

VIRAL SELF COMPONENTS  
OF THE *TECTIVIRIDAE*  
FAMILY

By

JULIA MEGAN BESHIRS

Bachelor of Science in Chemistry

Southeastern Oklahoma State University

Durant, OK

2001

Submitted to the Faculty of the  
Graduate College of the  
Oklahoma State University  
in partial fulfillment of  
the requirements for  
the Degree of  
MASTER OF SCIENCE  
July 2007

VIRAL SELF COMPONENTS  
OF THE *TECTIVIRIDAE*  
FAMILY

Thesis Approved:

Dr. Stacy Benson  
\_\_\_\_\_  
Advisor

Dr. Nicholas Materer  
\_\_\_\_\_  
Committee Member

Dr. Ziad El Rassi  
\_\_\_\_\_  
Committee Member

Dr. A. Gordon Emslie  
\_\_\_\_\_  
Dean of the Graduate College

## ACKNOWLEDGEMENTS

I have many people who have helped me get through my time here in graduate school. I would like to extend a grateful thank you to all who have offered help and advice during my time here. There are those few who deserve a special recognition of their role in my journey.

I would like to thank Amanda Nichols for being such a wonderful friend to me. While you may not have physically helped me carry out experiments, you offered much needed laughter and a listening ear. Thank you for all of your advice, both personal and professional. I would like to thank the other members of my laboratory, both the graduate and undergraduate students. Thanks for being such a great group to work with and for helping me out with my projects every now and then. I am incredibly grateful to have worked under such an understanding and knowledgeable advisor, Dr. Stacy Benson. I enjoyed working for your lab and learning from you. Thank you for allowing me to grow and explore new areas on my own, without being overshadowed. I also appreciate you listening to me about my personal matters and giving me counsel on matters involving my degree. You will continue to build a productive and thriving laboratory by being a well rounded leader who cares for his student's lives, as well as their educational pursuits. I would like to extend a thank you to my committee for helping me to obtain my degree.

I also need to express gratitude for those people in my personal life that have helped me get to this point. Not enough thanks could ever be said to my fiancé, Joe. However, I would like to express my appreciation to you for being so supportive of my dreams and goals. Thank you for being so understanding and enthusiastic towards my other life in the lab. I also need to send an acknowledgement to Joe's family, my new found family, for being so incredibly supportive of me. I appreciate the help that you all give. Most importantly, I would not have made it to this point in my life if it had not been for my wonderful, loving mother, Nona. You always impressed upon me how important education was and equipped me with the tools to succeed in that arena. You were always my biggest cheerleader, with you I felt that I could do anything. Though you are no longer with me in body, your spirit continues to live on in me and I know that I have and will continue to make you proud, just like you always told me I did.

## TABLE OF CONTENTS

Chapter	Page
1. LITERATURE REVIEW / BACKGROUND .....	1
2. Bam35 P3	
2.1 INTRODUCTION .....	13
2.2 RESEARCH DESIGN .....	18
2.3 RESULTS AND DISCUSSION .....	27
3. PRD1 P6	
3.1 INTRODUCTION .....	55
3.2 RESEARCH DESIGN .....	59
3.3 RESULTS AND DISCUSSION .....	65
4. CONCLUSION .....	68
REFERENCES .....	71

## LIST OF FIGURES

Figure	Page
1-1. Schematic of PRD1 virion .....	7
1-2. Tree of Life .....	11
2-1. Genome organization of Bam35c and PRD1 .....	15
2-2. Jelly roll fold .....	16
2-3. Single subunit of PRD1 P3 .....	17
2-4. P3 structure of PRD1 and model of Bam35 .....	18
2-5. Full chromatogram of P3 run through the 5 mL GE Healthcare Hi-Trap SP column. ....	28
2-6. Glucose comparison chromatogram overlay .....	29
2-7. SDS-PAGE gel analysis of P3 produced with no glucose and autoclaved media.....	30
2-8. SDS-PAGE gel analysis of P3 produced with no glucose and filtered media .....	31
2-9. SDS-PAGE gel analysis of P3 produced with glucose and autoclaved media.....	31
2-10. SDS-PAGE gel analysis of P3 produced with glucose and filtered media .....	32
2-11. SDS-PAGE gel analysis of P3 produced with no glucose and autoclaved and filtered media run on Bio-Rad DuoFlow .....	32
2-12. SDS-PAGE gel analysis of P3 produced with glucose and autoclaved and filtered media run on Bio-Rad DuoFlow .....	33
2-13. Chromatogram overlay of glucose/no glucose and autoclaved/filtered comparisons.....	33
2-14. SDS-PAGE gel reference for MALDI-TOF results.....	35

Figure	Page
2-15. MALDI-TOF results for band #1.....	35
2-16. MALDI-TOF results for band #2.....	36
2-17. MALDI-TOF results for band #3.....	36
2-18. MALDI-TOF results for band #4.....	37
2-19. MALDI-TOF results for band #5.....	37
2-20. SDS-PAGE gel analysis of P3 with protease inhibitors added.....	38
2-21. Chromatogram overlay of freeze/thaw isolation methods.....	39
2-22. Chromatogram overlay of the lysozyme isolation method.....	40
2-23. SDS-PAGE gel of P3 sonicated 10 times for 10 seconds in lysis buffer with PI added.....	41
2-24. Chromatogram overlay of sonication method (4x15).....	41
2-25. Chromatogram overlay of sonication method (10x10).....	42
2-26. Chromatogram overlay of sonication method (2x1).....	42
2-27. SDS-PAGE analysis of different boiling times at 60, 70, and 80°C.....	44
2-28. SDS-PAGE analysis of different boiling times at 90, 95, and 100°C.....	44
2-29. Crystal formed with 7 mg/mL protein concentration.....	46
2-30. Crystal formed with 10 mg/mL protein concentration.....	46
3-1. P6 sequence comparison for Bam35 and PRD1.....	58
3-2. SDS-PAGE gel analysis of PRD1 P6.....	66
3-3. SDS-PAGE analysis of P6 on the Hi-Trap Q column run on the Bio-Rad DuoFlow.....	67

## NOMENCLATURE

ASFV	African swine fever virus
bp	base pairs
BSA	bovine serum albumin
CIV	Chilo iridescent virus
DTT	dithiothrietol
ds	double stranded
EM	electron microscopy
EG	ethylene glycol
HSV-1	Herpes simplex virus 1
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	isopropyl-beta-D-thiogalactopyranoside
kDa	kilodaltons
kbp	kilobase pairs
LB	Luria Bertani
MALDI-TOF	Matrix assisted laser desorption/ionization – Time of flight
MW	molecular weight
MWCO	molecular weight cutoff
OD	optical density

PBCV-1	<i>Paramecium bursaria</i> chlorella virus 1
PMSF	phenylmethanesulphonylfluoride
PEG	polyethylene glycol
PI	protease inhibitors
ss	single stranded
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
STIV	Sulfolobus turreted icosahedral virus

## CHAPTER 1

### **LITERATURE REVIEW / BACKGROUND**

Viruses are relatively new organisms to us, only first being discovered in the late 19<sup>th</sup> century. Jacob Henle postulated in 1840 that there was a very small infectious agent that could not be seen with a microscope, unlike the protozoa, fungi, and bacteria which had been observed in this manner. His hypothesis would not be proven true for another 50-60 years. In 1892, Dimitri Ivanofsky passed extracts of tobacco mosaic infected leaves through the smallest filter known at the time, unglazed porcelain. He was trying to eliminate bacteria, which were bigger than the porcelain pores. However, he found that the filtrate still caused disease in new leaves. He thus determined that the filter must be defective or the disease was caused by a toxin. In 1898, Martinus Beijerinck, unaware of Ivanofsky's work, also did the same experiment, but came up with a different conclusion. He hypothesized that the disease was not caused by a toxin because it was able to grow and reproduce in the living plant cells. Thus, the idea of viruses had been established. The term filterable agents was used for classification of this new infectious component, but was gradually replaced with the term *virus* (Latin for "poison"). Bacteriophages, viruses that infect bacteria, were discovered by Frederick Twort (1915) and Felix d'Herelle (1917) and have since become important agents in obtaining knowledge about

viruses as a whole. This is due to the simplicity of growing the bacteria that they infect in the laboratory<sup>1, 2</sup>.

Even today, the complexity and number of bacteriophages, and other viral species, is just beginning to be appreciated. Studies have been performed to quantify the number of bacterial and viral populations in ocean waters. These findings have shown a bacterial concentration of  $10^6$  cells/mL and a viral concentration of  $10^7$  particles/mL. Globally, it has been estimated that at any one time there are  $10^{30}$  or more phage particles, thus surpassing available host organisms by at least one order of magnitude<sup>3, 4</sup>. This would suggest that all living cells are under a constant barrage of viral intrusions. There is evidence of this viral-host battle in the fact that the cellular organisms have been able to acclimate themselves to coexist with viruses and to support viral or virus-like genomes<sup>5</sup>. When using the global marine population numbers, calculations of gene transfer between organisms occur at a rate of 20 million billion times per second in the oceans. Actual numbers would be less due to smaller transduction efficiency and increased phage decay in the ocean versus the laboratory. However, even a fraction of this occurrence taking place still opens the door for an enormous amount of horizontal gene transfer from both the virus to the host and from the host to the virus<sup>3</sup>. Therefore, viruses need to be studied further to fully comprehend the effect that they have on cellular organisms.

The most popular molecular bacteriology species, *Escherichia coli* and *Bacillus subtilis*, are each hosts for about 10 phage species. Each of these phages varies widely in

genome organization, as well as DNA sequencing levels within each group. If it can be generalized that each bacterial species plays host to 10 or more phages, then viruses would certainly be the largest uncultivated reservoir of sequence information in the biosphere. A study was conducted to perform random sequencing of viral DNA obtained from two uncultured ocean water samples. The results, combined with statistical analysis, showed a presence of 400-7,000 different viral species, depending on sample variations, with the most abundant type representing just 3% of the total viral population<sup>3</sup>. This would indicate that our current knowledge has only begun to scratch the surface, so further work needs to be done to explore and understand the interesting world of viruses.

Present knowledge of viruses and viral infections has led to increased knowledge of cellular processes, cell biology, and genetics. They have proved important when exploited as tools in molecular biology. One area, in particular, is in protein production where bacteriophages are enormously efficient. Since viral programming is designed to carry genetic information into cells, they have been used to replace defective cellular genes, such as in gene therapy. A new field of research is occurring with genetically engineered viruses to aid in medical intervention and agricultural applications<sup>1</sup>. A more comprehensive understanding of phages, in both their processes and structure, can only lead to better utilization of this knowledge and the ability to enrich our medical, industrial, and agricultural institutions.

Viral families are organized by properties that they share with each other. These properties are defined by four different criteria. The first involves the nucleic acid contained within the virion, either DNA or RNA. The second incorporates the symmetry of the capsid, which can be described by three subcategories: helical, icosahedral, and complex. The third property is based on the presence or absence of a lipid membrane, or envelope. The final category is attributed to the dimensions of the virion and viral coat. The classification can then further be subdivided by applying knowledge of the genome architecture. The arrangements are divided into single stranded (ss) or double stranded (ds) RNA or DNA and whether the DNA is circular or linear<sup>2</sup>. Additional partitioning can be made via host preference. However, viruses are real, but the species designations are man-made<sup>6</sup>, and studies leading into this research is showing that these classifications might need to be revised. Bacteriophages, the main focus of this work, have representative members in all division types, making them a very diverse group.

Bacteriophages, from the Latin “eaters of bacteria”, range in size from 4 kilobase pairs (kbp) to 600 kbp. Tailed phages make up 96% of this group<sup>4</sup>. Currently, there are known to be 12 distinct groups of bacteriophages, which are very diverse structurally and genetically. Their mode of replication is similar to all other viruses. After attaching to a specific receptor on the bacterial cell wall, the phage genome enters the cell. The capsid proteins are usually left outside the cell by being stripped off in the entry process or the phage uses special viral components to inject the genome into the cell. Many of the tailed phages have a contractile sheath that is utilized like a syringe for genome injection. The "head and tail" morphology is unique to the tailed phages and is related to their mode

of cell penetration. Some "round" phages (icosahedral or lipid-coated) enter the bacterial cell by adhering to flagellae or pili on the host's cell wall and being drawn in to the cell by endocytosis. After entry, many phage genomes are degraded and destroyed. This is due to protective bacterial mechanisms which depend on the recognition and destruction of foreign DNA. Surviving phage genomes utilize the cellular apparatus for gene expression (transcription and translation) to various extents<sup>7</sup>.

Of all the known viruses, double-stranded DNA (dsDNA) tailed bacteriophages are the most common in the biosphere<sup>4</sup>. Therefore, these phages can be found in a vast array of environments and most likely infect the whole range of bacterial species. These viruses are classified as coliphage and can be regarded as being one super-organism, despite differences in morphological features, with the members exchanging genetic material globally through recombination<sup>8</sup>. Bacteriophage infections are estimated to occur at a rate of  $\sim 10^{25}$  globally per second<sup>9</sup>. This has led to the tailed bacteriophages having mosaic genomes<sup>8, 10, 11</sup>. Mosaicism occurs when one compares two genome sequences and finds regions of high sequence similarity with abrupt transitions to neighboring regions of little to no similarity<sup>11</sup>. This phenomenon points to the fact that viral evolution has been influenced by horizontal transfer, as well as impacting bacterial evolution through infection. Consequently, understanding more about viral evolution and origins will allow for more familiarity of viral processes as a whole.

This thesis involves icosahedral dsDNA bacteriophages with a lipid membrane and no tail, an individual property that segregates them from the tailed dsDNA

bacteriophages. More specifically, this study focuses on the family *Tectiviridae*, containing those isolates that infect Gram-negative hosts (PRD1, PR3, PR4, PR5, L17, and PR772) and Gram-positive hosts (Bam35, Ap50, and phi NS11). These isolates were originally placed together for morphological reasons<sup>5</sup>, even though their hosts diverged about two billion years ago<sup>12</sup>. Unlike the mosaicity of the tailed bacteriophages, PRD1 and all of the PRD1-like isolates are almost identical in their genomes, despite isolation at different dates and locations from around the world. Also, Bam35 and one of its isolates, GIL01, have proven to have essentially equivalent genome sequences<sup>5</sup>. This family is of interest due to the lipid membrane, which lies beneath their protein capsid. The membrane substitutes for the tail, seen in the majority of phages, by forming a tube with the membrane to generate a tail to allow for infection of its host. This aspect of their morphology is one reason they were investigated so as to study the protein-membrane interactions<sup>13</sup>. Other studies of interest focus on the viral structure. Determining the configuration of the virus as a whole can give insights into the assembly, and even disassembly, of the virus. Knowledge of the framework of the virus and its individual components can lead to further understanding of the functions of each; such as binding sites, host cell and viral component fusion, and packaging, among others. The uniqueness in structure of the *Tectiviridae* family makes it an interesting candidate for further studies into organization and biogenesis<sup>14</sup>.

PRD1 is the type species for this family and has been studied in detail. It is a lytic phage<sup>13, 15</sup>, which means that the cells that it infects are destroyed in the process, due to cell membrane disruption. Structures have been determined for its coat protein, receptor-

binding protein and virion<sup>14, 16-23</sup>. The virus consists of an icosahedral protein coat with receptor-binding spikes at its vertices. Underneath the capsid lies the membrane, which encloses a linear 14,935 bp genome<sup>5</sup> (see Figure 1-1). The capsid protein, known as P3, is a trimeric molecule. It is composed of two eight-stranded viral jelly rolls, or  $\beta$ -barrels, in each 394 residue subunit. The capsid measures 698 Å between opposing vertices. The organization of the coat proteins is in a pseudo T=25 arrangement with 240 P3 trimers that make up the 20 facets and a P31 pentamer located at each 5-fold vertex<sup>19, 21, 24</sup>. There are twelve total vertices on this phage, eleven regular vertices and one special packaging vertex. This packaging vertex functions as a channel through which the genome DNA is packaged into the empty particles<sup>25</sup>.

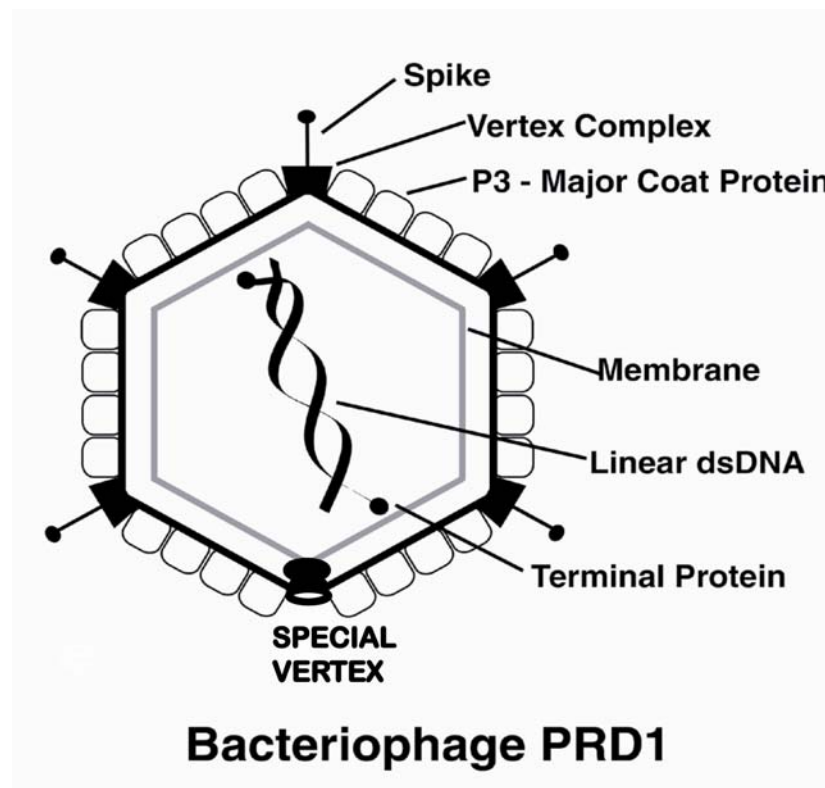


Figure 1-1. Schematic of the PRD1 virion.

Another important member of this family is Bam35 (14,935 bp). It has essentially an identical capsid organization as PRD1, as recently revealed by a 7.3 Å resolution image reconstruction from cryo-electron microscopy (cryo-EM) images<sup>26</sup>. Furthermore, its capsid protein, also denoted P3, is predicted to be similar to the coat of PRD1<sup>27</sup>. However, the two proteins share a limited amount of sequence similarity, despite a comparable protein size (356 residues for Bam35 P3 compared with 394 residues for PRD1 P3). The sequences of the P3 proteins share only 16% identity and 31% similarity. The sequence of one capsid protein will not retrieve the other when performing a PSI-BLAST<sup>28</sup> search. Nevertheless, 3D-PSSM<sup>29</sup>, a program that searches for structures congruent with the sequence and secondary structure estimates, has predicted with >95% confidence that the folds are analogous. CLUSTAL X<sup>30</sup> was used to align the two sequences and introduce gaps into the shorter Bam35 P3 sequence to optimize alignment. Due to the nominal sequence similarity, specific alignment was not possible. This alignment was evaluated by threading the Bam35 P3 sequence onto the structure of PRD1 P3. However, the sets of matching residues and their similar spacing were able to secure the alignment. The modeling showed that the Bam35 P3 sequence could adapt to the PRD1 P3 crystal structure.

The major focus of the resemblance between these two phages comes from the double barrel trimer motif in the major coat proteins. The exact composition of this structural aspect will be discussed in more detail in Chapter 2. The importance of this similarity is not just that it can be seen within these two species, but that structural studies have now found the double barrel trimer coat protein apparent in several other viruses,

with hosts spanning across the tree of life and forming a proposed lineage. *Paramecium bursaria* chlorella virus 1 (PBCV-1), a member of the family *Phycodnaviridae*, infects algae. The solved structure of its capsid protein, Vp54, has shown the presence of the double barrel trimer<sup>31</sup>. Recently, *Sulfolobus* turreted icosahedral virus (STIV) has also shown this same structural aspect. STIV infects an archaeobacterium that is known to grow optimally at pH 2-4 and >80°C<sup>32</sup>. Adenovirus was one of the first to show the double barrel trimer within its major capsid protein structure. It belongs to the family *Adenoviridae* and has mammalian and avian hosts<sup>33, 34</sup>. Models have been created for other viruses which were acquired with literature and PSI-BLAST searches. Chilo iridescent virus (CIV) is an insect infecting member of the *Iridoviridae* family, which infects vertebrates (fish, amphibian, and reptile) and invertebrates (insects). The threading of its coat protein, P50, onto the PBCV-1 Vp54 crystal structure shows the preservation of the double barrel trimer and major differences in the sequence alignments occurring in loop regions. The coat protein, P72, of African swine fever virus (ASFV), the sole member of *Asfarviridae*, is also consistent with the double-barrel framework and is a candidate for this proposed lineage. The recently discovered mimiviruses, infecting amoebae, and the ascoviruses, infecting moth and butterfly larva, have also shown modeling results that indicate that they too belong in this viral line<sup>13</sup>.

It is proposed that convergence of structure to perform the same function is highly improbable. Extremely similar structures with very different amino acid sequences is most likely the result of divergent, instead of convergent, evolution<sup>35, 36</sup>. This structural similarity, despite differences in hosts and genome sequences, points to an overall

divergence theory for PRD1 and Bam35. The divergence theory stems from the belief that viruses were present before the tree of life (see Figure 1-2) diverged into the branches that we observe today. There is a belief that viruses do not possess a single root, but instead different viral lineages span out and complement all branches that we currently observe on the cellular tree of life. The fact that viruses were present in primeval times means that they have been able to adapt to their hosts and influence the host's evolution.

If this theory is true, we need to be able to explain the conservation of structural aspects when all other features, such as genome sequence and host preference, show little to no similarities. This can be done with the concept of a viral “self”. The self preserves efficiency when interacting with hosts and yet conserves viral structures and functions<sup>37</sup>. It would contain determinants, transcended from the viral ancestor, which would encode for vital structural properties of the virion. These would be items such as architecture and even genome packaging, which would lead to phylogenetic relationships among the viruses. Characteristics such as host recognition and adaptation are not part of the viral self, instead they are most likely obtained through lateral gene transfer from the host, and can veil the true phylogeny of viruses<sup>13, 37</sup>.

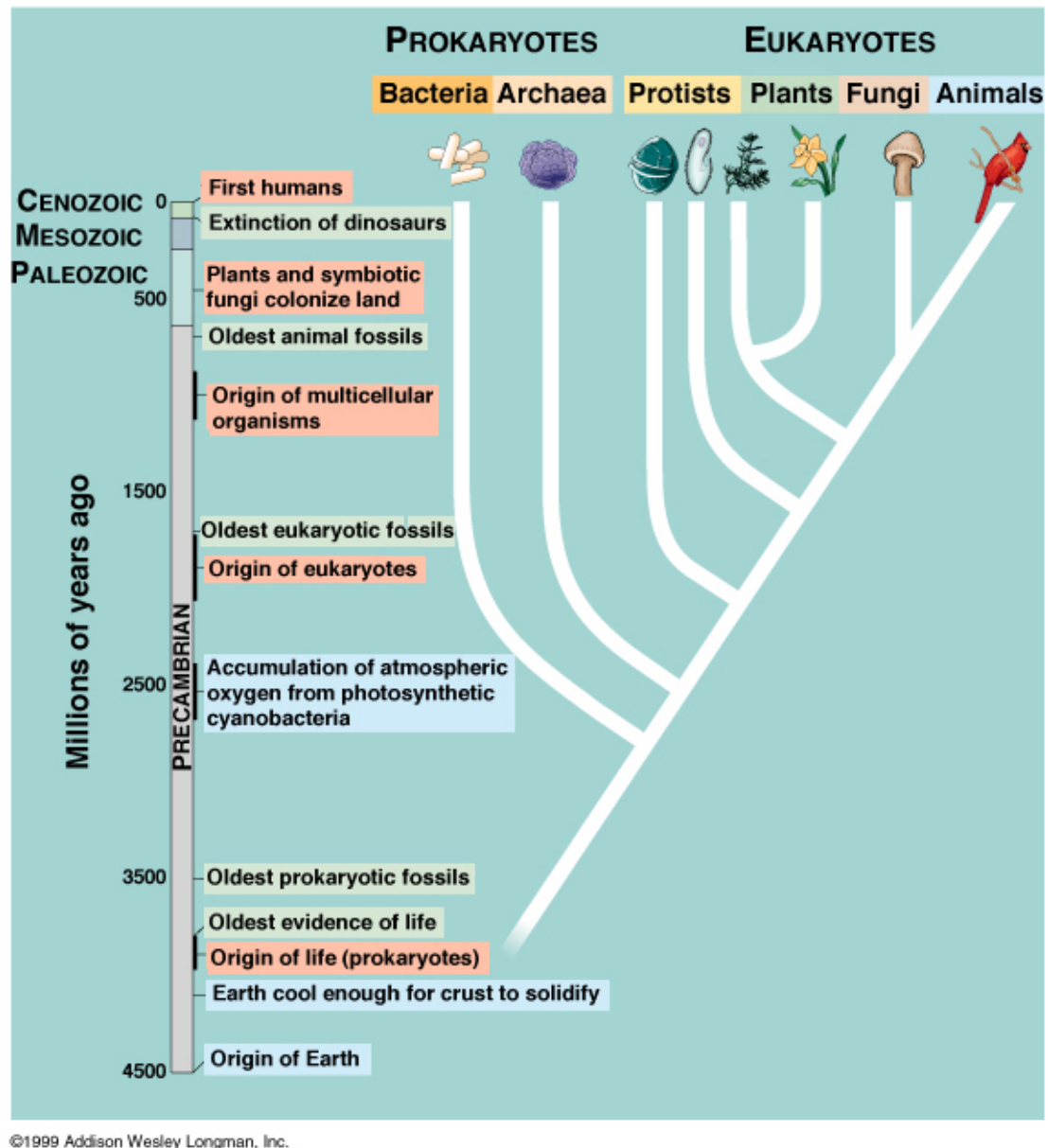


Figure 1-2. Tree of Life.

The main goal of this study is to find evidence to support the idea of the divergence theory. The first part of the study focuses on trying to show the true relationship of the two P3 coat proteins of the known PRD1 and the studied Bam35 (Chapter 2). The second part of the study focuses on a portal protein, located on the

special packaging vertex of PRD1, known as P6 (Chapter 3). The expectation here is to expand the idea of conserved viral capsid structures to other vital viral structures and functions. By adding additional structures of the major components of these dsDNA icosahedral viruses, we can further explain the similarities and differences between them and discover the characteristics that arose from the ancestral virus. Furthermore, these similarities can lead to a better understanding of how these viruses assemble and package their genomes. Perhaps it is possible to manipulate these “commonalities” to prevent the viruses from infecting their hosts and, thus, prevent disease.

## CHAPTER 2

### **Bam35 P3**

#### **2.1 Introduction**

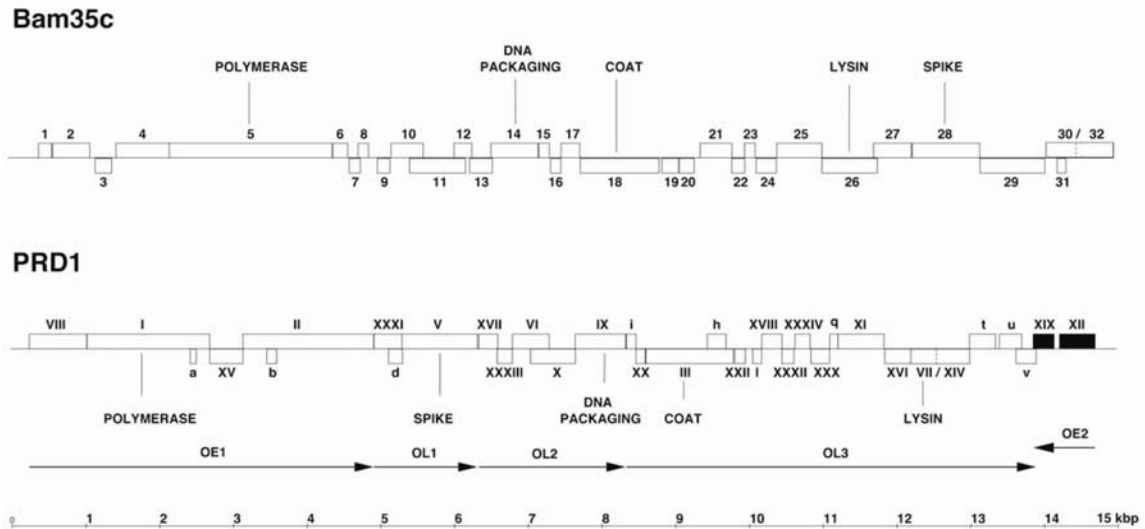
Bacteriophage Bam35, which infects Gram-positive bacteria, is a member of the *Tectiviridae* family. It is specific for *Bacillus thuringiensis*, a soil bacterium. The genome consists of linear dsDNA, which has 14,935 bp<sup>13, 27</sup> with 5'-covalently linked terminal proteins. It is a temperate phage, which means that it can either lyse all of the infected bacterial cells or remain as a linear dsDNA plasmid in the bacterium and replicated along with its host<sup>38</sup>.

Much is known genetically, biochemically, and structurally about the type species of the *Tectiviridae* family, bacteriophage PRD1, as detailed in Chapter 1<sup>5, 39</sup>. The major coat protein, P3, as well as other proteins have been determined by X-ray crystallography, and the entire PRD1 virion has been determined with both X-ray crystallography and cryo-EM image reconstructions<sup>14, 16-21</sup>. Models have been developed to give a more in depth picture on various structures that make up this virus. Despite differences in host preference and genome sequence, PRD1 and Bam35 have been placed in the same family due to gross morphological reasons. Two other Gram-positive infecting members exist in this family, AP50 and phi NS11. The anthracis phage, AP50,

is not an ideal model system and phi NS11 is no longer available. Thus, Bam35 was the ideal candidate of these three to try and ascertain if the previously linked Tectivirus-adenovirus lineage is maintained<sup>16</sup> and to determine whether this phage belongs to the same family as PRD1.

The divergence of Gram-positive and Gram-negative bacteria is thought to have occurred one to two billion years ago<sup>12,40</sup>. Therefore, any present resemblance amongst their phages would be of great interest. Negative-stain electron microscopy has shown that the Bam35 virion morphology closely resembles similarly derived images of the virion of PRD1<sup>41</sup>. As previously mentioned, the genome lengths of each phage are almost identical and they are similarly organized (See Figure 2-1). It has been discovered that the capsids have similar diameters. Bam35 has an internal membrane that is congruent with PRD1. This means that Bam35 has an equitable volume to PRD1 in which DNA packaging can occur.

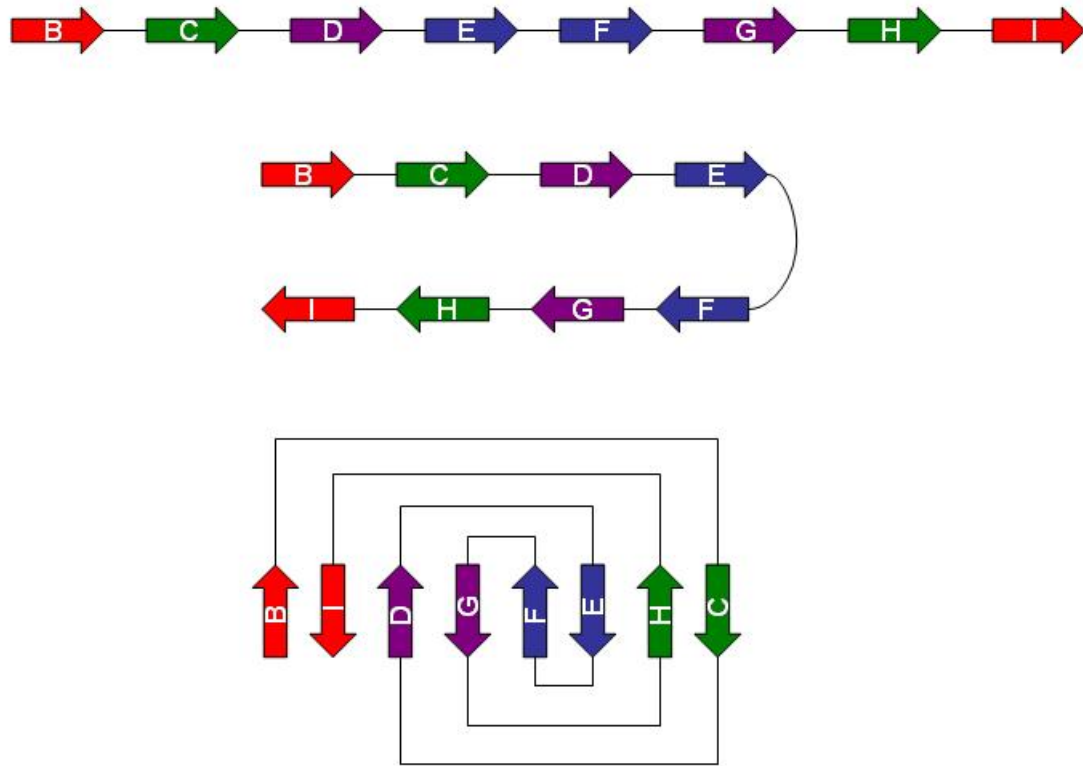
The major interest in this study is to be able to crystallize and obtain the atomic structure of the major coat protein, P3, of Bam35. This can then be compared to the previously determined structure of the P3 protein of PRD1<sup>14</sup> and the other structures that fall into this order (adenovirus hexon, PBCV-1 Vp54, and STIV MCP) in order to affirm the divergence theory. The coat structures previously mentioned have the conservation



**Figure 2-1. Genome organization of Bam35c (top) and PRD1 (middle). Note the similar genome length as measured by the kbp ruler (bottom).**

of a double-barrel trimer. A double-barrel trimer, as seen in PRD1, occurs when two 8-stranded viral jelly rolls, or  $\beta$ -barrels, are arranged in tandem in a single subunit. The jelly roll contains eight  $\beta$ -strands in sequence (B, C, D, E, F, G, H, and I). The A strand is omitted for historical reasons since the first example of a single viral jelly roll from tomato bushy stunt virus contained an extra strand before the jelly roll<sup>42</sup>. The jelly roll is folded down the middle, which causes E to pair up antiparallel with F via hydrogen bonds, D bonds with G, C bonds with H, and B with I. Hydrogen bonds occur within the sheet, but not at the edges. Therefore, B does not bond with C and F does not bond with G. Two sheets with four strands, BIDG and CHEF, comprise the folded jelly roll (See Figure 2-2). This situation makes it unique, unlike the true barrel seen with the  $\beta/\alpha$  motif. The strands are variable in length, but the width and separation of the sheet is rather stable. There are essentially four different ways in which an eight stranded jelly roll could fold<sup>43</sup>. However, only the form above can be seen in spherical viruses. This conserved form of folding of the major capsid protein, seen in several viruses of different

hosts, implies that the divergence theory can hold true. Although loss of similarity in sequences has occurred, the viral folds have been maintained<sup>13</sup>, thus preserving the viral self.

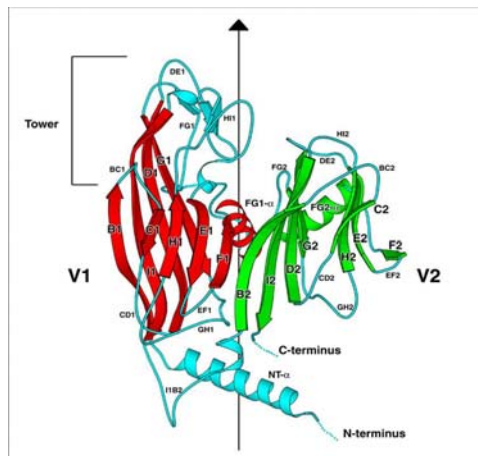


**Figure 2-2. Jelly roll fold.** The initial protein chain for the jelly roll contains 8 β-strands in sequence (top). The jelly roll folds down the middle to pair up B & I, C & H, D & G, and E & F (middle). Finally, the jelly roll folds (or rolls like a jelly roll) so that one sheet contains BIDG and another sheet contains CHEF (bottom). Strands G and F do not hydrogen bond.

It is known that the PRD1 P3 structure is a trimeric molecule with each subunit composed of two eight stranded viral jelly rolls, or β-barrels. The protein is made up of 394 residues. The P3 protein of Bam35 is made up of 356 residues. While the two show minimal sequence similarity (only 16% identity and 31% similarity), the folds are anticipated to be similar with >95% confidence by the program 3D-PSSM<sup>29</sup>. To prove

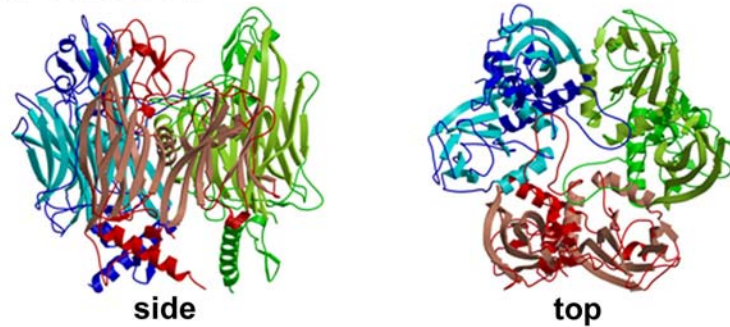
the assumption of similar folding and structure, a model of Bam35 was performed by threading the sequence onto the known structure of PRD1. CLUSTAL X<sup>30</sup> was used to insert gaps into the Bam35 sequence to optimize alignment. This model (see Fig 2-3 and 2-4) showed a conservation of the double barrel as a building block to the capsid protein, which could not have been determined from the sequence alignment alone<sup>13</sup>.

We hope to show that the above model of Bam35 P3 is indeed correct when crystallization and structure determination are performed. In doing so, we expect to enhance the evidence that the divergence theory is correct and that an icosahedral dsDNA viral ancestor does exist, from which a viral lineage can be traced. To determine the structure of Bam35 P3, it is important to have good expression of the Bam35 P3 molecule, to obtain it in high purity, and to develop well-formed and reproducible crystals that diffract to high resolution. The progress made on these points is discussed in this chapter.



**Figure 2-3. Single subunit of PRD1 P3.** This gives a detailed view of the jelly roll fold for one of the trimers of PRD1.

### A. PRD1 P3



### B. Bam35 P3 Model

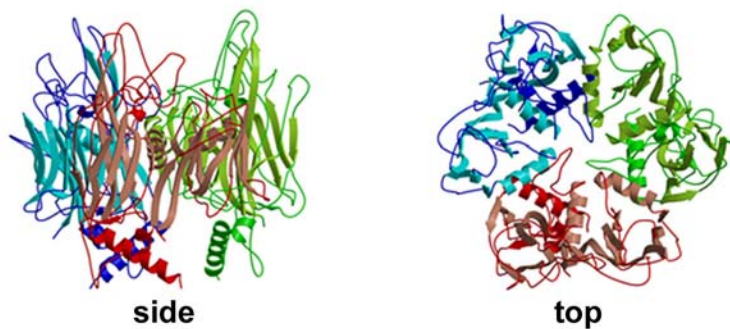


Figure 2-4. P3 structure of PRD1 and model of Bam35. (A) PRD1 P3 trimer structure (B) Bam35 P3 trimer model

## 2.2 Research Design

### Culture and Over-Expression of Protein P3:

For the growth of the Bam35 P3 protein, *E. coli* strain HMS174 (DE3) (pSK50) was used with the cloning vector being pJJ2 that was developed by our collaborator Dr. Jaana Bamford (University of Jyväskylä, Finland). A small amount was taken from the frozen stock (-80°C) solution and was streaked on a Petri dish containing LB (Luria Bertani) agar with Ampicillin. This dish was then incubated at 37°C for 24 hours.

A colony was picked and placed into a 100 mL baffled bottom Nalgene flask. This flask contained 20 mL of Luria broth spiked with 150 µg/mL of ampicillin. The flask, as well as all others in the growing process, was capped and the cap was opened a quarter of a turn to allow for adequate oxygenation. This was then placed in the Barnstead International Max Q 5000 incubator/shaker and incubated overnight at 37°C and 200 rpm. The next day, the overnight culture was placed into a 1 L baffled bottom Nalgene or glass flask containing 480 mL of Luria broth spiked with 150 µg/mL of ampicillin. This was allowed to incubate at 28°C and 200 rpm until proper optical density was reached ( $OD_{600} \approx 0.8$ ).

The temperature was reduced to 18°C one hour before addition of isopropyl-beta-D-thiogalactopyranoside (IPTG), at a concentration of 1 mM, to induce the recombinant protein expression. The culture was maintained at 18°C and 200 rpm overnight. The cells were pelleted (Beckman Coulter Allegra X-15R, 4500 rpm, 30 min, 4°C), collected into two 50 mL conical tip tubes, and each was re-suspended in approximately 10 mL of 20 mM Tris-HCl (pH 7.2) to be later purified.

Different variations of media were explored for culturing *E. coli*. When I first began this project, the media was ordered as a pre-made solution from Ward's through VWR. It provided for excellent culturing and protein production. However, it sometimes took awhile to receive the shipment and it was decided to look for ways to cut costs. So, a premixed powdered form (Amresco) of the media was ordered, which contained yeast extract, tryptone and sodium chloride in a 2:1:2 ratio. For a 1 L solution, 25 g of the

powder needed to be combined with distilled water and autoclaved at 120°C and 15 psi for about 15 minutes. This combination did not give the same results as the VWR media, so it was decided to add glucose to try to increase the bacterial production. In the beginning, 0.1 % glucose was added to the media mixture. Later, the addition was increased to 1% glucose. After several trials of getting very low yields in protein production, a determination of the problem needed to be made. Three batches of media were prepared with 0%, 0.1%, and 1% glucose and all were cultured under the same conditions. It was also decided to compare the previous VWR purchased media with our formulation and to test against a formulation available through Sigma as well. We have never been able to determine all of the components available in the VWR media, but Sigma let their formulation be known and the components could be purchased individually. The Sigma media was made up of yeast extract, tryptone, and sodium chloride and was filtered through a 0.2 µm filter, instead of being autoclaved. Three batches of bacteria in each of the three broths were cultured under the same conditions. The Sigma formulation was noted as giving comparable results to the VWR media. Four batches of media were then prepared for further analysis of the Sigma formulation. Two were made up of the same three components, but one was autoclaved while the other was filtered. Two other batches had the addition of 0.1% glucose with one treated by autoclaving while the other was filtered. This trial showed the best result to be the batch with no glucose added and being autoclaved. After narrowing down these conditions, we began to see better protein production.

**Isolation and Purification of Protein P3:**

The original protocol called for using a French press to isolate the P3 protein from the collected cells. This technique was too cost prohibitive, so it was replaced with sonication. A Branson Sonifier 150 set at level three was used for a total of two minutes. The sample was sonicated for one minute and allowed to cool down (approximately 5 minutes) before sonicating for the other minute.

The cell debris was centrifuged (Beckman Coulter Allegra X-15R, 4500 rpm, 30 min, 4°C) and the supernatant was decanted into a 15 mL conical tip vial. A sample was taken for sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE). The supernatant was further clarified by centrifugation (Beckman Coulter Allegra X-15R, 4500 rpm, 2-17 hr, 4°C) and a sample was taken for SDS-PAGE.

A 5 mL Pharmacia HiTrap Q cation column was equilibrated by applying 25 mL of 20 mM Tris-HCl (pH 7.2) with a pump into the column. The sample was then applied in the same way and the flow through was collected. The column was washed with 10 mL of 20 mM Tris-HCl (pH 7.2), with the flow through being collected in three equal fractions. The column was then washed with 25 mL of 1 M NaCl followed by 50 mL of 20 mM Tris-HCl (pH 7.2) to reequilibrate the column. Samples were taken from all collections to run on SDS-PAGE.

Next, a 5 mL Pharmacia HiTrap SP anion column was used. The column was equilibrated as above and the combined flow throughs were applied to the column. This

column was then placed onto the Bio-Rad BioLogic DuoFlow system. The buffers used are 40 mM Tris base, 40 mM Tris-HCl, 2 M NaCl, and de-ionized water. All buffers were degassed prior to placing them into the system. The program was set to collect 1 mL fractions and run at pH 7.2 during the course of the run. First, an isocratic flow was used with 0% salt for a volume of 5 mL at 1 mL/min. Next, a linear gradient of 0-30% (0 mM – 300 mM) salt ran for a volume of 30 mL at 1 mL/min. This was followed by an isocratic flow of 100% (1 M) salt for 5 mL at 1 mL/min. A final wash of the column came with an isocratic flow of 0% salt for 10 mL at 1 mL/min. Samples (20  $\mu$ L) were taken for SDS-PAGE from the indicated fractions corresponding with the peak(s) on the chromatogram determined by absorbance at 254 nm. Desired fractions were combined based on the findings of the SDS-PAGE gel.

The SDS-PAGE gels showed several bands in addition to the P3 protein present in the purified solution. Mass spectroscopy was performed to identify the other bands present. It was discovered that the other bands were still all associated with Bam35. This led us to believe that during the isolation of the P3 from its bacterial host, the protein was being cleaved at different intervals, possibly by cellular proteases. This led us to look at the possibility of using protease inhibitors. A cocktail of three protease inhibitors was initially used, comprised of leupeptin, pepstatin, and phenylmethanesulphonylfluoride (PMSF). Stock solutions were made by dissolving 2 mg of leupeptin in 1 mL of distilled water, dissolving 1 mg of pepstatin in 900  $\mu$ L of 200 proof ethanol and 100  $\mu$ L of glacial acetic acid, and dissolving 17 mg of PMSF in 1 mL of 200 proof ethanol. A volume of 50  $\mu$ L, 100  $\mu$ L, and 100  $\mu$ L, respectively, of stock solutions, was added per 100 mL of

cell suspension. Three variations of cell lysis were attempted to try and reduce protein degradation. First was the freeze/thaw method. Two samples, one with protease inhibitors and one without, were placed in liquid nitrogen and allowed to freeze. Then the two samples were placed in a 37°C water bath to thaw. This procedure was repeated two more times. Two other samples were also frozen in the same manner, but allowed to thaw more slowly in a 0°C ice bath. Another method involved the use of lysozyme. One sample contained protease inhibitors and the other did not. These samples were suspended in lysozyme lysis buffer (50 mM Tris-HCl (pH 7.2), 5% glycerol, and 1 mM dithiothreitol (DTT)), rather than the previously mentioned 20 mM Tris-HCl (pH 7.2) buffer. The lysozyme was added to a final concentration of 300 µg/mL and allowed to incubate for approximately four hours at 4°C. Finally, different methods of sonication were investigated as well. Two samples contained the previously mentioned protease inhibitors, while one was suspended in 20 mM Tris-HCl (pH 7.2) buffer and the other was suspended in another lysis buffer (50 mM Tris-HCl (pH 7.2) and 5 mM DTT). Two other samples were prepared in the two buffers, but without protease inhibitors. These four samples were sonicated for 15 second intervals with 30 second rests, for a total of four times. The samples were kept on ice during the entire process. Another set of four samples, treated as previously mentioned, were sonicated for 10 second intervals with 30 second rests for a total of ten times. Finally, two samples, both in 20 mM Tris-HCl (pH 7.2) buffer were treated in the original manner, one without protease inhibitors and one with to act as controls.

**Protein Concentration for Crystallization Trials:**

A concentration calibration curve for the Bradford Assay was performed according to the manufacturer's (Bio-Rad) instructions against known concentrations of bovine serum albumin (BSA). This was done on a Beckman Coulter DU 530 UV/Vis Spectrophotometer at a wavelength of 595 nm. Then, a 20  $\mu$ L aliquot was combined with 1 mL of Bradford Assay solution (at room temp) and was shaken vigorously in the cuvette. A blank solution, containing only Bradford Assay, was placed into another cuvette. The samples were incubated at room temperature for 10-15 minutes. The blank was placed into the sample holder of the DU 530 and was analyzed. The sample cuvette was then placed into the sample holder and the absorbance value was obtained. This value was used to determine the concentration of protein in the solution by use of the previously derived calibration curve. Once the initial concentration was determined, a final volume was calculated to concentrate the sample down to achieve the desired concentration of 10 mg/mL. To concentrate the sample, the solution was placed into a 15 mL Vivaspinn tube with a 30,000 molecular weight cutoff (MWCO) and centrifuged (Beckman Coulter Allegra X-15R, 3050 rpm, 4°C) until the volume desired was shown within the tube's graduations. This solution was then put in a microcentrifuge tube, labeled, and placed in storage at 4°C.

**Crystallization:**

Crystallization trials were carried out in Hampton VDX trays, which are modified 24-well cell culture trays. The hanging drop method was used with initial amounts of reservoir and protein solution in the drop being 2  $\mu$ L each. To set up a tray for

crystallization, the edges of the wells must be greased to allow for a closed system when sealed with a coverslip. The grease was applied by way of a syringe. Next, each reagent needed to create the reservoir solution is added one by one. The total volume of the well solution used was 1 mL, so distilled water was added last to meet this volume requirement. The tray is lightly shaken to allow for adequate mixing of all chemicals. To prepare the crystallization drop, 2  $\mu$ L of the reservoir solution is first pipetted onto a coverslip. Then, 2  $\mu$ L of the concentrated, purified protein solution is added and mixed into the reservoir drop. The coverslip is then placed over the well and pressed down, so as to form a tight seal. All 24 wells were prepared in a similar manner and the tray was set inside a cabinet to be isolated from sunlight.

A reservoir solution is typically made up of a precipitant (usually an organic compound) of which there can be more than one, a buffer, and possibly a salt. To determine the identity and concentrations of reagents needed to crystallize the protein, two pre-made crystal screens were purchased from Hampton<sup>44, 45</sup>. The protein solution was introduced to drops from the 98 different combinations of chemicals in the screens and sealed with a well solution containing 1 mL of the chemical combination. The trays were checked periodically for crystal growth. After several conditions were identified as possible candidates, more defined trials were performed. The concentrations of the individual components, the pH of the buffer, the use of different salts, and temperature were altered to try to hone in on a good method for crystallization. Several conditions were found to produce crystals. The P3 protein crystallized quickly, usually within 24 hours.

**X-ray Diffraction:**

Once crystals were formed, a test for X-ray diffraction was done. The coverslip was taken from over the well and placed with the drop facing up. The crystal mounting is performed with the aid of a stereomicroscope. For room temperature data collection, the crystal is mounted into a 0.5 mm glass capillary. The crystal is drawn into the capillary along with some of the drop solution. Excess liquid is removed from around the crystal with a small wick created from filter paper. The wet portion of the wick is cut with a razor and left in the capillary tube to supply moisture to the crystal after the capillary is sealed. The ends of the capillary are cut with a glass cutting board and sealed with bee's wax. For data collection at cryo temperatures, a loop was used to lasso a crystal from the drop. Once the crystal was seated, the loop was placed onto the goniometer head of the X-ray diffractometer (Bruker Microstar) under a nitrogen cryo stream (100 K; Oxford Cryosystems COBRA). The power was set at 2.5 kW (45 kV and 55 mA). The remaining steps were the same for both the room temperature and cryo-preserved crystals. The crystal was then centered as best as possible within the center of the X-ray beam. A camera focusing on the crystal gave streaming video to the computer monitor. The crystal could be spun in 90° angles and adjusted up and down and back and forth until it stayed within the centered area indicated on the computer screen. This center corresponds to the center of the X-ray beam. Then various modes for detection were set with the systems software (Proteum 2). The detector was adjusted to a distance of 70 mm from the goniometer head (the sample holder). The length of time could be adjusted to allow for better intensity of diffraction. If no diffraction was observed, the crystal was

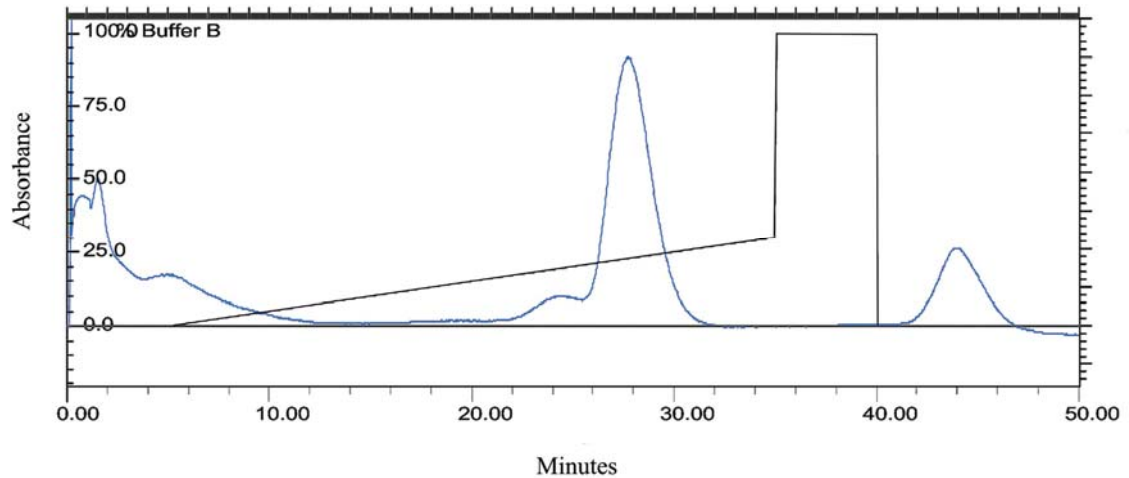
readjusted for additional collection or removed. However, if diffraction was detected and to a decent resolution, further exposure of the crystal followed to allow for preliminary data to be collected. Proteum 2 was used to create a collection scheme to expose the crystal to X-rays for several frames and at different angles. After this was completed, manipulation of the data gave a preliminary space group. If decent results were evident at this stage, even more exposure could be performed. The crystal could then be exposed to more X-rays through a program executed by the computer program. This usually resulted in hundreds of frames that could then be read by SAINT NT<sup>46</sup> and XPREP<sup>47</sup> to give better results. This would give a better idea of the space group involved and the degree of certainty with which the results could be trusted.

## **2.3 Results and Discussion**

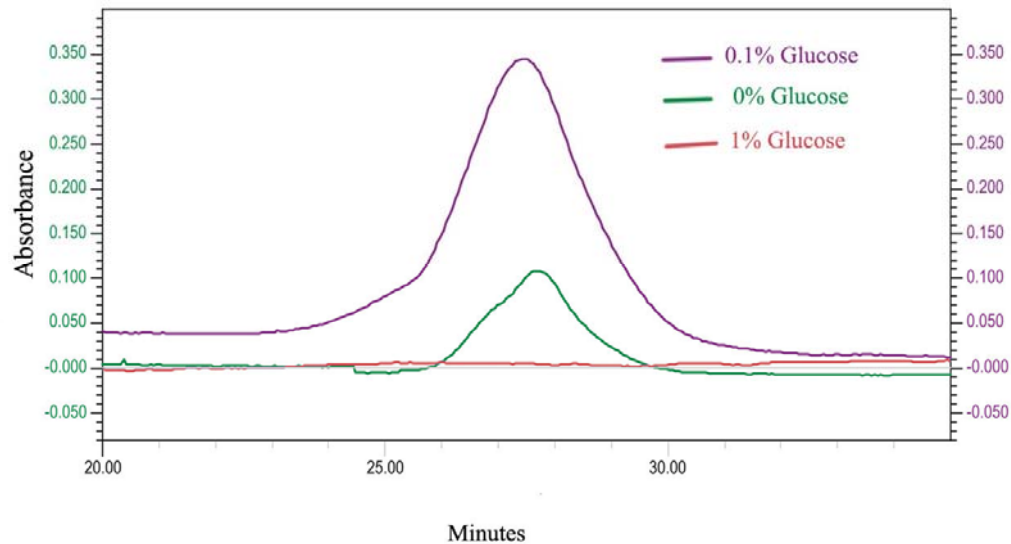
### **Culture and Over-Expression of Protein P3:**

In order for crystallographic studies to be done on Bam35 P3, it is important to have a large quantity of highly pure protein. In order to obtain the required protein, it is important to have optimal growth conditions for the bacteria that are over-expressing the protein. Thus, great interest was placed in finding the best media conditions for bacterial growth and protein expression so that optimal protein production could be achieved. A pre-made LB broth, purchased through VWR, provided good protein expression, but it was costly and delivery was not guaranteed, so other methods were investigated. Three batches of media, each containing yeast extract, tryptone, and sodium chloride (purchased as a pre-mixed composite from Amresco), were prepared and spiked with either 0%, 0.1%, or 1% glucose. All three batches were autoclaved and then used for culturing

bacteria under the same conditions. After running SDS-PAGE gels to confirm the presence of P3 and approximating the protein amounts from the UV chromatograms obtained on the Bio-Rad DuoFlow (See Figure 2-5), it was determined that 0.1% glucose gave the best results (See Figure 2-6).



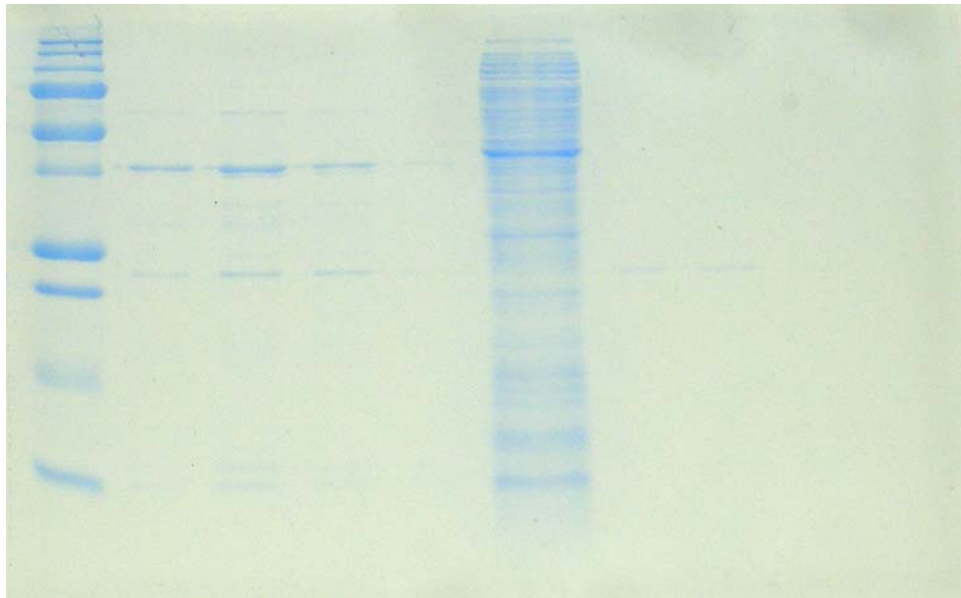
**Figure 2-5. Full chromatogram of P3 run through the 5 mL GE Healthcare Hi-Trap SP column. The P3 peak occurs in the middle of the run, starting around 22 minutes and ending around 30 minutes. The other peaks in the chromatogram were analyzed on a SDS-PAGE gel and were found to not be P3. The black line indicates the salt (Buffer B) as it was programmed to run. The SP column is a cation exchange column, as purification is carried out using ion exchange chromatography. All other chromatograms shown are analyzed in the same manner, but are zoomed into the peak of interest.**



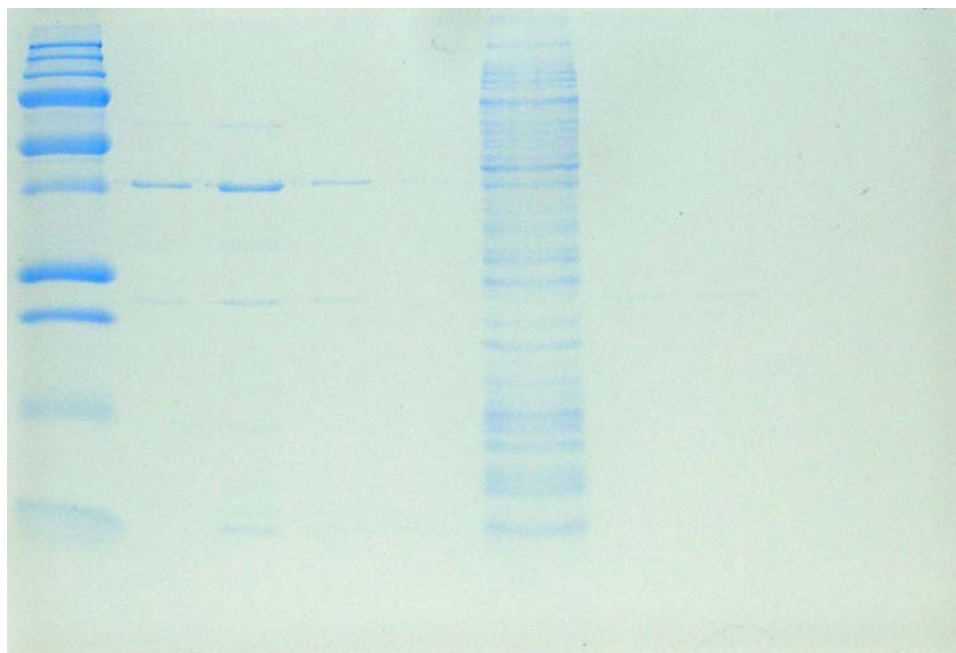
**Figure 2-6. Glucose comparison chromatogram overlay. The peak corresponding to the 0.1% glucose concentration gave the highest absorbance reading of all three trials. The 1% trial shows no peak, indicating that the concentration of glucose has caused the bacteria to multiply rapidly in the rich medium. After the population could not be supported any longer, the bacteria began to die and the P3 protein was lost.**

Since the formulation of the Sigma media had been determined (yeast extract, tryptone, and sodium chloride filtered through a 0.2  $\mu\text{m}$  filter, instead of being autoclaved), we wanted to compare these results with the VWR media, to see if similar results could be obtained. Three batches of bacteria were cultured under the same conditions. One batch was made with the VWR media, one was made with the Sigma media, and the last batch was the Amresco premixed formulation, which was spiked with 0.1% glucose. Due to bad weather interrupting the growing process, only initial results were available. It was determined that the quantity of bacteria produced with the Sigma media gave equivalent results to the VWR media. Using this knowledge, four batches of media were prepared. Two of the batches contained the three components used in the

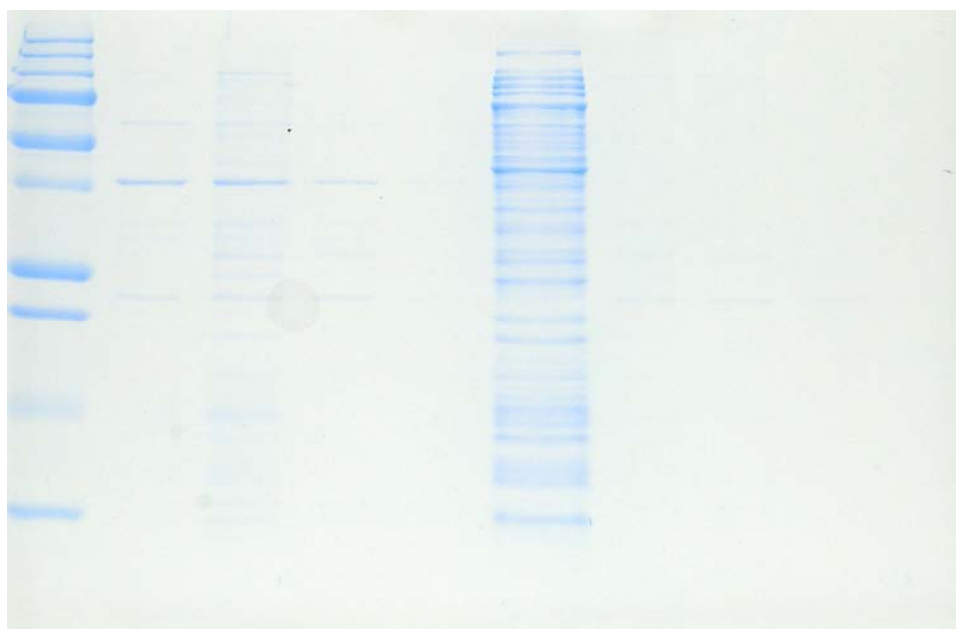
Sigma formulation. One of these batches was autoclaved and the other was filtered through a 0.2  $\mu$ m filter. The other two batches contained the three components with the addition of 0.1% glucose. Again, one batch was autoclaved and the other filtered. The best results occurred with the batch that contained no glucose and was autoclaved (See Figures 2-7 – 2-13). These results were not run in triplicate. However, these optimal results have been repeated numerous times and the same results have been obtained.



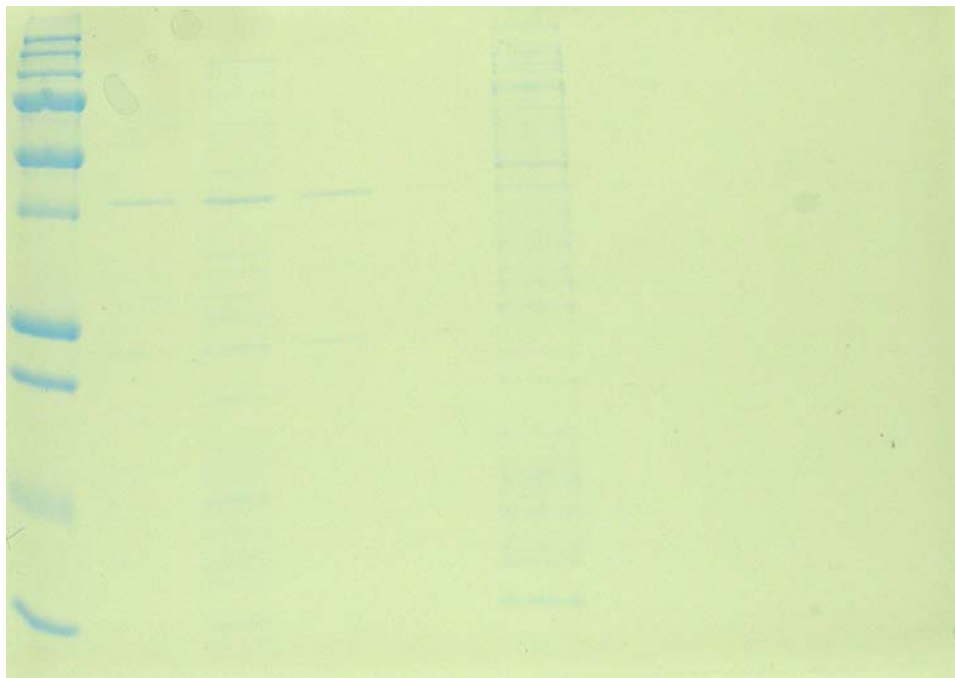
**Figure 2-7. SDS-PAGE gel analysis of P3 produced with no glucose and autoclaved media. The first lane of the 15% gel corresponds to the molecular weight (MW) standard. The next four lanes correspond to the flow through of the isolated protein and washes with 20 mM Tris-HCl pH 7.2 through the Hi-Trap Q column. The P3 protein is seen as the strongest band (second from the top) in the first three of the four lanes and can be seen very faintly in the fourth. The sixth lane contains the 1 M NaCl wash, which washes off any components still sticking to the column. The last four lanes are the flow through collected from the Q column and the three washes with the Tris buffer onto the Hi-Trap SP column. Notice that the P3 protein band is not evident, as it binds to the SP column.**



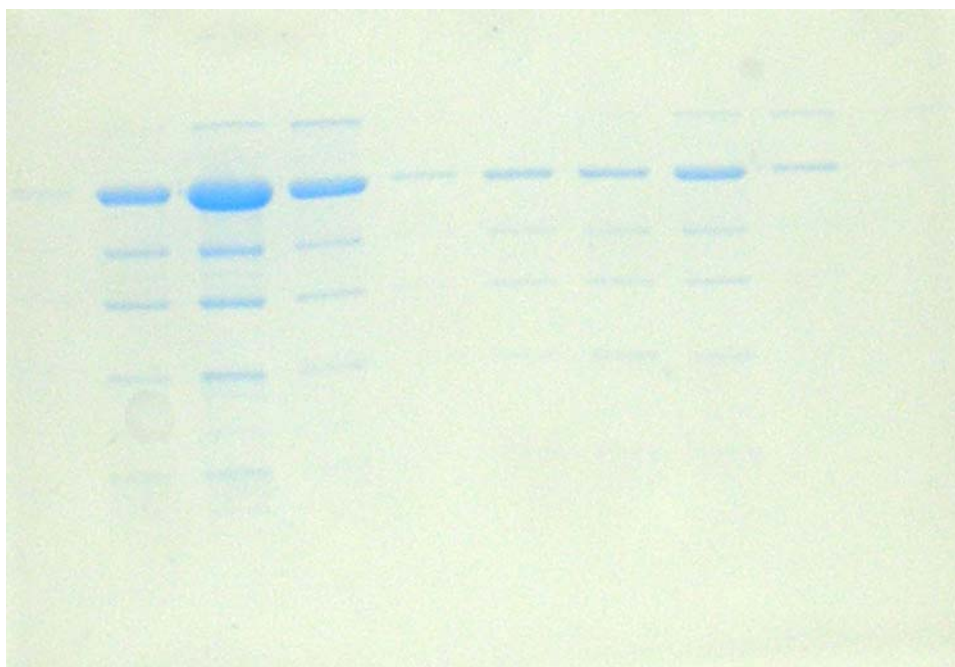
**Figure 2-8. SDS-PAGE gel analysis of P3 produced with no glucose and filtered media. Refer to Figure 2-7 for lane designations.**



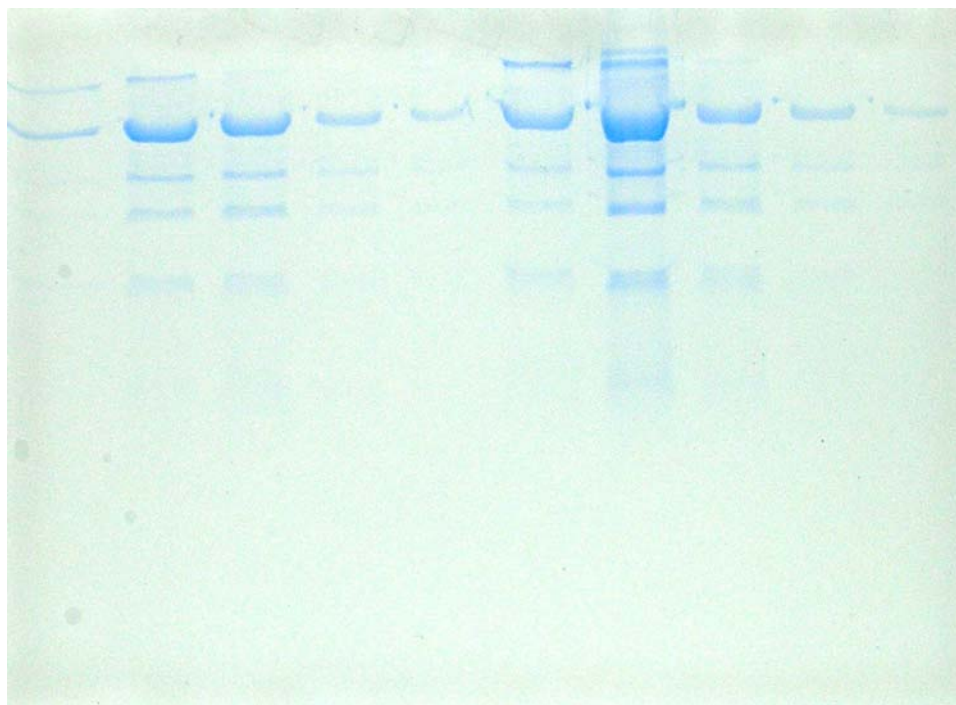
**Figure 2-9. SDS-PAGE gel analysis of P3 produced with glucose and autoclaved media. See Figure 2-7 for lane designations.**



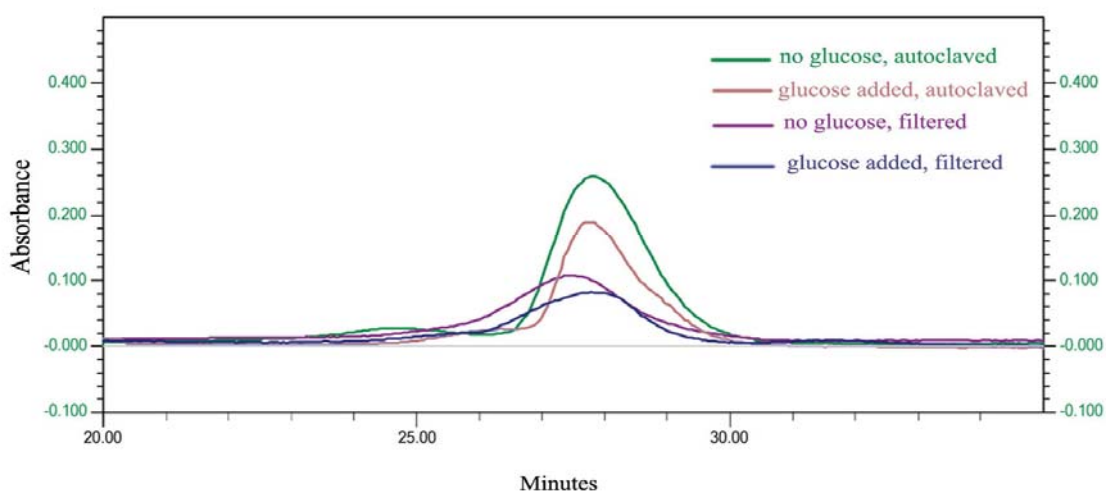
**Figure 2-10. SDS-PAGE gel analysis of P3 produced with glucose and filtered media. See Figure 2-7 for lane designations.**



**Figure 2-11. SDS-PAGE gel analysis of P3 produced with no glucose and autoclaved and filtered media run on Bio-Rad DuoFlow. This gel shows the results of the analysis of the SP column run on the Bio-Rad DuoFlow. The first five lanes contain samples taken from five 1 mL culture tubes corresponding to the P3 protein as shown on the chromatogram for the autoclaved, glucose deficient conditions. The last five lanes contain samples taken for the filtered, glucose deficient conditions. This gel shows that the autoclaved conditions gave a higher P3 production.**



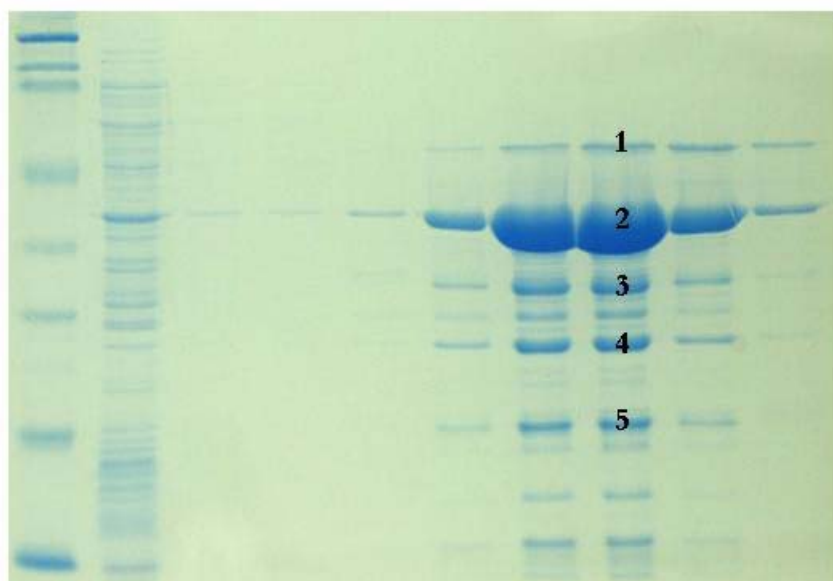
**Figure 2-12. SDS-PAGE gel analysis of P3 produced with glucose and autoclaved and filtered media run on Bio-Rad DuoFlow.** This gel shows the results of the analysis of the SP column run on the Bio-Rad DuoFlow. The first five lanes contain samples taken from five 1 mL culture tubes corresponding to the P3 protein as shown on the chromatogram for the autoclaved, glucose containing conditions. The last five lanes contain samples taken for the filtered, glucose containing conditions. This gel shows that the P3 production in these two instances is fairly equal, but is less than the glucose deficient, autoclaved condition seen in Figure 2-11.



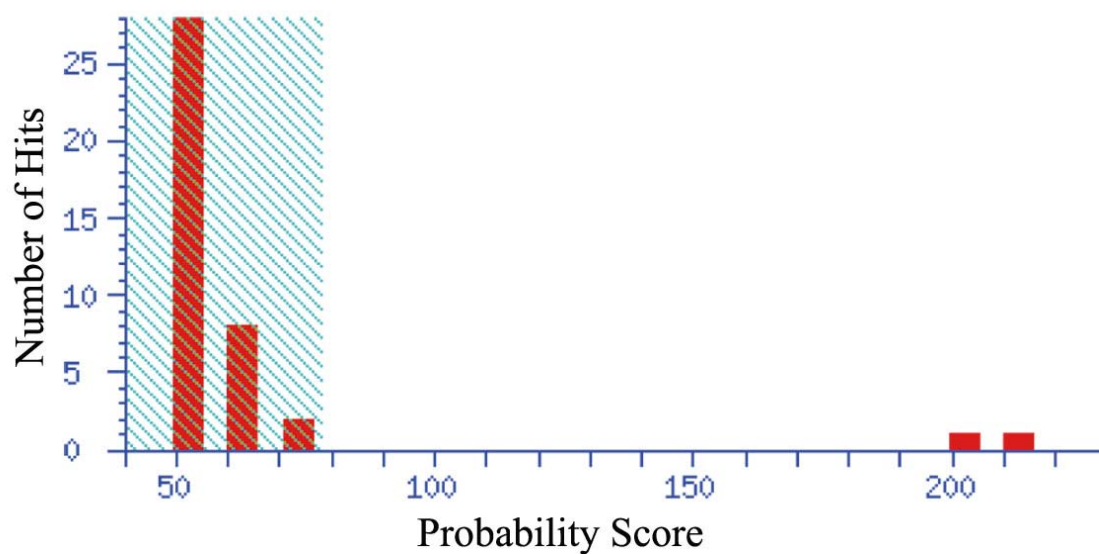
**Figure 2-13. Chromatogram overlay of glucose/no glucose and autoclaved/filtered comparisons.** This chromatogram overlay of the Bam35 P3 peak from the SP column run on the Bio-Rad DuoFlow from all four conditions confirms the SDS-PAGE analysis that the optimal condition for P3 production occurred in the media containing no glucose and was autoclaved.

### **Isolation and Purification of Protein P3:**

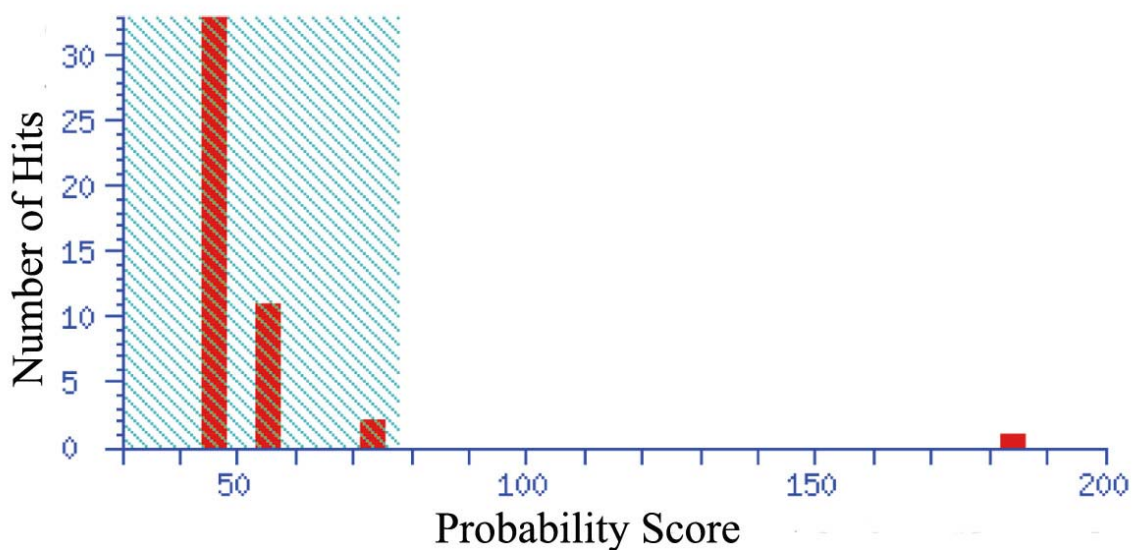
Every SDS-PAGE gel of the purified protein showed several bands in addition to the expected position of the P3 protein. Gel filtration was unable to remove any of the contaminants with the protein eluting as a single peak corresponding to the molecular weight expected for the P3 trimer (~120 kDa). Since gel filtration did not seem to help with the purity of the protein and only resulted in some loss of protein quantity, this step was removed from future purification protocols. It was decided to analyze the other bands present with MALDI-TOF, Matrix Assisted Laser Desorption /Ionization-Time of Flight, mass spectroscopy. The bands were cut from the gels, dissolved, and trypsinized before analysis – a service provided by the Department of Biochemistry and Molecular Biology Core Facility. It was discovered that one band, the highest molecular weight (around 70 kDa), was most likely oligopeptide-binding protein precursor from our *E. coli* that was co-eluting with our protein of interest. However, the other bands were all Bam35 P3 that had been cleaved, though all were predicted as a hypothetical protein since Bam35 P3 is not annotated in the Mascot database (See Figures 2-14 – 2-19). This led us to believe that during the isolation of the P3 from the bacterial host, the protein was being cleaved in multiple positions.



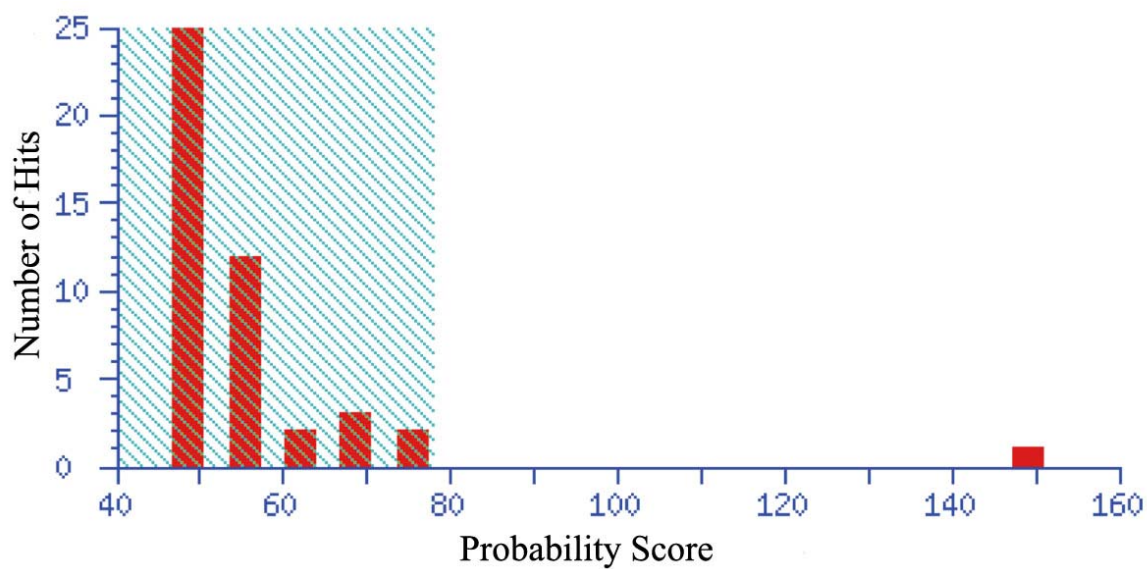
**Figure 2-14. SDS-PAGE gel reference for MALDI-TOF results. Numbered bands from this SDS-PAGE gel reference should be used to relate to MALDI-TOF results in Figures 2-15 to 2-19. The P3 protein was highly concentrated to bring out “contaminants” and there are other bands present along with the protein than the ones noted. However, only the more evident bands were analyzed with MALDI-TOF.**



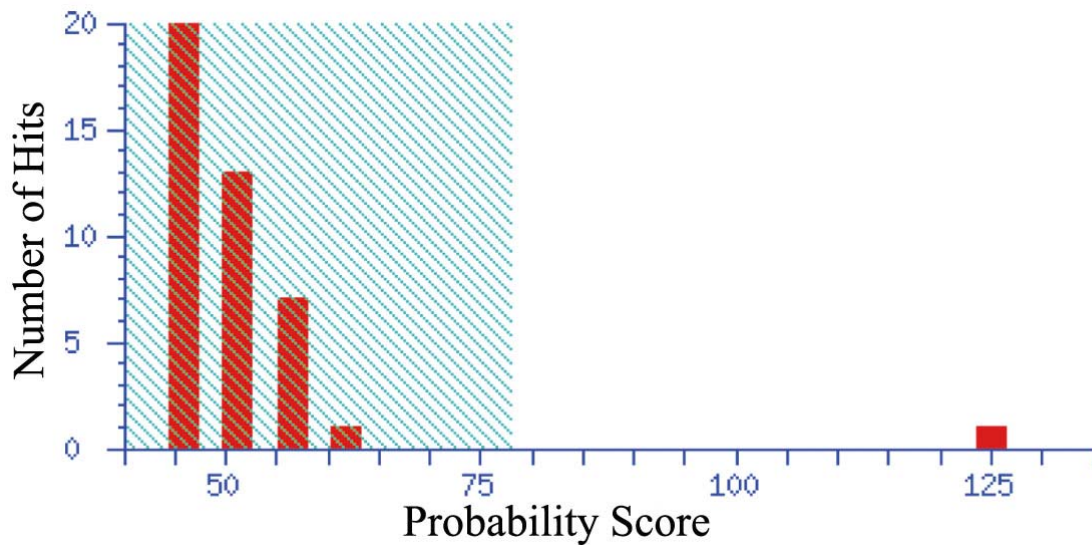
**Figure 2-15. MALDI-TOF results for band #1. Scores of greater than 78 are considered significant. The high score here correlates with the band being oligopeptide-binding protein precursor from our *E. coli*, a trace component that we are unable to eliminate from the pure P3 protein solution.**



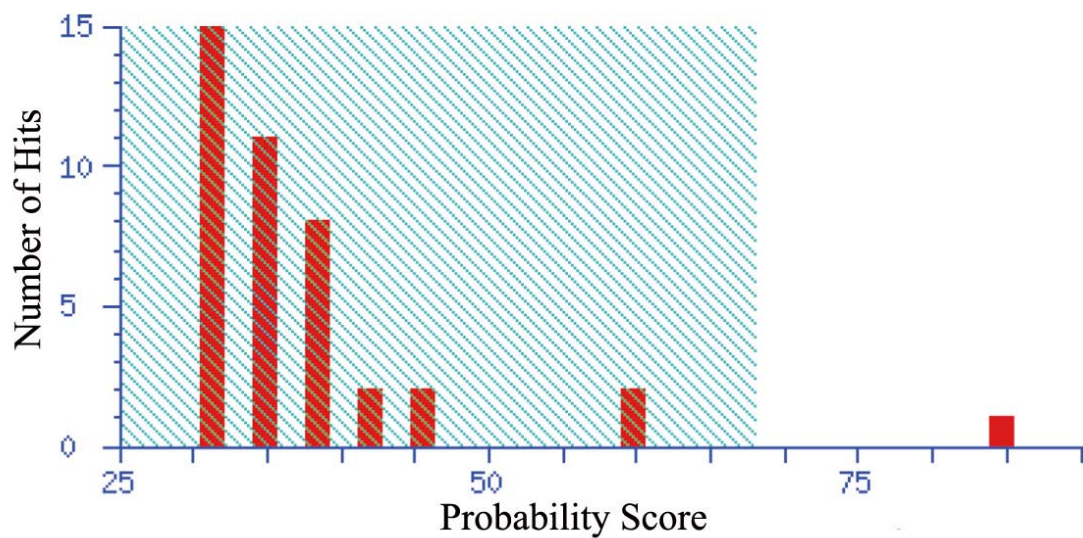
**Figure 2-16. MALDI-TOF results for band #2. The MALDI-TOF results confirm that it is indeed Bam35 P3.**



**Figure 2-17. MALDI-TOF results for band #3. The MALDI-TOF results confirm that it is hypothetical protein Bam35 P3.**



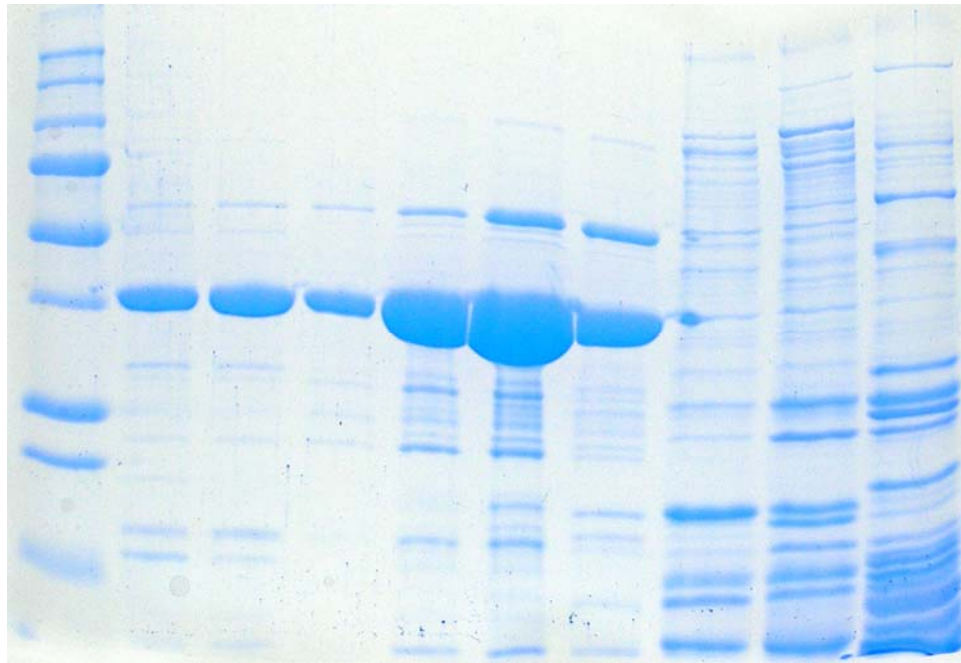
**Figure 2-18. MALDI-TOF results for band #4. The MALDI-TOF results confirm that it is hypothetical protein Bam35 P3.**



**Figure 2-19. MALDI-TOF results for band #5. The MALDI-TOF results confirm that it is hypothetical protein Bam35 P3.**

Since the mass spectrometric analysis showed that P3 was being cleaved at multiple positions, we thought that it might be proteases from the *E. coli* being released when the bacteria are broken apart during protein purification, even though the bacterial

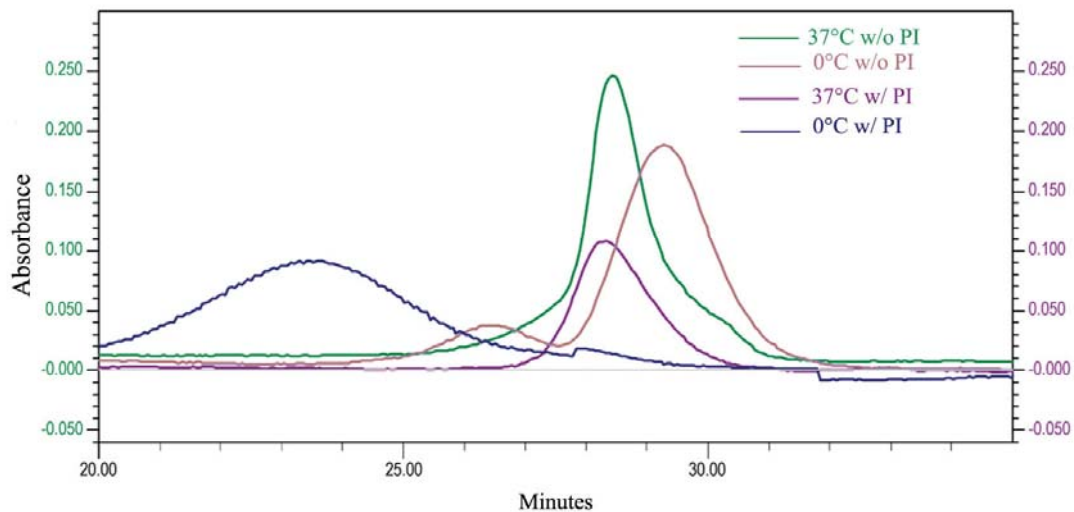
host is specially designed to reduce proteases. We decided to use protease inhibitors (PI) to see if this will eliminate the protein cleavage. However, it was established that the addition of PI did nothing to improve the apparent protein degradation (See Figure 2-20).



**Figure 2-20. SDS-PAGE gel analysis of P3 with PI added.** This is an SDS-PAGE gel analysis from samples taken from the SP column run on the Bio-Rad DuoFlow. Protease inhibitors were spiked into the cultured bacteria before sonication. The first lane contains the MW standard. The next six lanes contain samples from the peak on the chromatogram corresponding to P3. The last three lanes are taken from another peak which appeared at the end of the run. This gel shows that there are still degradation bands present along with the P3 (indicated as the large bands in lanes 2-7), despite the PI addition. Refer to Figure 2-14 for a reference gel with no PI added.

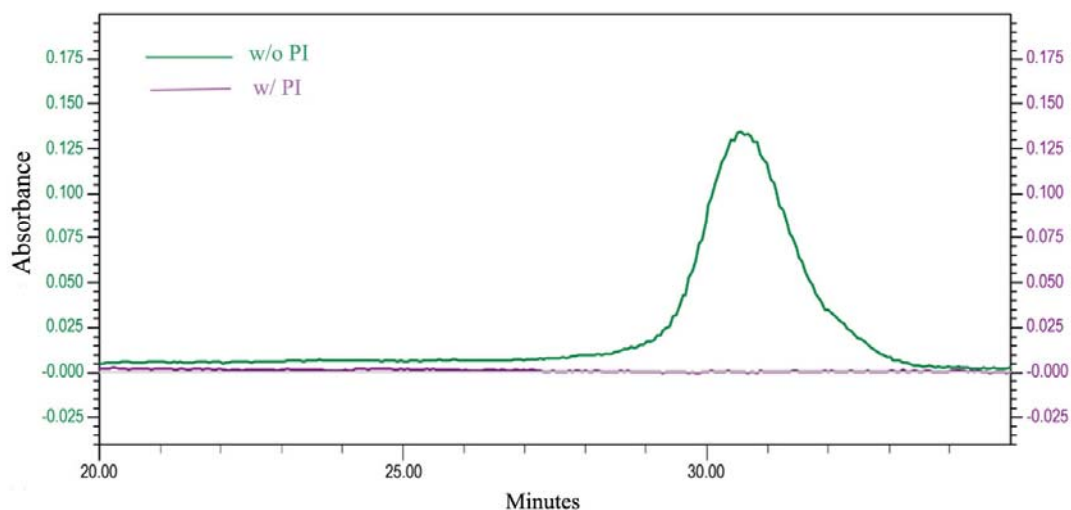
Therefore, other methods of isolation from the bacterial host were performed to attempt a reduction in degradation. The original method used sonication (two 1-minute sonications with an approximately five minute resting period on ice between them) to break open the cells and release the Bam35 P3 protein. First, the freeze/thaw method was performed. It was chosen because it is one of the more gentle isolation techniques, due to the lessened amount of stress on the bacteria. After spiking two of the four samples with

PI and thawing at different temperatures, it was determined from the SDS-PAGE gels and the ion exchange chromatograms that this method did not seem to diminish the other Bam35 P3 bands present (See Figures 2-23). Also, it was not adequate in releasing a decent amount of protein (See Figure 2-21).



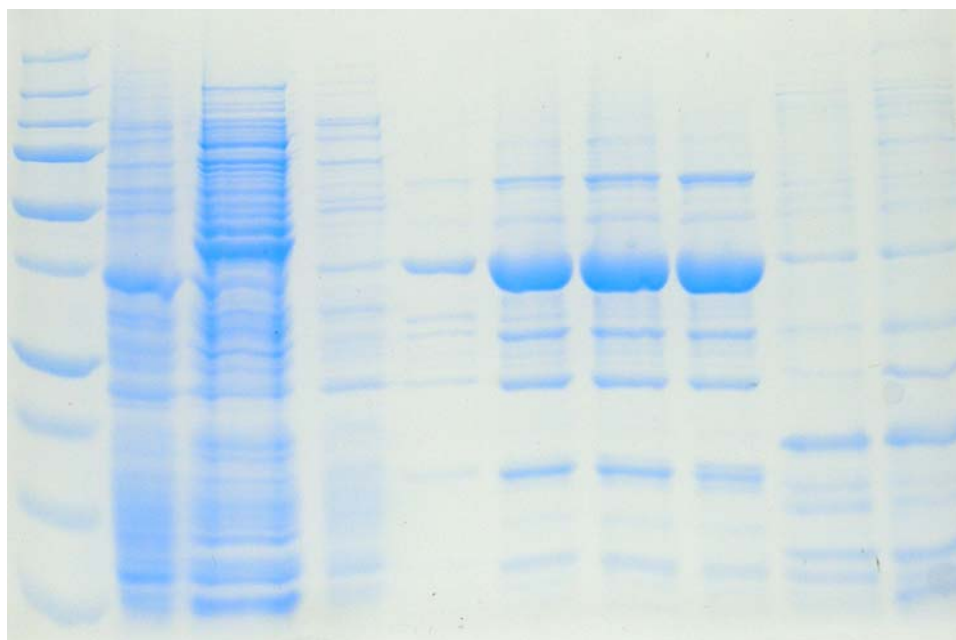
**Figure 2-21. Chromatogram overlay of freeze/thaw isolation methods.** The chromatogram overlay of all four analyses of the freeze/thaw methods shows that in relation to actual protein recovery, these methods are quite poor. The thawing at 37°C gave the highest amount, which is not surprising. It is surprising to note that the PI seem to reduce the amount of P3 protein recovered.

Isolation by means of lysozyme was investigated. This is another well known method of isolation that is used because it is more gentle than sonication. Again, this method showed small amounts of protein were isolated and no reduction in the other Bam35 P3 bands were achieved on the SDS-PAGE gels (See Figures 2-22 and 2-23).

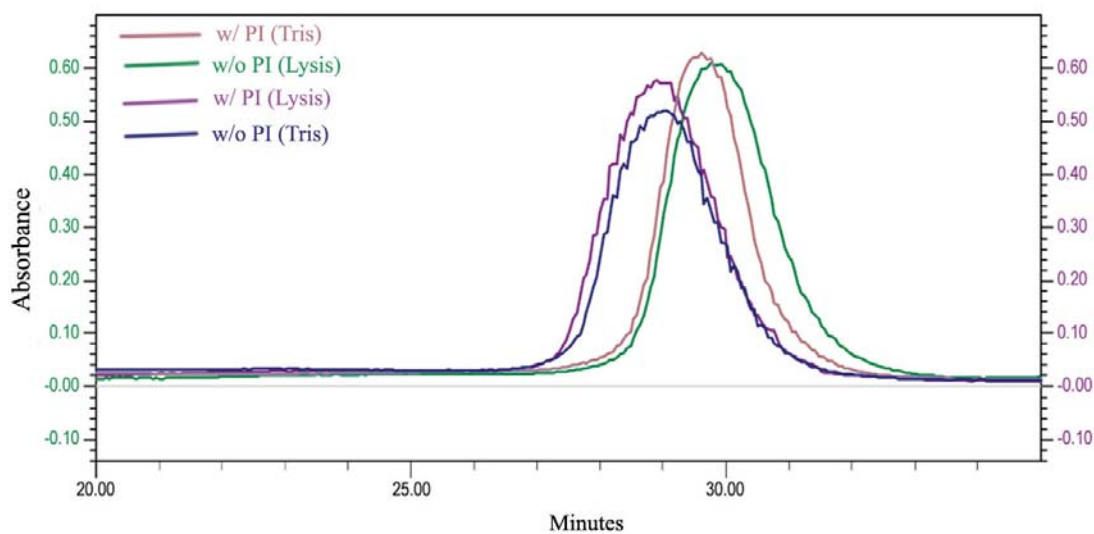


**Figure 2-22. Chromatogram overlay of the lysozyme isolation method. This lysozyme chromatogram overlay shows that the P3 peaks do not occur in the same places for the two analyses. There is a P3 production for the PI addition, but that peak occurs very early in the run. As far as protein recovery, again this method is quite poor.**

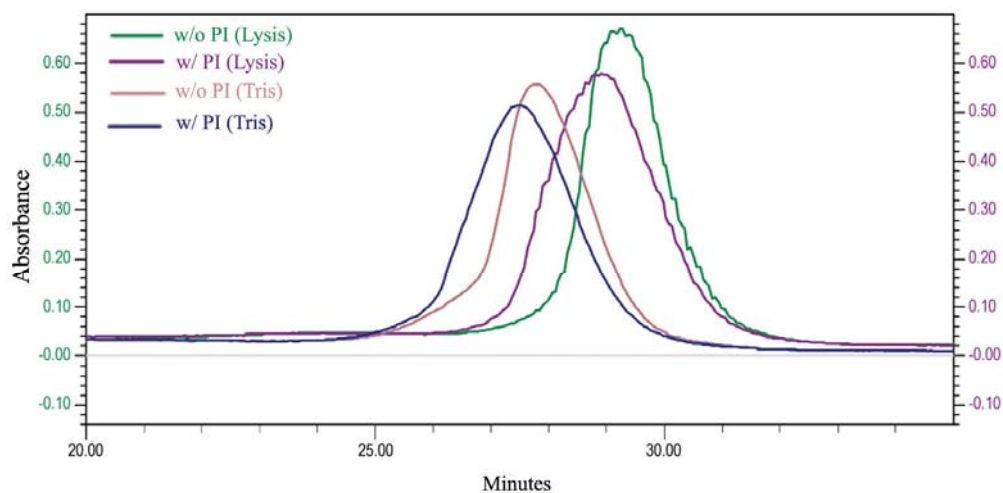
While the two previously mentioned methods are considered to be gentler, it was evident that protein release was hindered and the gentler methods did not seem to inhibit protein degradation. Different sonication times and repetitions were performed to see if these would produce less degradation and possibly even produce more protein. A marked improvement was noted in released protein in the two new sonication timed trials (4 x 15 secs and 10 x 10 secs) versus the previous method (2 x 1 min). However, each of the three different variations showed no improvement or worsening of the protein degradation issue (See Figures 2-23 – 2-26). Also, all trials performed, including the freeze/thaw and lysozyme methods, showed again that the addition of PI did nothing to improve on degradation. It was noted from comparing protein amounts from the chromatograms that the addition of the PI actually reduced the amount of P3 released.



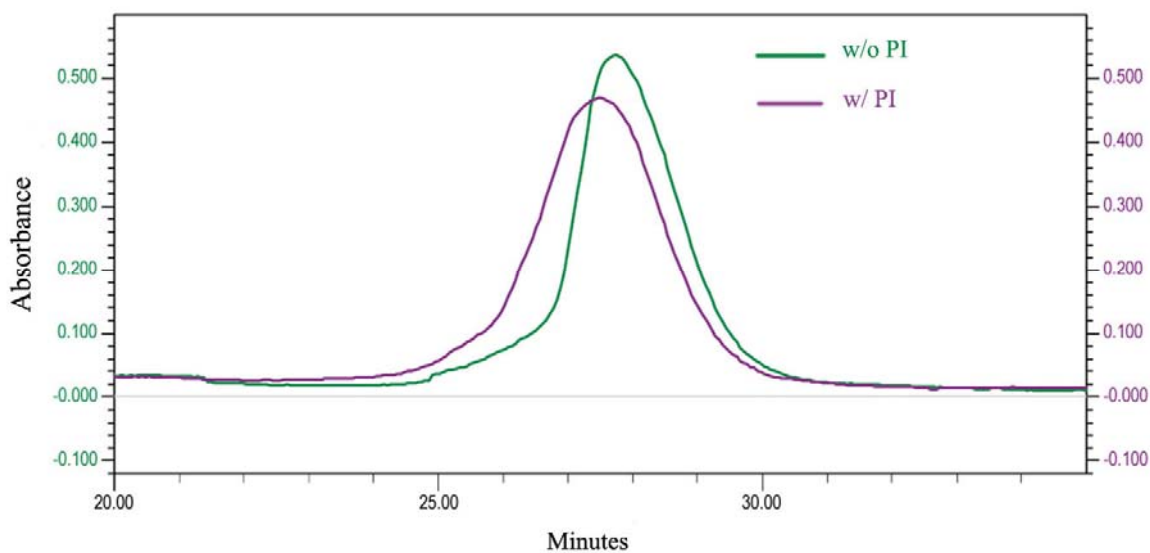
**Figure 2-23. SDS-PAGE gel of P3 sonicated 10 times for 10 seconds in lysis buffer with PI added. This gel is indicative of all SDS-PAGE gels analyzed for all isolation methods. No method or PI addition was able to reduce or eliminate the degradation bands.**



**Figure 2-24. Chromatogram overlay of sonication method (4x15). As far as protein recovery, this results in a mixed message. However, it appears that the addition of PI in Tris buffer, as opposed to lysis buffer, gives the best recovery amount.**



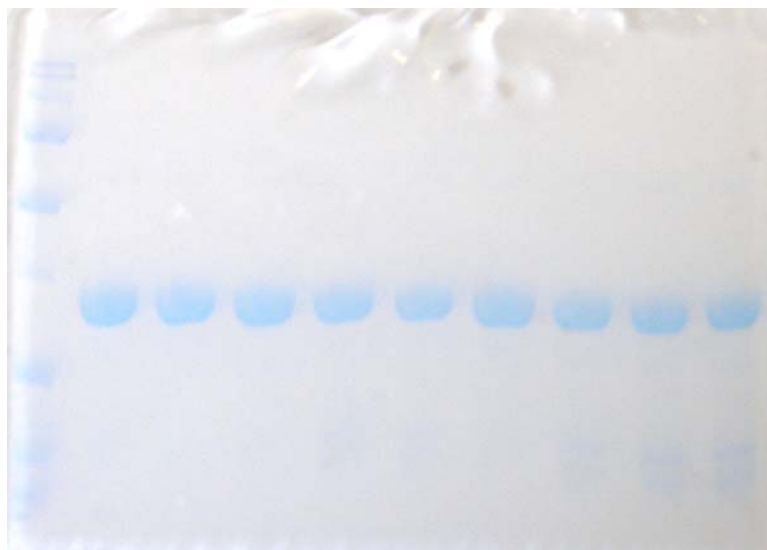
**Figure 2-25. Chromatogram overlay of sonication method (10x10). This method showed that lysis buffer was preferred over Tris buffer. This method gave the highest protein recovery, with the optimal isolation method being carried out with no addition of PI and the protein being in lysis buffer.**



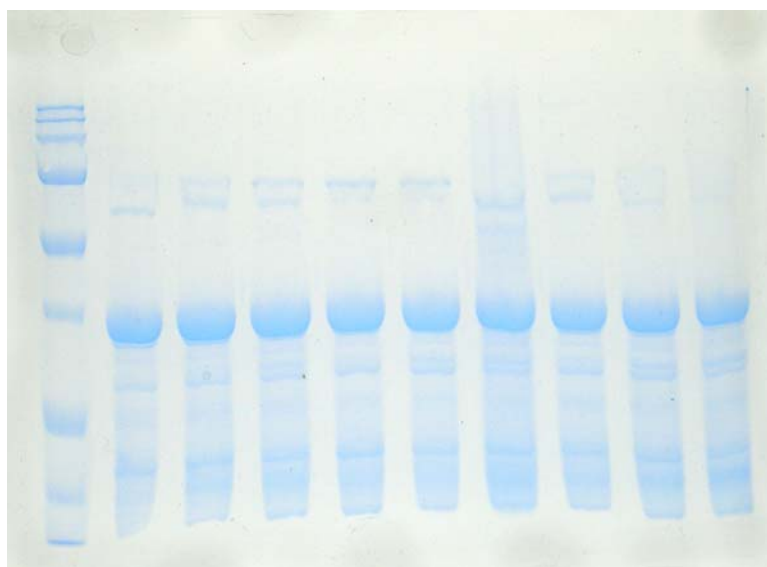
**Figure 2-26. Chromatogram overlay of sonication method (2x1). Both analyses were carried out in Tris buffer. Again, there is an indication of no PI giving the best protein recovery.**

After being unsuccessful with finding the cause of the degradation bands, a search was done to try to find other explanations than the ones explored. An article was located that described the boiling process in the preparation of the samples for the SDS-PAGE gels to be the root cause of degradation<sup>48</sup>. This led to an investigation of various boiling times and temperatures. Six different heating temperatures were investigated, as were three different heating times. This led to eighteen separate samples which were analyzed with SDS-PAGE gels. This analysis finally gave an answer to the cause of the degradation bands seen on the gels. All three times taken at 60°C showed no degradation bands. These bands could be seen faintly at all three heating times for 70°C and 80°C. The analysis of the samples taken at 90°C, 95°C, and 100°C showed a profound and almost equal degradation of the protein sample (See Figure 2-27 and 2-28). This means that the protein sample itself is extremely pure and is not degraded. The apparent cleavage occurs during sample preparation.

Even after obtaining pure P3 protein, there were problems noted with its stability. At lower concentrations, it would last for about a month at 4°C before “crashing out”, i.e. precipitate out as a white solid. When the protein was stored at higher concentrations, the length of time was shortened to about two weeks. It was found that the addition of 2% ethylene glycol to the Tris buffer lengthened the P3 solution lifetime from two weeks to three. Therefore, it became necessary to only make small amounts of P3 protein solutions at a time, but more frequently. This instability is in contrast to the PRD1 P3 protein which was determined to be much more stable than Bam35 P3 in previous work.



**Figure 2-27. SDS-PAGE analysis of different heating times at 60, 70, and 80°C. The first lane contains the MW standard. Lanes 2-4 contain samples heated at 60°C for 2, 5, and 10 minutes, respectively. Lanes 5-7 contain samples heated at 70°C for the three trial times. The last three lanes are the samples heated at 80°C for the three trial times. The 60°C samples show no degradation bands. The 70°C samples show the beginnings of the cleaved bands and they can be seen more readily in the 80°C sample lanes. This indicates that the P3 purified solution is indeed pure and free of degradation products. These are only seen because of the way the samples are prepared for SDS-PAGE.**

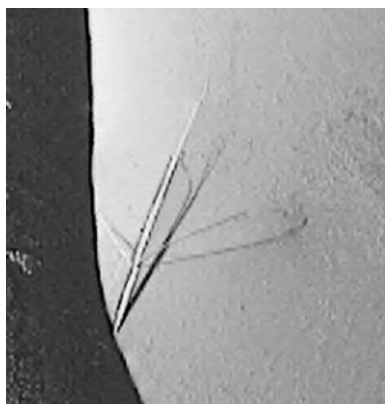


**Figure 2-28. SDS-PAGE analysis of different heating times at 90, 95, and 100°C. The first lane contains the MW standard. Lanes 2-4 contain samples heated at 90°C for 2, 5, and 10 minutes, respectively. Lanes 5-7 contain samples heated at 95°C at all three trial times. The last three lanes are the samples heated at 100°C for all three trial times. Degradation bands are very evident in all samples viewed on this gel.**

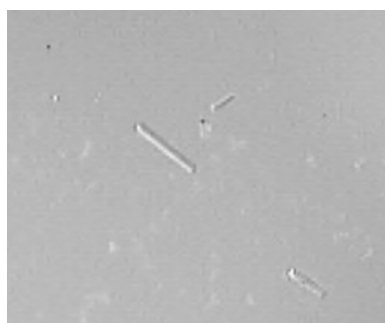
The increased stability for PRD1 is indicated to be from its host preference, thus allowing it to exist in more hostile environments<sup>27</sup>.

### **Crystallization:**

We used previously known crystallization conditions, discovered by our collaborators, in our beginning stages of crystallization trials for Bam35 P3 (given at the beginning of the next paragraph). In the initial stages of the crystallization trials, 7 mg/mL protein concentration was used. This produced crystals that were very thin and shaped like plates (see Figure 2-29). A lower concentration, 5 mg/mL, was tried to see if crystal shape and quality could be improved. It was hypothesized that lower concentrations might reduce the amount of nucleation leading to fewer, more singular crystals. However, very poor results were noted. There was mostly precipitate in the hanging drop instead of crystals. A higher concentration was attempted by increasing it to 10 mg/mL. The thinking here was that the crystals were being starved of material and thus not continuing to form in the same direction, thus leading to twinning. However, if the concentration was increased, the nucleation could continue in the same order producing singular crystals. Fewer crystals were produced than with the 7 mg/mL concentration, and there was a marked improvement in crystal shape. The crystals were much thicker and in the shape of a 3-D prism (See Figure 2-30).



**Figure 2-29.** Crystal formed with 7 mg/mL protein concentration. Notice the thinness of the crystals. (There are actually three crystals viewed here. Two stacked on top of each other and one protruding to the right.)



**Figure 2-30.** Crystal formed with 10 mg/mL protein concentration. Notice that these are rod shaped and are 3-dimensional, unlike those seen in Figure 2-29.

Two initial conditions showed promise in the formation of P3 crystals: 8.5% polyethylene glycol (PEG) 8K, 50 mM sodium acetate, 50 mM potassium phosphate (pH 7.0); and 8.0% PEG 8K, 55 mM sodium acetate, and 25 mM sodium cacodylate (pH 7.0). With these conditions not only are crystals formed, but they appear rapidly, usually within 24 hours or less. After performing trials where concentrations of all species present were manipulated and pH ranges were adjusted, I found the second condition to be best. One problem that occurred from the rapid formation of the P3 crystals was a showering of crystals. In other words, a great number of crystals formed, but they were

too small to analyze. Also, as mentioned earlier, they were very thin and fragile. Additionally, the crystals also degraded soon after forming. It was proposed that slowing down the nucleation process would allow for better crystal formation. Therefore, trials were performed at cooler temperatures (4°C and 11°C) than room temperature. However, these trials never produced crystals, instead eventually precipitation occurred within the drop.

It was decided that a sparse matrix crystal screen would be done to see if there were any other possible conditions for crystal formation. The crystal screens were purchased from Hampton Research and contain 48 to 50 conditions that have been taken from publications as popular and productive crystallization conditions<sup>44, 45</sup>. After performing this trial, I did discover a couple of new conditions that formed crystals. After further work on these only one condition ended up working well, 10% PEG 8K, 8% ethylene glycol (EG), and 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5). It was decided to screen this condition with different salts to see if crystallization could be enhanced. Of the ones tried, two worked well, sodium chloride and sodium citrate. The condition that worked well with the sodium chloride was 10% PEG 8K, 8% EG, 80 mM HEPES (pH 8.0), 80 mM NaCl. With the addition of sodium citrate, the condition that worked well was 10% PEG 8K, 8% EG, 80 mM HEPES (pH 8.0), 10 mM sodium citrate. Using these two conditions, thicker crystal were produced than had been seen previously. However, when in trying to diffract these crystals, the diffraction pattern suggested that they were twinned. A common type of crystal twinning occurs when two different crystals share some of the same crystal lattice points in an

equal manner. The result is an intergrowth of the two crystals in a variety of specific arrangements. Twinning can also occur when the plates lay down on top of each other, but are not adequately aligned. The latter was the most likely cause of the twinning observed for P3.

The best condition that repeatedly gave good crystal formation seemed to be 8.0% PEG 8K, 55 mM sodium acetate, and 25 mM sodium cacodylate (pH 7.0). We were able to collect diffraction data on these crystals, as well as those from some of the other conditions.

Currently, the Index Screen<sup>49-51</sup> (96 conditions), purchased from Hampton Research, is being used to try and determine new and better crystallization conditions. This screen is a combination of several screens and uses classic, contemporary, and new reagents<sup>49-51</sup>. Currently, there have been several conditions which have caused crystallization to occur. Two of the best are 12% PEG 3350, 0.1 M HEPES (pH 7.5), 0.005 M CoCl<sub>2</sub>, 0.005 M NiCl<sub>2</sub>, 0.005 M CdCl<sub>2</sub>, 0.005 M MgCl<sub>2</sub> and 10% PEG MME 5000, 0.1 M HEPES (pH 7), 5% Tacsimate (pH 7). Additional reagents are being purchased and further trials will continue in this area.

### **X-ray Diffraction:**

Early on a rod shaped crystal, which formed from the sodium cacodylate conditions, was prepared for analysis with the X-ray diffractometer. The crystal, along with some of the mother liquor, was placed into a capillary tube, sealed with bee's wax,

aligned into the X-ray beam, and run at room temperature on the diffractometer. It diffracted to a moderate resolution of 3.5 Å. The space group was determined to be orthorhombic ( $P2_12_12_1$ ). The unit cell was determined to have sides of  $a = 128.04$  Å,  $b = 146.44$  Å, and  $c = 168.42$  Å. Due to the space group determined, this makes all angles ( $\alpha$ ,  $\beta$ , and  $\gamma$ ) 90°. However, because the data was taken at room temperature and the crystal was not protected from the intense radiation, the crystal deteriorated before complete data could be collected. It was at this point that it was decided that cryo temperatures were needed to preserve any future crystals for analysis. Cryo-preservation would allow for the crystal to be stabilized and to be protected from the radiation, which dramatically extends its lifetime (essentially being immortal on laboratory X-ray sources). This allows for longer data collection times.

After the nitrogen cryo stream was in place, crystals were looped up, placed under the nitrogen stream, and analyzed initially to see if any diffraction could be detected. Before being placed in the nitrogen stream, the crystals were introduced to a cryo-preservant, usually the mother liquor with a small percentage (~10%) of ethylene glycol added. This allows for the formation of vitreous ice, not crystalline ice, which will not interfere with diffraction. If decent diffraction was achieved (i.e. diffraction out to at least 3 Å, no obvious twinning, good intensity), then the crystal was further analyzed to obtain data for indexing. Indexing gives preliminary information on the crystal's space group and unit cell. If this information is adequate, then further analysis can be done to determine more about the crystal structure in detail.

A crystal which was grown with 10% PEG 8K, 8% EG, 80 mM HEPES pH 8.0, and 80 mM NaCl was analyzed. Diffraction was sufficient and a run was performed to allow for indexing. This indicated that the sides were  $a = 67.658 \text{ \AA}$ ,  $b = 121.072 \text{ \AA}$ , and  $c = 123.851 \text{ \AA}$ . The angles are  $\alpha = 112.026^\circ$ ,  $\beta = 104.195^\circ$ , and  $\gamma = 103.529^\circ$ . The space group was determined as triclinic P1. After indexing, the prediction of where diffraction spots should be located did not identify all the spots that were observed. The confidence in this unit cell determination was not high. We were unable to collect further data from this crystal and another crystal from the same well did not show diffraction. Another well was looked at, with a different set of conditions, 10% PEG 8K, 8% EG, 20 mM HEPES pH 8.0, and 10 mM NaCl. A crystal from this well gave diffraction and was able to be indexed; however, again with not much confidence. The information determined from it gave side lengths of  $a = 65.51 \text{ \AA}$ ,  $b = 123.4 \text{ \AA}$ , and  $c = 124.3 \text{ \AA}$ , angles of  $\alpha = 119.75^\circ$ ,  $\beta = 101.8^\circ$ , and  $\gamma = 98.92^\circ$ , and a triclinic P1 space group. Further analysis of this crystal was initiated, but was unable to be completed due to low intensity of the spots. Later, a crystal that was formed from 10% PEG 8K, 8% EG, 80 mM HEPES pH 8.0, and 80 mM NaCl was analyzed again. However, the drop contained 4  $\mu\text{L}$  of protein solution, rather than the standard 2  $\mu\text{L}$ , and the well contribution was left at 2  $\mu\text{L}$ . The diffraction test indicated that the crystal was twinned, but further analysis was conducted anyway. Indexing showed  $a = 213.5 \text{ \AA}$ ,  $b = 124.1 \text{ \AA}$ , and  $c = 299.9 \text{ \AA}$  side lengths,  $\alpha = 90^\circ$ ,  $\beta = 113.99^\circ$ , and  $\gamma = 90^\circ$  angles, and monoclinic C for the space group. Another crystal grown with 10% PEG 8K, 8% EG, 80 mM HEPES pH 8, 10 mM NaCl gave indexing information of  $a = 68.6 \text{ \AA}$ ,  $b = 125.7 \text{ \AA}$ , and  $c = 206.4 \text{ \AA}$  and  $\alpha = \beta = \gamma = 90^\circ$  with a space group of orthorhombic P. We have high confidence in the unit cell and

space group determination of the capillary mounted crystal. However, there is much less confidence in the determinations from all the other crystals. This is due to the fact that the predicted locations of the diffraction patterns did not cover all the collected diffraction spots.

I was finally able to collect information about two crystals, besides the earlier capillary mounted crystal, beyond the indexing stage. This included using two programs to further analyze the crystal data, SAINT-NT and XPREP. SAINT-NT<sup>46</sup> is a data processing program that converts raw crystallographic frame data into integrated intensity sets with standard deviations, direction cosines, and XYZ centroids. It has the ability to do spatial correction in the frames, background subtraction, and error analysis, among others. XPREP<sup>47</sup> is a program that performs an analysis and manipulation of intensity data. The operations include space group determinations, intensity statistics, scaling and merging of different data sets, derivation and filtering of anomalous differences, as well as several others. It can be used to determine a reduced primitive unit cell, known as the Niggli cell. One of the crystals, produced with 10% PEG 8K, 8% EG, 80 mM HEPES pH 8.0, and 10 mM sodium citrate, was indicated by XPREP as having  $a = 125.2 \text{ \AA}$ ,  $b = 216.7 \text{ \AA}$ , and  $c = 161 \text{ \AA}$  and  $\alpha = \gamma = 90^\circ$ ,  $\beta = 100.77^\circ$ . Monoclinic C was determined to be the space group. The other crystal, which was formed by 10% PEG 8K, 8% EG, 80 mM HEPES pH 8.0, and 80 mM NaCl, gave indexing in XPREP as  $a = 69.085 \text{ \AA}$ ,  $b = 121.554 \text{ \AA}$ , and  $c = 122.545 \text{ \AA}$ ,  $\alpha = 60.48^\circ$ ,  $\beta = 76.60^\circ$ , and  $\gamma = 81.53^\circ$ , and triclinic P1 for the space group.

The Matthew's coefficient,  $V_m$ , allows for the determination of the number of molecules within the asymmetric unit by using a simple calculation.

$$V_m = \text{volume of unit cell} / (MW * A * X)$$

MW= molecular weight in daltons

A= the number of asymmetric units in the unit cell (the number of symmetry operators in the predicted space group)

X= the number of molecules in the asymmetric unit

The most probable values of X are those that give a  $V_m$  output that falls within the expected range of 1.68 to 3.53 Å<sup>3</sup>/Dalton. The expected solvent content, calculated as  $1 - 1.23/V_m$ , should lie in the range of 27-65%<sup>52</sup>. However, there certainly are exceptions to these ranges.

Using Matthew's coefficient, it was determined that for the capillary mounted crystal, the  $V_m = 3.31$  Å<sup>3</sup>/Dalton and 63% solvent for three trimers within the asymmetric unit. Repeating the calculation for two trimers,  $V_m = 2.48$  Å<sup>3</sup>/Dalton and 50% solvent. Three trimers within the asymmetric unit seems probable, but two molecules can not be ruled out. The monoclinic C space group had a  $V_m = 3.00$  Å<sup>3</sup>/Dalton and 59% solvent for three trimers within the asymmetric unit. Two trimers gives a  $V_m = 4.50$  Å<sup>3</sup>/Dalton and 73% solvent. Therefore, this crystal most likely has three trimers within each asymmetric unit. The second crystal, space group of triclinic P1, gave a  $V_m = 3.65$  Å<sup>3</sup>/Dalton and 66% solvent for two trimers. A second calculation with three trimers gave a  $V_m = 2.44$  Å<sup>3</sup>/Dalton and 50% solvent content. Again, three

trimers in the asymmetric unit seemed most probable, but two molecules could not be ruled out.

### **Conclusion:**

I have been able to obtain a highly pure protein solution in adequate quantities, due to several findings. First, a reproducible, reliable medium was needed to produce large amounts of bacteria, which would in turn produce sufficient quantities of P3 protein. Through the course of several experiments, this was determined to be the combination of the individual components, tryptone, yeast extract, and sodium chloride, with no glucose addition needed and the solution needed to be autoclaved. Protein degradation appeared to be a problem in the purified protein solution. Attempts to correct this situation proved fruitless. However, the finding that indeed the problem was due to SDS-PAGE preparation and not in the solution itself was very exhilarating. Once we knew for sure that we were producing an extremely pure protein solution, all efforts could be placed on crystal production. So far, crystal production has been both encouraging and frustrating. Even in initial crystallization trials, crystals have been produced and sometimes in large quantities. However, there have been obstacles that we have been working to overcome. Early crystals proved to be numerous, but very small. Even after conditions were improved to increase crystal size, they were too thin and produced very weak diffraction. Conditions were further improved to produce thicker crystals, but these proved to be twinned. The small amount of diffraction data has not been sufficient for further structure analysis. Current work with metal additions to crystallization conditions is proving to be quite promising. Weak diffraction was

detected in crystals taken from the crystal screen (i.e. have not been optimized), but was detected at fairly high resolution, around 3 Å. Continued work in optimizing these conditions or in determining other beneficial conditions will certainly lead to the final goal of structure determination.

## CHAPTER 3

### PRD1 P6

#### 3.1 Introduction

The discussion up to now has focused on structural studies of the major coat protein of Bam35 and related viruses. A major reason to examine the major coat protein is due to the fact that it is the largest part of the virion. However, additional valuable information could be acquired by investigating other viral structures for comparison purposes. One area that we have begun studying is the genome packaging machinery of PRD1. Bam35 and PRD1 both are suspected of having 11 regular vertices and one special vertex, used for packaging. One of the proteins making up this vertex, P9, has been known for years and if it is absent, empty capsids are formed. P9 functions as an ATPase and is involved in the active packaging of the genome. Though the knowledge of P9 has been around for quite some time, the unique vertex is a relatively new discovery only recently confirmed with immunogold labeling and mutational studies<sup>25, 53</sup> and warrants further investigation. As has been previously noted, much is now known about PRD1's structural, genetic, and biochemical properties. However, the special vertex has still not been visualized. This is due to the fact that previous X-ray crystallography and image reconstructions from cryo-EM for PRD1 used icosahedral

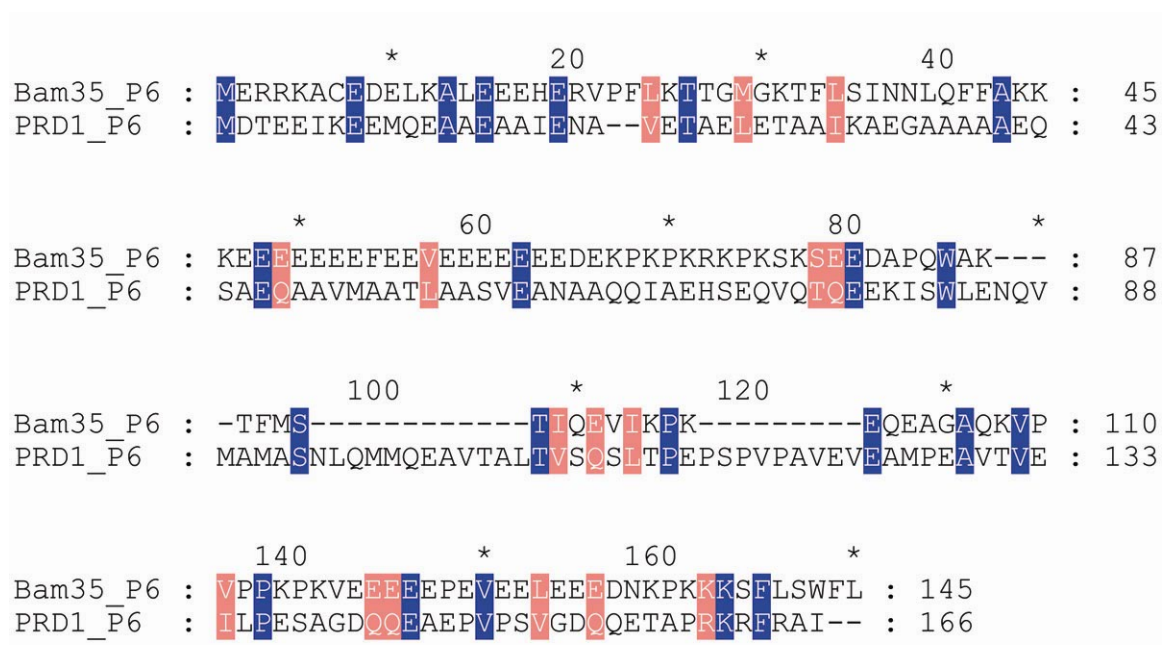
averaging for the electron densities, which would obscure any proteins not obeying the icosahedral symmetry<sup>23</sup>. This chapter looks at PRD1 and trying to isolate and crystallize one of the proteins making up the special vertex, P6, which was chosen due its ability to be over-expressed and its solubility. P9 is being cloned by our collaborator for future structural studies.

This special vertex of PRD1 is involved in genome packaging into the empty capsids and is comprised of four individual proteins: P6, a minor structural capsid protein; P9, the packaging ATPase; and the small membrane proteins P20 and P22. Though currently the exact role of P6 is undetermined, it has been found to play an important role in the makeup of the special vertex. A study has been performed using mutants that lacked the P20 or P22 membrane proteins. It was found that in this case the mutant lacked the P6 protein as well. In addition, it was noted that with the lack of P6 also came the lack of P9. This suggests that the membrane proteins are interacting via protein-protein interactions with P6, which then is associated with P9. This also leads to the conclusion that this is indeed a unique, portal vertex for PRD1 and contains all the proteins necessary for DNA packaging and preservation within the particle. This vertex also extends to the internal membrane, allowing for access to the genome of the phage<sup>25</sup>. Another study has investigated the membrane-associated protein P16<sup>54</sup>. When the PRD1 virion was determined by X-ray crystallography<sup>23</sup>, P16 was shown to be located beneath the other eleven vertices. When mutations are introduced that eliminate P16, the virion is not able to correctly assemble the spike complex (located at the other eleven vertices)

needed for infectivity, but the packaging complex is correctly assembled, further supporting the idea that there is a unique vertex.

It has been suggested that P6 plays a major role in DNA packaging. A mutant form of PRD1, free of any P6, has shown a severe deficit in the packaging of the phage DNA<sup>55</sup>. Viral genes, which code for similar protein functioning, occur in close vicinity on their genome. Gene *VI* codes for P6 and is located in front of gene *IX*, which codes for the known packaging ATPase P9<sup>55</sup>. This close genomic relationship therefore insinuates a close operational relationship.

A sequence alignment was performed with CLUSTAL X<sup>30</sup> to compare the two P6 proteins from Bam35 and PRD1. This information was then displayed with GeneDoc<sup>56</sup> and the two were determined to have 17% identity and 32% similarity (See Figure 3-1). These two percentages correspond very closely to the relationship of these two bacteriophages in respect to their P3 sequences; although, the P6 sequences could not be grounded on the basis of structure as the P3 sequences have. Structures of these two proteins are needed to confirm a similar fold. The ATPases for PRD1 and Bam35, along with several members of the proposed lineage, have also previously been compared. They show a similarity in the Walker A and Walker B motifs, which are characteristic ATP binding sequences, and show a unique motif characteristic to this group of viruses, labeled the P9-specific motif<sup>57</sup>. Thus indicating that similarity in structure extends beyond just the major capsid proteins.



**Figure 3-1. P6 sequence comparison for Bam35 and PRD1. The sequence alignment was done with CLUSTAL X and displayed with GeneDoc. The blue areas represent identity (17%) and the pink areas represent similarity (32%).**

The tailed bacteriophages have an obvious special vertex, but it was thought that most dsDNA eukaryotic viruses did not. It has been discovered that there is a unique vertex in Herpes simplex virus, HSV-1, which is a dsDNA virus infecting higher organisms<sup>25</sup>. DNA packaging is thought to be one of the components of the viral “self”, resistant to evolutionary modification. Therefore, it is of interest to extend our knowledge of viral protein coats, and their apparent similarities, and branch out to other structural properties, such as this special vertex. Could adenovirus, PBCV-1, PRD1, Bam35, and the other proposed members of this lineage also show familial relationships in this area as well? Sequence alignments with the ATPases from these viruses have shown the Walker A, Walker B, and P9-specific motif<sup>57</sup> signatures in all of them. Extensive research has been performed on the special vertex of HSV-1<sup>58</sup> and tailed

bacteriophages and strong similarities have been observed. However, these two show no structural similarity, other than possession of dsDNA genomes, with our proposed lineage. If the virus is prevented from packaging its DNA, this will hinder its ability to infect a host, thus preventing viral disease. If it is shown that there is a conservation of structure and function with this special vertex, a generalized mechanism could be discovered which could prevent its function and thus infectivity among any member of this virus lineage. This thesis is attempting to begin to put together the pieces of the unique vertex of PRD1. Because P6 currently remains a mystery as to its structure and true function, it was decided to try and crystallize and determine the structure of this protein.

### **3.2 Research Design**

#### **Bacteria Growth and Protein Over-Expression:**

For the growth of the PRD1 P6 protein, *E. coli* strain HMS174 (DE3) (pNS62) was used with the cloning vector being pJJ2 containing gene *VI* from PRD1 that was developed by our collaborator Dr. Jaana Bamford (University of Jyväskylä, Finland). A bacterial colony was mailed as a stab agar culture. A small amount was looped and spread onto an agar plate containing 50 µg/mL of ampicillin. A single colony was selected and placed in 2 mL of LB media. A glycerol stock was created by mixing the 2 mL aliquot with an equal volume of 70:30 LB media/glycerol. The mixture was vortexed, to ensure proper dispersion of the glycerol. This was dispensed into 1 mL aliquots and stored at -80°C. For additional bacterial growth for protein expression, a small amount was taken from the frozen stock solution and streaked on a Petri dish

containing agar and 50 µg/mL ampicillin. This dish was then incubated at 37°C for 24 hours.

A colony was then picked and placed into a 100 mL baffled bottom Nalgene flask. This flask contained 20 mL of Luria broth spiked with 150 µg/mL of ampicillin. The flask, as well as all others in the growing process, was capped and the cap was opened a quarter of a turn to allow for adequate oxygenation. This was then placed in the Barnstead International Max Q 5000 incubator/shaker and incubated overnight at 28°C and 200 rpm. The next day, the overnight culture was placed into a 1 L baffled bottom Nalgene or glass flask containing 480 mL of Luria broth spiked with 150 µg/mL of ampicillin. This was allowed to incubate at 28°C and 200 rpm until proper optical density was reached ( $OD_{600} \approx 0.8$ ).

IPTG (100 µM) was then added to induce the recombinant protein expression. The temperature and shaking were held constant at 28°C and 200 rpm overnight. The cells were pelleted (Beckman Coulter Allegra X-15R, 4500 rpm, 30 min, 4°C), equally distributed between two 50 ml conical tip tubes, and each was re-suspended in approximately 10 mL of 20 mM Tris-HCl (pH 7.2).

#### **Purification of Protein:**

The cells were sonicated using a Branson Sonifier 150 set at level three. The sample was sonicated for 15 seconds and allowed to rest on ice for 30 seconds before additional sonication. This procedure was repeated three more times. The cell debris

was centrifuged (Beckman Coulter Allegra X-15R, 4500 rpm, 40 min, 4°C) and the supernatant was decanted into a 15 ml conical vial tube. A sample was taken for SDS-PAGE. The supernatant was further clarified by centrifugation (Beckman Coulter Allegra X-15R, 4500 rpm, 2-17 hr, 4°C) and a sample was taken for SDS-PAGE.

The protein solution was then subjected to a salt cut using ammonium sulfate. Proteins will precipitate at varying concentrations of ammonium sulfate. First, a pre-cut was done with 22% ammonium sulfate. This was done by adding the necessary amount of grams to the given volume very slowly over a period of several minutes with vigorous stirring at room temperature. After all of the ammonium sulfate had dissolved, the solution was allowed to stir for ten minutes. At this point, the solution was centrifuged (Beckman Coulter Allegra X-15R, 4500 rpm, 15 min, 4°C) and decanted into a new 15 ml conical vial. Next, the ammonium sulfate was adjusted to 26% and stirred at room temperature for 10 minutes. This concentration had been shown to precipitate PRD1 P6. The protein precipitation was pelleted as before and the solution decanted. Samples of the supernatant and pellet were taken both times for SDS-PAGE. The second precipitate was dissolved in 20 mM piperazine (pH 5.5).

Next, a 5 mL Pharmacia HiTrap Q cation column was used. The column was equilibrated by applying 25 mL of 20 mM piperazine (pH 5.5) with a syringe into the column. The sample was applied in the same way. This column was then placed onto the BioLogic DuoFlow system. The buffers used were 40 mM piperazine (pH 5.5), 40 mM piperazine (pH 5.5)/600 mM NaCl, and de-ionized water. All buffers were degassed

prior to placing them into the system. The program was set to collect 1 mL fractions and run at pH 5.5 during the course of the run. First, an isocratic flow was used with 0% salt for a volume of 5 mL at 1 mL/min. Next, was a linear gradient of 33-100% (99 mM – 300 mM) salt which ran for a volume of 30 mL at 1 mL/min. A final wash of the column came with an isocratic flow of 0% salt for 10 mL at 1 mL/min. Fractions were taken, from the indicated collected tubes corresponding with the peak(s) on the chromatogram, for SDS-PAGE. Desired fractions were combined based on the findings of the SDS-PAGE gel.

A 15 mL Vivaspin concentrator tube (10,000 MWCO) was used to aid in buffer exchange from 20 mM piperazine (pH 5.5) to 20 mM Tris-HCl (pH 7.2) in the centrifuge (Beckman Coulter Allegra X-15R, 3050 rpm, 4°C) and ran until the solution was below a volume of 1 mL.

The concentrated solution was introduced into the 1 mL sample loop of the Bio-Rad DuoFlow. A gel filtration column was set up to analyze the solution. The buffers used were 40 mM Tris base, 40 mM Tris-HCl, 2 M NaCl, and de-ionized water, all previously de-gassed. Fractions of 1 mL were collected during the run, and the Tris-buffer was maintained at a pH of 7.2. An isocratic flow with 8% (80 mM) salt was used for 1 mL at a rate of 0.5 mL/min. This was followed by the injection of the sample from the loop for 2 mL at 0.5 mL/min. Another isocratic flow followed for 20 mL at 8% salt concentration at 0.5 mL/min. Fractions were taken from the indicated collected tubes

corresponding with the peak(s) on the chromatogram for SDS-PAGE. Desired fractions were combined based on the findings of the SDS-PAGE gel.

A concentration calibration curve of the Bradford Assay against BSA was performed according to the manufacturer's (Bio-Rad) instructions. This was done on a Beckman Coulter DU 530 UV/Vis Spectrophotometer at a wavelength of 595 nm. Then a 20  $\mu$ L aliquot of the purified PRD1 P6 protein was combined with 1 mL of Bradford Assay solution (at room temp) and was shaken vigorously in the cuvette. A blank solution, containing only Bradford Assay solution, was placed into another cuvette. The blank was placed into the sample holder of the DU 530 and was read. The sample cuvette was then placed into the sample holder and the absorbance value was obtained. This value was used to determine the concentration of protein in the solution by use of the previously derived calibration curve. Once the initial concentration was determined, a final volume was calculated to concentrate the sample down to achieve the desired concentration of 7 mg/mL. To concentrate the sample, the solution was placed into a 15 mL Vivaspin tube and centrifuged (Beckman Coulter Allegra X-15R, 3050 rpm, 4°C) until the volume desired was shown with the tube's graduations. This solution was then put in a microcentrifuge tube, labeled, and placed in storage at 4°C.

### **Crystallization:**

Crystallization was carried out in Hampton Research VDX trays, which are modified 24-well cell culture trays. The hanging drop method was used with initial amounts of reservoir and protein solution in the drop being 2  $\mu$ L each. To set up a tray

for crystallization, the edges of the wells must be greased to allow for a closed system when sealed with a coverslip. The grease was applied by way of a syringe. Next, each reagent needed to create the reservoir solution is added one by one. The total volume of the well solution needed was 1 mL, so distilled water was added last to meet this volume requirement. The tray is lightly shaken to allow for adequate mixing of all chemicals. Then, 2  $\mu$ L of the reservoir solution is pipetted onto a plastic coverslip. Then, 2  $\mu$ L of the concentrated, purified protein solution is pipetted and mixed into the reservoir drop. The coverslip is then placed over the well and pressed down, so as to form a tight seal. All 24 wells were treated in the same manner, though each contained different quantities and identities of solutions, and the tray was set inside a cabinet to be isolated from sunlight.

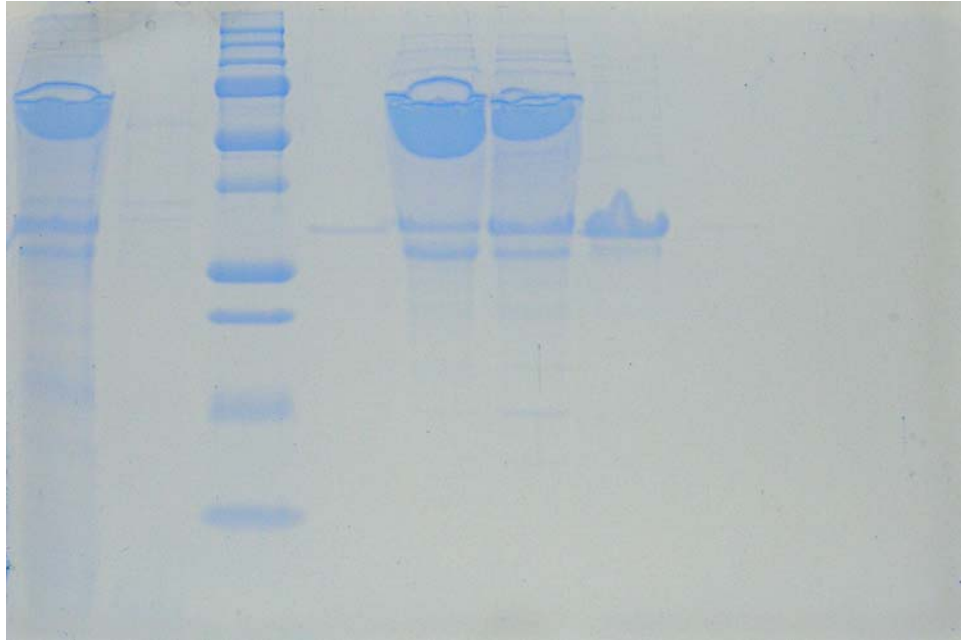
A reservoir solution is typically made up of a precipitant (usually an organic compound) of which there can be more than one, a buffer, and possibly a salt. To determine the identity and concentrations of reagents needed to crystallize the protein, two pre-made crystal screens were purchased from Hampton<sup>44, 45</sup>. The protein solution was introduced to drops of 98 different combinations of chemicals and checked periodically for crystal growth. Several different conditions were found to produce crystals. However, before more trials could be performed, MALDI-TOF results showed that we did not have the P6 protein present. There were eight bands analyzed with mass spectrometry, though some were duplicates from two separate gels. This led to a reevaluation of where the protein was, if it was present at all.

### **3.3 Results and Discussion**

#### **Purification and Isolation of protein:**

The P6 protein has a molecular weight of 17573.38 daltons. When the SDS-PAGE gels were run after protein purification from the Hi-Trap Q column and the gel filtration column, the gels both showed a band in the range of the expected molecular weight. However, this band was very light, indicating that hardly any protein was present. Another band appeared at approximately 34-35 kDa, which could indicate a dimer of P6. Even more interesting, a comparatively enormously intense band appeared at approximately 70 kDa. This could indicate a tetramer of our P6 protein. Several fainter bands appeared on the gels, but these three bands mentioned seemed to be the main proteins present (see Figures 3-2 and 3-3). If these bands were related to P6, it would appear that the P6 protein solution was much more pure than the Bam35 P3 solution had been in the early stages before identification of the protein cleavage caused by the sample preparation for SDS-PAGE analysis. However, if there really was a presence of a monomer, dimer, and tetramer, especially a tetramer in the amounts indicated, there might be problems in using this solution to induce crystallization. Additionally, this was an SDS-PAGE gel where the protein is run in a denatured state. An SDS-PAGE running buffer is added that contains SDS, which helps denature proteins, and 2-mercaptoethanol that would disrupt any disulfide bonds, even though PRD1 P6 does not contain any cysteine residues. The protein solution with the SDS-PAGE running buffer is boiled for 10 minutes to further denature the protein. We should expect to see the PRD1 P6 protein band corresponding to about 17.5 kDa on the gel, but the

previously run gels of purified protein from both intact viruses and recombinant proteins run around 37 kDa<sup>55</sup>.



**Figure 3-2. SDS-PAGE gel analysis of PRD1 P6. The first lane is a sampling of the pellet as it was re-suspended in 20 mM Tris-HCl pH 7.2, before it was purified. The second lane contains the Q column flow through (note in this case P6 does not show on the gel as it is stuck on the column). The third lane contains the MW standard. All remaining lanes correspond to samples taken from the gel filtration analysis performed in the Bio-Rad DuoFlow. Lanes 4-7 contain the peak thought to be P6 and the remaining lanes are from a second peak that appeared at the end of the run.**

It was necessary to find out for sure if these three bands obtained from our purified sample really indicated three different forms of P6 or if they are bacterial proteins. To do this, MADLI-TOF mass spectrometry was used on the trypsinized bands cut from the gels. The results from the spectrometric analysis told a much different story than what we had expected. None of the bands were associated with PRD1, much less the P6 protein. The analysis showed that though the scores were too low to be considered significant, three of the samples seemed to be related to the *E. coli* that our protein was



**Figure 3-3. SDS-PAGE analysis of P6 on the Hi-Trap Q column run on the Bio-Rad DuoFlow. The first lane is the MW standard. Lanes 2-6 are the peak thought to be P6. Lanes 7-10 are from a second peak.**

produced in. The other samples proved to be inconclusive. We then needed to find out why our protein was not being expressed.

### **Crystallization:**

All crystallization trials were performed before we knew the true identity of our protein solution. Hampton crystal screen 1 & 2<sup>44, 45</sup> indicated five different conditions in which the protein solution formed small crystals. One condition contained two different PEG components. The other four all were comprised of a different organic precipitant and salt, but all contained 0.1 M Tris (pH 8.5). Further trials were delayed until the identity of the bands from the mass spectrometric results was obtained. After learning that the P6 protein was not actually present, these crystallization trials were suspended.

## CHAPTER 4

### CONCLUSION

A highly pure protein solution is needed to obtain favorable and reproducible crystals. This has been achieved through several important discoveries. First, the most ideal media conditions were obtained through several experiments. We were able to ascertain that combining the individual components (tryptone, yeast extract, and sodium chloride) without the presence of glucose and autoclaving the resultant media solution gave the best bacterial growth, thus Bam35 P3 production. While there were initial fears of protein degradation in our working sample, it was determined that indeed this was not the case. This effort took much trial and error to obtain, but the end result has shown to be quite rewarding. The protein solution that is obtained for the purification of Bam35 P3 is not just mostly pure, but instead extremely pure. This leads into crystal production that is definitely P3 and not a mixture of other proteins. These results have been verified by the MALDI-TOF mass spectrometric analysis and the SDS-PAGE gels (with adequate heating times). Crystals have been obtained from these pure protein solutions. Crystal production has been achieved with great regularity, but with many problems encountered. In some cases they are too small to analyze and will even be present in dozens or hundreds in the drop. In other cases they are larger in size, but are too thin to diffract with much intensity. There have been some crystals that are larger

and more 3-dimensional, but they are usually twinned or show no or weak diffraction at low resolution. Trials are still being performed and will need to be expanded upon in the future to try and produce better crystals. Current work with divalent metal chloride additions in crystallization solutions has shown some promising preliminary results, meaning that the crystals have a nice shape and have shown weak diffraction, but to higher resolution (around 3 Å). Work in this area will continue. Finding the right crystallization conditions is all that stands in the way currently of obtaining a visualized structure of Bam35's capsid protein, P3.

Initial work with PRD1 P6 was slow and discouraging. However, with recent progress in expressing and purifying this protein, structure determination is not far off. Recently, an undergraduate in the lab, Edward T. El Rassi, with Dr. Benson have transformed P6 into another vector that inserts a histidine tag for purification. It has been confirmed that the cloning is correct and the protein is currently being over-expressed. The tag also allows for easy purification with a nickel affinity column. Previous crystal screens have shown some promising crystallization possibilities, but as was done with Bam35 P3, more extensive work will need to be performed to find the ideal crystallization conditions for the P6 protein.

Obtaining structures for these two proteins is the ultimate goal. However, this work will have to be carried out in the future. Preliminary data from Bam35 P3 does continue to point in the direction of putting it into the proposed lineage. Verifying the

structure of the PRD1 P6 protein will need to be followed up with other proteins of the special vertex of other viruses to see if the ancestral lineage continues to be validated.

The knowledge of structures, while certainly giving insight into functions, packaging, and assembly, would hopefully lead into a much broader advantage. That being the ability to use the knowledge of shared traits among viruses to interrupt their ability to package their nucleic acid and/or infect their hosts. Thus, being able to prevent disease in humans, animals, plants, and so forth. In essence, this would lead to the enrichment of all who inhabit our biosphere.

## REFERENCES

1. Virus (life science). Microsoft® Encarta® Online Encyclopedia. <http://encarta.msn.com>. 2007.
2. Flint, S. J.; Enquist, L. W.; Racaniello, V. R.; Skalka, A. M., *Principles of Virology: Molecular Biology, Pathogenesis, and Control of Animal Viruses*. 2nd ed.; ASM Press: 2004.
3. Chibani-Chennoufi, S.; Bruttin, A.; Dillmann, M. L.; Brussow, H., Phage-host interaction: an ecological perspective. *J. Bacteriol.* **2004**, 186, 3677-3686.
4. Brussow, H.; Hendrix, R. W., Phage genomics: small is beautiful. *Cell* **2002**, 108, 13-16.
5. Saren, A.-M.; Ravantti, J. J.; Benson, S. D.; Burnett, R. M.; Paulin, L.; Bamford, D. H.; Bamford, J. K. H., A snapshot of viral evolution from genome analysis of the *Tectiviridae* family. *J. Mol. Biol.* **2005**, 350, 427-440.
6. Van Regenmortel, M. H., Viruses are real, virus species are man-made, taxonomic constructions. *Arch. Virol.* **2003**, 148, 2481-2488.
7. <http://www.microbiologybytes.com/virology/Phages.html> Bacteriophages.
8. Lawrence, J. G.; Hatfull, G. F.; Hendrix, R. W., Imbrolios of viral taxonomy: genetic exchange and failings of phenetic approaches. *J. Bacteriol.* **2002**, 184, 4891-4905.
9. Pedulla, M. L.; Ford, M. E.; Houtz, J. M.; Karthikeyan, T.; Wadsworth, C.; Lewis, J. A.; Jacobs-Sera, D.; Falbo, J.; Gross, J.; Pannunzio, N. R.; Brucker, W.; Kumar, V.; Kandasamy, J.; Keenan, L.; Bardarov, S.; Kriakov, J.; Lawrence, J. G.; Jacobs, W. R., Jr.; Hendrix, R. W.; Hatfull, G. F., Origins of highly mosaic mycobacteriophage genomes. *Cell* **2003**, 113, 171-182.
10. Hendrix, R. W., Evolution: the long evolutionary reach of viruses. *Curr. Biol.* **1999**, 9, R914-R917.
11. Hendrix, R. W., Bacteriophages: evolution of the majority. *Theor. Popul. Biol.* **2002**, 61, 471-480.
12. Maidak, B. L.; Cole, J. R.; Lilburn, T. G.; Parker, C. T., Jr.; Saxman, P. R.; Farris, R. J.; Garrity, G. M.; Olsen, G. J.; Schmidt, T. M.; Tiedje, J. M., The RDP-II (Ribosomal Database Project). *Nucleic Acids Res.* **2001**, 29, 173-174.

13. Benson, S. D.; Bamford, J. K. H.; Bamford, D. H.; Burnett, R. M., Does common architecture reveal a viral lineage spanning all three domains of life? *Mol. Cell* **2004**, 16, 673-685.
14. Benson, S. D.; Bamford, J. K. H.; Bamford, D. H.; Burnett, R. M., The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 Å resolution. *Acta. Crystallogr.* **2002**, D58, 39-59.
15. Bamford, J. K.; Hanninen, A. L.; Pakula, T. M.; Ojala, P. M.; Kalkkinen, N.; Frilander, M.; Bamford, D. H., Genome organization of membrane-containing bacteriophage PRD1. *Virology* **1991**, 183, 658-676.
16. Benson, S. D.; Bamford, J. K. H.; Bamford, D. H.; Burnett, R. M., Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell* **1999**, 98, 825-833.
17. Xu, L.; Butcher, S. J.; Benson, S. D.; Bamford, D. H.; Burnett, R. M., Crystallization and preliminary X-ray analysis of receptor-binding protein P2 of bacteriophage PRD1. *J. Struct. Biol.* **2000**, 131, 159-163.
18. Xu, L.; Benson, S. D.; Butcher, S. J.; Bamford, D. H.; Burnett, R. M., The receptor binding protein P2 of PRD1, a virus targeting antibiotic-resistant bacteria, has a novel fold suggesting multiple functions. *Structure* **2003**, 11, 309-322.
19. Butcher, S. J.; Bamford, D. H.; Fuller, S. D., DNA packaging orders the membrane of bacteriophage PRD1. *EMBO J.* **1995**, 14, 6078-6086.
20. San Martin, C.; Huiskonen, J. T.; Bamford, J. K. H.; Butcher, S. J.; Fuller, S. D.; Bamford, D. H.; Burnett, R. M., Minor proteins, mobile arms and membrane-capsid interactions in the bacteriophage PRD1 capsid. *Nat. Struct. Biol.* **2002**, 9, 756-763.
21. San Martin, C. S.; Burnett, R. M.; de Haas, F.; Heinkel, R.; Rutten, T.; Fuller, S. D.; Butcher, S. J.; Bamford, D. H., Combined EM/X-ray imaging yields a quasi-atomic model of the adenovirus-related bacteriophage PRD1 and shows key capsid and membrane interactions. *Structure* **2001**, 9, 917-930.
22. Cockburn, J. J.; Abrescia, N. G.; Grimes, J. M.; Sutton, G. C.; Diprose, J. M.; Benevides, J. M.; Thomas, G. J., Jr.; Bamford, J. K. H.; Bamford, D. H.; Stuart, D. I., Membrane structure and interactions with protein and DNA in bacteriophage PRD1. *Nature* **2004**, 432, 122-125.

23. Abrescia, N. G.; Cockburn, J. J.; Grimes, J. M.; Sutton, G. C.; Diprose, J. M.; Butcher, S. J.; Fuller, S. D.; San Martin, C.; Burnett, R. M.; Stuart, D. I.; Bamford, D. H.; Bamford, J. K. H., Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* **2004**, 432, 68-74.
24. Rydman, P. S.; Caldentey, J.; Butcher, S. J.; Fuller, S. D.; Rutten, T.; Bamford, D. H., Bacteriophage PRD1 contains a labile receptor-binding structure at each vertex. *J. Mol. Biol.* **1999**, 291, 575-587.
25. Strömsten, N. J.; Bamford, D. H.; Bamford, J. K. H., The unique vertex of bacterial virus PRD1 is connected to the viral internal membrane. *J. Virol.* **2003**, 77, 6314-6321.
26. Laurinmäki, P. A.; Huiskonen, J. T.; Bamford, D. H.; Butcher, S. J., Membrane proteins modulate the bilayer curvature in the bacterial virus Bam35. *Structure* **2005**, 13, 1819-1828.
27. Ravantti, J. J.; Gaidelyte, A.; Bamford, D. H.; Bamford, J. K. H., Comparative analysis of bacterial viruses Bam35, infecting a gram-positive host, and PRD1, infecting gram-negative hosts, demonstrates a viral lineage. *Virology* **2003**, 313, 401-414.
28. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc. Acids Res.* **1997**, 25, 3389-3402.
29. Kelley, L. A.; MacCallum, R. M.; Sternberg, M. J., Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **2000**, 299, 499-520.
30. Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G., The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nuc. Acids Res.* **1997**, 25, 4876-4882.
31. Nandhagopal, N.; Simpson, A. A.; Gurnon, J. R.; Yan, X.; Baker, T. S.; Graves, M. V.; Van Etten, J. L.; Rossmann, M. G., The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, 99, 14758-14763.
32. Khayat, R.; Tang, L.; Larson, E. T.; Lawrence, C. M.; Young, M.; Johnson, J. E., Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 18944-18949.

33. Athappilly, F. K.; Murali, R.; Rux, J. J.; Cai, Z.; Burnett, R. M., The refined crystal structure of hexon, the major coat protein of adenovirus type 2, at 2.9 Å resolution. *J. Mol. Biol.* **1994**, 242, 430-455.
34. Rux, J. J.; Kuser, P. R.; Burnett, R. M., Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.* **2003**, 77, 9553-9566.
35. Lewin, R., Fish to bacterium gene transfer. *Science* **1985**, 227, 1020.
36. Bannister, J. V.; Parker, M. W., The presence of a copper/zinc superoxide dismutase in the bacterium *Photobacterium leiognathi*: a likely case of gene transfer from eukaryotes to prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, 82, 149-152.
37. Bamford, D. H., Do viruses form lineages across different domains of life? *Res. Microbiol.* **2003**, 154, 231-236.
38. Gaidelyte, A.; Cvirkaitė-Krupovic, V.; Daugelavičius, R.; Bamford, J. K. H.; Bamford, D. H., The entry mechanism of membrane-containing phage Bam35 infecting *Bacillus thuringiensis*. *J. Bacteriol.* **2006**, 188, 5925-5934.
39. Bamford, D. H.; Caldentey, J.; Bamford, J. K. H., Bacteriophage PRD1: a broad host range DSDNA tectiviruses with an internal membrane. *Adv. Virus Res.* **1995**, 45, 281-319.
40. Feng, D. F.; Cho, G.; Doolittle, R. F., Determining divergence times with a protein clock: update and reevaluation. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, 94, 13028-13033.
41. Ackermann, H.-W.; Roy, R.; Martin, M.; Murthy, M. R. V.; Smirnov, W. A., Partial characterization of a cubic *Bacillus* phage. *Can. J. Microbiol.* **1978**, 24, 986-993.
42. Harrison, S. C.; Olson, A. J.; Schutt, C. E.; Winkler, F. K.; Bricogne, G., Tomato bushy stunt virus at 2.9 Å resolution. *Nature* **1978**, 276, 368-373.
43. Stirk, H. J.; Woolfson, D. N.; Hutchinson, E. G.; Thornton, J. M., Depicting topology and handedness in jellyroll structures. *FEBS Letters* **1992**, 308, 1-3.
44. Cudney, R.; Patel, S.; Weisgraber, K.; Newhouse, Y.; McPherson, A., Screening and optimization strategies for macromolecular crystal growth. *Acta crystallograph.* **1994**, D50, 414-423.

45. Jancarik, J.; Kim, S.-H., Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **1991**, 24, 409-411.
46. *SAINT-NT*, 6.0; Bruker AXS: Madison, WI, 2003.
47. Sheldrick, G. M. *XPREF*, 6.12; Bruker AXS: Madison, WI, 2001.
48. Cookson, E. J.; Beynon, R. J., Degradation artefacts during sample preparation for sodium dodecyl sulphate polyacrylamide gel electrophoresis. *Biosci. Rep.* **1987**, 7, 209-215.
49. D'Arcy, A.; Mac Sweeney, A.; Stihle, M.; Haber, A., The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. *Acta crystallograph.* **2003**, D59, 396-399.
50. Simanshu, D. K.; Murthy, M. R., Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of propionate kinase (TdcD) from *Salmonella typhimurium*. *Acta crystallograph.* **2005**, D61, 52-55.
51. Fu, T. M.; Zhang, X. Y.; Li, L. F.; Liang, Y. H.; Su, X. D., Preparation, crystallization and preliminary X-ray analysis of the methionine synthase (MetE) from *Streptococcus mutans*. *Acta crystallograph.* **2006**, D62, 984-985.
52. Matthews, B. W., Solvent content of protein crystals. *J. Mol. Biol.* **1968**, 33, 491-497.
53. Gowen, B.; Bamford, J. K. H.; Bamford, D. H.; Fuller, S. D., The tailless icosahedral membrane virus PRD1 localizes the proteins involved in genome packaging and injection at a unique vertex. *J. Virol.* **2003**, 77, 7863-7871.
54. Jaatinen, S. T.; Viitanen, S. J.; Bamford, D. H.; Bamford, J. K. H., Integral membrane protein P16 of bacteriophage PRD1 stabilizes the adsorption vertex structure. *J. Virol.* **2004**, 78, 9790-9797.
55. Karhu, N. J.; Ziedaite, G.; Bamford, D. H.; Bamford, J. K. H., Efficient DNA packaging of bacteriophage PRD1 requires the unique vertex protein P6. *J. Virol.* **2007**, 81, 2970-2979.
56. Nicholas, K. B.; Nicholas H.B. Jr.; Deerfield, D. W. I., GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW.NEWS* **1997**, 4, 14.

57. Strömsten, N. J.; Bamford, D. H.; Bamford, J. K. H., In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses. *J. Mol. Biol.* **2005**, 348, 617-629.
58. Newcomb, W. W.; Juhas, R. M.; Thomsen, D. R.; Homa, F. L.; Burch, A. D.; Weller, S. K.; Brown, J. C., The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J. Virol.* **2001**, 75, 10923-10932.

## VITA

Julia Megan Beshirs

Candidate for the Degree of

Master of Science

Thesis: VIRAL SELF COMPONENTS OF THE *TECTIVIRIDAE* FAMILY

Major Field: Chemistry

Biographical:

Personal Data: Born in Durant, OK on January 7, 1977

Education: Received Bachelor of Science degree in Chemistry from Southeastern Oklahoma State University in May 2001. Completed the requirements for the Master of Science degree in Chemistry at Oklahoma State University in July 2007.

Experience: Worked as a laboratory assistant as an undergrad at SOSU. Worked for ManTech Environmental as an Associate Scientist from May 2001 to June 2002. Worked as Laboratory/Stockroom Manager for SOSU from July 2002 to June 2004. Worked as a Teaching Assistant for OSU from August 2004 to July 2007.

Name: Julia Megan Beshirs

Date of Degree: July 2007

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: VIRAL SELF COMPONENTS OF THE *TECTIVIRIDAE* FAMILY

Pages in Study: 76

Candidate for the Degree of Master of Science

Major Field: Chemistry

Scope and Method of Study: One purpose of this study was to isolate, purify, and crystallize the major capsid protein, P3, for bacteriophage Bam35, which belongs to the *Tectiviridae* family. The diffraction data from a stable, well developed crystal is needed to determine the structure of P3. The hope is that this structure can confirm a model, produced earlier, and to further confirm Bam35's place in a double trimer viral lineage which has begun to emerge with several other viruses from all three domains of life: Bacteria, Archaea, and Eukarya. Another aspect of this study was to return to *Tectiviridae*'s type species, PRD1, and branch out to other vital viral structures, such as its special vertex, which is used for DNA packaging. One of the four proteins comprising this vertex, P6, was focused on to isolate, purify, and crystallize. Further structures would have to be determined from this arena for comparison of the viral self components. The viral self is an explanation of how viruses can diverge, rather than converge, to share extremely similar structural aspects, but have vastly different genome sequences and host preferences. Being able to solve these structures will allow for the confirmation of the divergence theory and an affirmation of the double barrel trimer viral lineage.

Findings and Conclusions: A very pure solution of Bam35 P3 has been obtained. Crystallization has been productive as well and some initial diffraction data has been collected. However, ongoing trials continue to try and determine optimal crystallization conditions for this protein. Early work with PRD1 P6 proved fruitless, as it was discovered that the protein was not being produced. Further efforts have led to the successful production of this protein and crystallization trials have begun to find working conditions for crystal production.

ADVISER'S APPROVAL: Dr. Stacy Benson

---