ways but 8 columns of 6 steps is quite convenient with even steps of 100 µL per well (Table 2A). The 96-well Corning-96 sitting-drop plate is also very useful in high density expansions but only with larger drop sizes, such as 1 µL + 1 µL. The simplest division of the 96-well tray is to divide each of the eight rows of 12 wells into two sections of 6 wells, resulting in 16 rows of 6 wells. The pipetting is then simplified into 20 µL incremental steps for each trial gradient (Table 2B).

2.3. Use of the Rainin EDP-plus pipette.

For most of us, the expense of a liquid handler to prepare custom grids is difficult to justify. The Rainin EDP-plus programmable pipettes (Rainin, Oakland, CA) have proved to be inexpensive alternatives. The Rainin EDP-plus pipettes have a programmable gradient feature which allows the storage of up to 12 variable steps. An entire tray can be prepared in a matter of minutes, avoiding the long set-up times which are plagued with problems of evaporation. Preparation of the 4x6 and 2x12 gradients in the 24-well and the 8x6 gradients in the 48-well trays can be performed in a single program of the 10 ml pipettor.

The 1x24 gradient is extremely useful for shallow gradients in the preparation of production crystals but does require two programming steps per tray. This large volume pipette is by no means as accurate as a smaller volume pipette. However, since the initial and final conditions are preformulated, the pipetting errors are confined within the boundaries of the A and B solutions. The entire gradient can be pipetted forward for A and backward for B using the same pipetting map and pipette tip. Gradients adapted to the 96-well format can be easily prepared using a 1 ml Rainin EDP-plus (Model EP-1000) programmable pipette as shown (Figures 1C and 1D).

3.0. Results and Discussion.

The use of A/B gradients has improved the speed and reproducibility of crystal setups. Coarse and shallow gradients can be prepared using the same pipetting map. Coarse gradients are
used to refine the optimal region for crystal growth and shallow gradients are used to prepare crystalline structures. A coarse 0% to 30% PEG 8000 gradient and a shallow 19% to 20% PEG 8000 in a 2x12 format (Figures 1A and 1B respectively) were both prepared using the Table 1B pipetting map. The preparation of premixed initial and final solutions places a boundary on pipetting errors reducing the variability in component concentrations.

When a protein has the propensity to crystallize, many positive crystallization conditions can be obtained from the initial screens. When we choose to expand a few of the best hits, we typically base our choice on aesthetics rather than diffraction quality. As many can attest, the best looking crystals may not always yield the best diffraction. It is useful to use a higher density expansion tray format and expand all potential hits. A single 96-well sitting-drop tray can be used to expand conditions of 16 hits in the proposed pipetting layout. Coarse and shallow grids (Figures 1C and 1D) were prepared following the 6x16 format (Table 2B).

When preparing crystals for diffraction studies, a shallow gradient can be set up in a 24-well format to step through the refined crystallization zone. Using the same A and B solutions, two trays set up on different days display different regions of quality growth. The shallow gradient is most likely compensating for variability in other parameters such as temperature and perhaps nucleation. Step gradients are the simplest way to define regions for crystal growth. This method can also be used to prepare other styles of gradients, such as a 1x24 pH gradient, with a wide or narrow search range.


We have used standard pipetting maps for a number of years and benefit from a substantial increase in the number of trails conducted. It is a simplistic and practical approach to the preparation of crystal trays. This technique reduces the time and effort required while also reducing the waste associated with multiple pipetting. Robotic liquid handlers are very useful but impractical in most academic laboratories. The use of the Rainin EDP pipette with a programmable variable volume option is an affordable alternative. For those laboratories with a liquid handler, the pipetting maps are easily adapted to all platforms and each new tray does not have to be reprogrammed, a definite plus for high throughput screening.

**Figure 1** Coarse and shallow step gradients can be prepared using the same pipetting map. To demonstrate the concentration gradient created with the pipetting maps, yellow dye was added to the A solution and blue dye added the B solution. 1A. A coarse gradient of 0% to 30% PEG 8000 was prepared in the 2x12 format with wells A1 and B6 as the initial and final conditions respectively. A Rainin 10 ml programmable
EDP pipette was used to add decreasing amount of solution A to consecutive wells. Using the same pipette tip, the incremental addition of decreasing amounts of solution B was added in the reverse direction. The tray was mixed for 15 seconds on an orbital shaker and shows complete mixing of the colours in the gradient. **1B.** A shallow gradient of 19% to 20% PEG 8000 was prepared as described in 1A. A similar color gradient is seen upon mixing. **1C.** A coarse gradient of 0% to 30% PEG 8000, was prepared in one row of the 6x16 format of the 96-well Corning 96 tray with A1 and B6 as the initial and final conditions, respectively. **1D.** A shallow gradient of 19% to 20% PEG 8000 was prepared as described in 1C.

Table 1  **Standard pipetting maps for 24-well trays.** Presented are the volumes added to consecutive wells to prepare the grid screen step-gradient. **A.** Three milliliters each of solutions A and B are required to prepare each gradient when the 24-well tray is divided into four rows of 6 wells; the 4x6 format. **A**
complete tray will require 4 A and 4 B solutions prepared separately. Solution A is added to consecutive wells, decreasing the volume at each step by 200 µL. Using the same tip, solution B can be added with increasing volume with the same step size. B. Division of the 24-well tray into two experiments, the 2x12 format, requires 6.6 mls each of A and B solution. Step sizes of 100 µL are used with 1.1 ml total volume. This format is used in the color gradients shown in Figure 1A and 1B. C. The entire 24-well tray, the 1x24 format, can be used for the final preparation of diffraction quality crystals. A shallow gradient across the entire ensures the crystallization zone has been covered.

A. 4x6 (1.0 ml well volume, 200 µL steps, rows)

<table>
<thead>
<tr>
<th>Well #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total volume (mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution A</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Solution B</td>
<td>0.0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

B. 2x12 (1.1 ml well volume, 100 µL steps, rows)

<table>
<thead>
<tr>
<th>Well #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total volume (mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution A</td>
<td>1.1</td>
<td>1.0</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>6.6</td>
</tr>
<tr>
<td>Solution B</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>6.6</td>
</tr>
</tbody>
</table>

| Well # | 7   | 8   | 9   | 10  | 11  | 12  | |
|--------|-----|-----|-----|-----|-----|-----||
| Solution A | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 | 0.0 | 6.6               |
| Solution B | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 | 1.1 | 6.6               |

C. 1x24 (1.0 ml well volume, 40 µL steps, rows)

<table>
<thead>
<tr>
<th>Well #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total volume (mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution A</td>
<td>0.96</td>
<td>0.92</td>
<td>0.88</td>
<td>0.84</td>
<td>0.8</td>
<td>0.76</td>
<td>12</td>
</tr>
<tr>
<td>Solution B</td>
<td>0.04</td>
<td>0.08</td>
<td>0.12</td>
<td>0.16</td>
<td>0.2</td>
<td>0.24</td>
<td>12</td>
</tr>
</tbody>
</table>

| Well # | 7   | 8   | 9   | 10  | 11  | 12  | |
|--------|-----|-----|-----|-----|-----|-----||
| Solution A | 0.72| 0.68| 0.64| 0.6  | 0.56 | 0.52 | 12               |
| Solution B | 0.28| 0.32| 0.36| 0.4  | 0.44 | 0.48 | 12               |

| Well # | 13  | 14  | 15  | 16  | 17  | 18  | |
|--------|-----|-----|-----|-----|-----|-----||
| Solution A | 0.48| 0.44| 0.4  | 0.36 | 0.32 | 0.28 | 12               |
| Solution B | 0.52| 0.56| 0.6  | 0.64 | 0.68 | 0.72 | 12               |

| Well # | 19  | 20  | 21  | 22  | 23  | 24  | |
|--------|-----|-----|-----|-----|-----|-----||
| Solution A | 0.24| 0.2  | 0.16| 0.12| 0.08| 0.04| 12               |
| Solution B | 0.76| 0.8  | 0.84| 0.88| 0.92| 0.96| 12               |
Table 2  Standard pipetting maps for 48 and 96-well trays. Presented are the volumes required to prepare the higher density format crystal screen trays. A. The 48-well hanging drop tray can be subdivided into eight, 6-well gradients. The setup requires 1.5 mls each of solutions A and B for each experiment with 0.5 ml total volume per well. B. The 96-well sitting drop format, normally used for initial screens, can be easily adapted to expansion trials. The tray can be subdivided into sixteen, 6-well gradients. The setup requires 0.3 mls each of solutions A and B for each experiment with a total volume of 0.1 ml per well. This tray formulation is highly conducive for adaptation to robotic liquid handlers since it does not require reprogramming for each experiment.

A. 48-well 8x6 (0.5 ml well volume, 100 µL steps, columns)

<table>
<thead>
<tr>
<th>Well #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total volume (mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution A</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Solution B</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

B. 96-well 16x6 (0.1 ml well volume, 20 µL steps, rows)

<table>
<thead>
<tr>
<th>Well #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total volume (mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution A</td>
<td>0.1</td>
<td>0.08</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Solution B</td>
<td>0.0</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Acknowledgements  The authors wish to thank Aude Izaac, Juliette Devos, Steve Tomanicek, and Bob Cudney for proofreading this manuscript and for helpful suggestions.

References


