# Biochemical studies of Ea22, a protein involved in the bacteriophage lysogenic-lytic decision and the virulence of clinically relevant serotypes of *E. coli*

**by**

**Jinge Tong**

A Thesis Submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science

Graduate Program in Biology
York University
Toronto, Ontario

# TABLE OF CONTENTS

## Abstract

Enterohemorrhagic *Escherichia coli* (EHEC) infections are associated with short term gastrointestinal problems, but in some cases, it can lead to more life-threatening conditions. The pathogenicity of EHEC Is supported by an endogenous, dormant bacteriophage that persists in the bacterium as a lysogen until it is activated by a stress-related event. While the lysogenic-lytic development decision has been characterized for decades, there are still genes that are unknown in every way, from structure to function. These genes are found in a region between *exo* and *xis* in the l bacteriophage genome which may serve as a model for understanding other related phages in EHEC. This thesis focuses on the protein encoded by the *ea22* gene, the largest conserved gene in the *exo-xis* region. The most significant finding of my research was the expression and characterization of several Ea22 proteins and protein fragments from l and two phages associated with EHEC. Structural differences in the C-terminal domains of these proteins may help explain the broad range of effects the Ea22 proteins mediate during infection.

# Acknowledgements

I would like to thank my supervisory committee who were more than generous with their expertise and precious time. A special thanks to Dr. Logan Donaldson, my supervisor for his consistent support, great patience, and countless hours of proofreading throughout the entire process. Thank you, Dr. Philip Johnson, for serving on my committee and inspiring me with your feedback. Your willingness to provide guidance and encouragement made the completion of my projects an enjoyable experience.

I appreciate that my loving parents Feng and Bin support me in every decision I make. Their words of encouragement, care, and love make me feel like the luckiest person in the world. I also would like to thank my sister Yvonne who is my best friend and my strongest motivation to become a better person.

I also dedicate this dissertation to many friends on the 3rd floor of LSB: Daria Taskina, Nick Bragagnolo, Christina Rodriguez, Wen Yi, Gaby Gr, Naveed Samari Kermani, and Ramsha who shared new insights and techniques with me, spent many hours of proofreading for me, kept me company and brought so much joy in my life.

# Abbreviations

CD.....Circular Dichroism
DSC.....Differential Scanning Calorimetry
EHEC..... Enterohemorrhagic *Escherichia coli*
ESI.....Electrospray Ionization Mass Spectrometry
HSQC..... Heteronuclear Single Quantum Coherence
HUS..... Hemolytic Uremic Syndrome
NMR.....Nuclear Magnetic Resonance
ORF.....Open Reading Frame
PEG.....Polyethylene Glycol
SEC-MALS.....Size Exclusion Chromatography with Multi-Angle Laser Scattering
SONICC.....Second-Order Nonlinear optical Imaging of Chiral Crystals
STEC.....Shiga Toxin producing *Escherichia coli*
Stx.....Shiga toxin / Shiga toxin-converting
TAPS..... Tris(hydroxymethyl)methylamino] propane sulfonic acid

# Table of Contents

# List of Tables



# List of Figures

# Chapter One — Introduction

## 1.1 EHEC outbreaks throughout the world

*Escherichia coli* is a common and benign resident of the human digestive system. There are, however, other strains of *E. coli* that have gained additional genetic elements giving them pathogenic properties. Infections of enterohemorrhagic *E. coli*, O157:H7 and O104:H4 [1], occur by consumption of contaminated food or water with cattle as serving as a reservoir or intermediate host [2]. Approximately twenty years ago, two thousand people in Walkerton, Ontario became seriously ill from drinking water contaminated with *E. coli* O157:H7 [3]. A year later, an outbreak of *E. coli* O104:O4 in Germany led to four thousand infections and fifty-three deaths [4]. Of these cases, eight hundred people were afflicted with hemolytic uremic syndrome (HUS), a complication that may lead to death. Much of the pathophysiology of HUS is due to the release of Shiga toxin by dying bacteria that in turn, kills surrounding tissues and blood vessels [5]. The gene encoding Shiga toxin does not belong to the bacterial genome; rather, it is part of a bacteriophage that has integrated into the bacterial genome. These endogenous bacteriophages, belonging to a class that includes I , are activated when

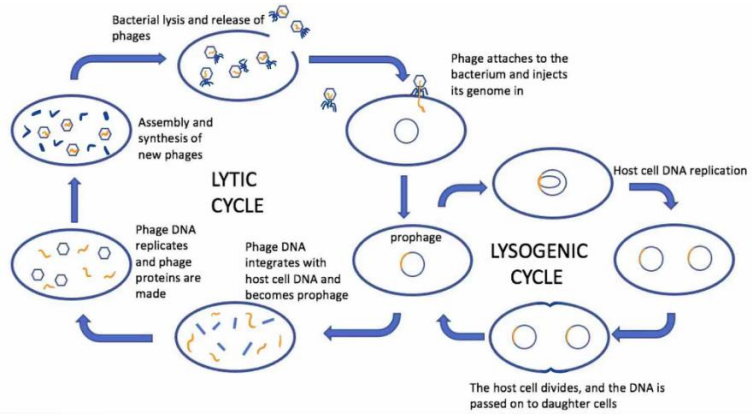



Figure 1.1
Bacteriophage lysogenic and lytic cycle. Changes were made to *Lytic cycle, compared to lysogenic cycle* by Suly12 is licensed under the CC BY-SA 3.0v

EHEC are exposed to stressors such as immune system attack or the application of antibiotics. The objective of my thesis is to explore one aspect of the bacteriophage - bacterial host relationship through the study of a bacteriophage protein called Ea22.

### 1.2 Bacteriophage-host life cycles and interactions

Bacteriophages have been studied for over one hundred years and have served as a critically important model in molecular biology discovery [6]. As shown in Figure 1.1, the bacteriophage life cycle is characterized by two distinct states, an infectious state where viral particles are produced and a prophage state where the viral genome is stably integrated into the host genome [7] . Bacteria harboring the viral genome are called lysogens.

In the laboratory, the lysogenic-lytic developmental switch is inducible by UV irradiation, antibiotics, and hydrogen peroxide [8] . The latter type of stress produces oxidative damage that most resembles the type of threat faced by the bacterial host when it is under attack the immune system within the human intestine. Once the lytic development program is initiated, a burst of gene expression occurs within minutes leading to replication and excision of the viral genome, production of structural proteins and release of new virions. Occasionally, host genes and gene fragments may come with the viral genome upon excision, leading to viruses with new traits and new evolutionary advantages. Given the time and scale in which replication and gene transfer occurs, the resulting genetic diversity is unlike anything else on the planet. The Shiga toxin gene carried by EHEC and STEC may confer an evolutionary advantage to the virus and host alike [9].

### 1.3 Bacteriophage λ as a model for investigating Stx phages

Bacteriophage λ is a useful model for studying related Shigatoxigenic (Stx) phage owing to its similar life cycle, similar genome, and overall ease of propagation. It is remarkable that despite over fifty years of molecular genetic studies, not all λ genes have been characterized. Many of these unknown reading frames map to a segment of

the genome between the *exo* and *xis* genes [10]. One explanation for the under-characterization of the *exo-xis* region is that its genes are known to be dispensable for viral development [11]. While this is true under standard laboratory conditions, *exo-xis* gene may provide functions that help the virus better adapt to its host. It is this idea that I will pursue throughout my thesis, bringing in *exo-xis* genes not only from λ, but also from other Stx+ phages. Among the Shigatoxigenic phages found in EHEC, $\varphi 24_B$ and 933W are two that have emerged as model phages. When the genomes of $\varphi 24_B$ and λ

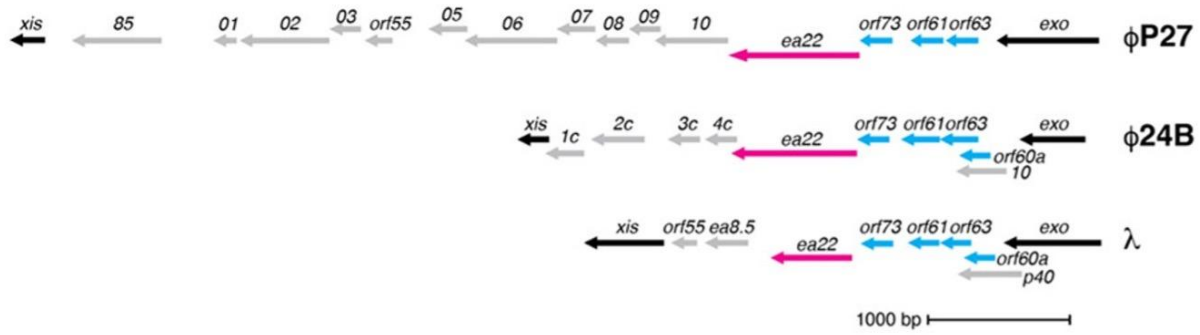



Figure 1.2 The *exo-xis* regions of selected bacteriophages
Conserved phage genes are shown in blue (Orf60a, Orf61, Orf63) and magenta (Ea22). Genes are read in the leftward direction arising from the pL promoter.

are compared, it is apparent that even less of $\varphi 24_B$ is understood [12]. Generally, the virulence of EHEC strains requires additional genetic elements such as LEE (Locus of Enterocyte Effacement) [13]. Bacteriophage 933W is notable due to its ability to turn non-pathogenic strains into pathogenic strains without any additional genetic elements [14].

**1.4 Open reading frames (ORFs) of the exo-xis region**

In Figure 1.2, the *exo-xis* regions of the λ, $\varphi 24_B$ and 933W genomes are presented. By virtue of its location in the bacteriophage genome, the *exo-xis* region genes are among the earliest genes transcribed during infection from the pL promoter.

Within the *exo-xis* regions, there is a cluster of gene near *exo* that appear to be conserved not only in kind but also in the same order. These genes are *orf60a*, *orf61*, *orf63*, *orf73* and *ea22*. After ea22, there is considerable variability. To date, the only structural information available is from the Donaldson laboratory who previously solved the structure of λ Ea8.5. The structure of Ea8.5 suggested that it might be a transcription factor due to its composition of a homeodomain fold fused with a zinc cluster motif. It is possible that the other *exo-xis* proteins may be transcription factors, as well, or tie into other major bacterial host processes like DNA replication and translation.

**1.5 The effect of Ea22 on the lysogenic-lytic decision**

The Wegrzyn laboratory (Univ Gdansk) has performed several functional assays to explore the roles of *exo-xis* genes and the proteins they encode. With the presence of the *exo-xis* genetic region, the efficiency of induction was exchanged in bacteriophage λ and ϕ24B phage. Different from other characterized *exo-xis* open reading frames (ORFs) such as *orf60*, *orf61* and *orf63*, instead of promoting the lytic development, *ea22*, the focus of the thesis, tends to maintain the lysogenic state of phage, interacting with currently unknown host proteins. Evidence for a pro-lysogenic role comes from quantitative assays of gene expression by qRT-PCR methods and a determination of how many lysogens are formed after infection.

In the laboratory setting, UV irradiation and antibiotics such as mitomycin C are commonly used as inducing agents owing to their efficacy in promoting prophage excision during induction. In human EHEC infection cases, immune cells neutrophils produce hydrogen peroxide. This inducing agent poses biological oxidative stress on phage; provokes prophage induction and *stx2* gene expression. This process requires the activation of bacterial S.O.S response and is RecA dependent [15] . The *exo-xis* genetic region was found to be particularly significant for the Shiga toxin converting phage under the oxidative stress stimulus posed by hydrogen peroxide [16]. It was also demonstrated that the *exo-xis* genetic region affects the expression of specific phage genes that are critical for all stages of phage development. These effects appear to be

more pronounced in ϕ24B phage associated with EHEC and less pronounced in bacteriophage λ suggesting that *ea22* is adapted to suit the environment and nature of its bacterial host. These adaptations may occur either at the level of gene expression or the in the sequence and structure of the protein itself.

### 1.5.2 The deletion of *exo-xis* genetic region impaired phage lysogenization

Regardless of the type of stressor used (mitomycin C antibiotic, hydrogen peroxide mediated oxidative stress or DNA damage produced by UV irradiation), *exo-xis* deletion mutants in both a bacteriophage λ and ϕ24B phage genetic background resulted in impaired lysogenization and delayed prophage induction with varying outcomes again suggesting that Ea22 is specifically suited to its host and the physiological conditions that it experiences[47]. In a l phage model system, hydroxyl peroxide treatment delayed prophage induction, UV irradiation promoted induction and no effect was observed using mitomycin C as an inducing agent. In contrast, induction of ϕ24B prophages were delayed to no matter what inducing methods were used[47] .

## 1.6 Thesis hypothesis and objectives

It is hypothesized that the *exo-xis* genetic region plays a significant role in the lysogenic-lytic decision and therefore the viral development and expression of stx genes. The prolysogenic function of ea22 suggests its potential therapeutical function in combating EHEC infections through controlling the expression of Shiga toxin.

The expression levels [17] , functional outcomes and possibly even the structures of Ea22 proteins are mosaic requiring a multifaceted approach to understanding its role in viral development. My thesis draws its inspiration from structural genomics following the belief that one path to understanding function is to first understand structure. Towards this objective, I focused much of my studies upon the manufacture and characterization of Ea22 proteins and protein fragments. While it was outside the scope of my research, a C-terminal fragment was eventually solved by NMR spectroscopy and shown to be very similar to what was predicted by Google AlphaFold,

a new machine learning technique that is able to accurately predict structure from sequence. The structures are described in Chapter 4 and their biological implications. Milestones achieved throughout my thesis research include:

1. the production of several Ea22 fragments from three bacteriophage representatives
2. the characterization of these fragments by numerous biophysical techniques
3. domain mapping of φ27 Ea22 by limited proteolysis
4. attempting to refine crystals of the λ Ea22 C-terminal domain

### 1.7 Thesis organization

The following thesis chapters are organized in terms of published work in Chapter 2, followed by unpublished work in Chapter 3 and conclusions in Chapter 4.

# Chapter Two – Published Manuscript

*"Ea22 proteins from lambda and Shiga toxin-producing bacteriophages balance structural diversity with functional similarity"*

Jinge Tong, Bożena Nejman-Faleńczyk, Sylwia Bloch, Alicja Węgrzyn, Grzegorz Węgrzyn, and Logan W Donaldson

## Author Contributions

Jinge Tong performed protein purifications, assisted in biophysical analyses and performed limited proteolysis on a series of proteins and protein fragments described in the study. BNF, SB, AW and GW performed biological assays on gene and gene fragments and provided associated text and figures. LWD designed the study, performed the NMR analysis, manufactured proteins and wrote the manuscript.

## Other Contributions

The Hospital for Sick Children SPARC Centre performed Multi-angle laser scattering assays and calorimetry assays for the study. The University of Toronto Mass Spectrometry Centre provided services to aid the identification of protein fragments.

## Citation

Tong J, Nejman-Faleńczyk B, Bloch S, et al (2020) Ea22 Proteins from Lambda and Shiga Toxin-Producing Bacteriophages Balance Structural Diversity with Functional Similarity. *ACS Omega*. *https://doi.org/10.1021/acsomega.0c00894*.

## Notice

## 2.1 Abstract

Enterohemorrhagic *Escherichia coli* (EHEC) outbreaks are commonly associated with contaminated food sources. Unlike normal intestinal bacteria, EHEC are lysogens of lambdoid bacteriophages that also carry a gene for Shiga toxin. Oxidative attack by the immune system or other stressors on the bacterial host can activate the lytic pathway of the latent phage genome to produce phage progeny and the release of Shiga toxin into the surrounding tissues. Within the genomes of bacteriophage λ and Shiga toxin expressing ($Stx^+$) phage such as $\varphi 24_B$ and φP27, there is a conserved set of open reading frames that is located between the *exo* and *xis* genes that influences the lysogenic-lytic decision. In this report, we have focused on the largest *exo-xis* region open reading frame termed *ea22* that has been shown previously to have pro-lysogenic properties. Using a variety of biophysical and bioinformatic methods, we demonstrate that λ and φP27 Ea22 proteins are tetrameric in solution and can be considered in terms of an amino-terminal region, a central coiled-coil region, and a carboxy-terminal region. The carboxy-terminal region of λ and $\varphi 24_B$ Ea22, expressed on their own, form dimers with exceptional thermostability. Limited proteolysis of φP27 Ea22 also identified a C-terminal region along the predicted boundaries. While the three Ea22 proteins all appear to have the hallmarks of a domain in their respective C-terminal regions, each sequence is remarkably dissimilar. To reconcile this difference among Ea22 proteins from λ and $Stx^+$ phages alike, we speculate that each Ea22 may achieve the same function by targeting different components of the same regulatory process in the host.

## 2.2 Introduction

*Escherichia coli* (EHEC) outbreaks due to contaminated meat, vegetable, and water supplies have been recorded since the early 1980 with a spectrum of health issues ranging from mild gastrointestinal problems to kidney failure [18, 19] . Unlike normal intestinal bacteria, EHEC carry a stably integrated bacteriophage genome that encodes a ribosome-inactivating Shiga toxin (typically *Stx2*) [20, 21] . When EHEC are exposed to a variety of environmental stressors such as oxidative attack from the

immune system, the endogenous phage genome transitions from a dormant lysogenic state to a lytic state along with production and release of Shiga toxin [22, 23]. Antibiotic treatment, typically the first choice in a bacterial infection, can be counterproductive in EHEC infections because this treatment represents another type of stress that my drive the endogenous phage towards a lytic response. Since Shiga toxin producing ($Stx^+$) phages belong to the same family as bacteriophage λ [24], a molecular level comparison may reveal new ways in which $Stx^+$ phages contribute the pathogenicity of EHEC and potentially new therapeutic leads to combat infection [25, 26].

A relatively unknown region between the *exo* and *xis* genes of λ and $Stx^+$ phages is known to affect the host cell-cycle [27, 28] and the lysogenic-lytic transition [29, 30] yet, paradoxically, is also dispensable for normal viral development. This so-called *exo-xis* region consists of four open reading frames termed *ea22* [31], *orf73*, *orf61* [32], and *orf63* [33]. While the overall number of *exo-xis* region genes in $Stx^+$ phages can be larger, these four genes are conserved suggesting there may be some evolutionary pressure to maintain them. The hybrid zinc finger / homeodomain fold of Ea8.5 suggests that it may serve as a transcription factor or serve another role associated with nucleic acid binding [34]. Less is known about Orf63, except that is has an oligomeric two-helix fold[33].

Ea22 is the largest protein of the *exo-xis* region and is expressed early in the lytic cycle by the *pL* promoter. Unlike other *exo-xis* genes studied to date that accelerate development of the lytic state [32, 33, 35], *ea22* promotes the maintenance of the lysogenic state [31]. Before we embarked on this study, it was not known if *ea22* even encoded a protein although the functional data appeared to suggest it was the case. With set of expressed and purified Ea22 proteins and protein fragments from λ phage and two representative $Stx^+$ phages, φP27 [36] and $\varphi 24_B$ [29], we demonstrate that Ea22 is multidomain protein that can be considered in terms of an amino-terminal region, central coiled-coil region, and carboxy-terminal region. In this report, we have

used circular dichroism spectroscopy, nuclear magnetic resonance spectroscopy, differential scanning calorimetry, and limited proteolysis to show the C-terminal region of each Ea22 protein has the qualities of a protein domain, but each C-terminal domain is not necessarily similar. We consider this observation in term of the functional differences between λ and $Stx^+$ phage.

## 2.3 Materials and Methods

### 2.3.1 Cloning and expression

Several full-length genes and gene fragments of λ_Ea22, φP27_Ea22, and $φ24_B$_Ea22 were synthesized and placed in a N-terminal 6xHis-tagged, T5 promoter-driven expression vector by ATUM (Newark CA): Clone #134893 λ_Ea22(1-182), #179655 λ _Ea22(93-182), #179656 λ_Ea22(102-182), #179657 λ_Ea22(109-182), #179658 λ_Ea22(119-182), #350149 φP27_Ea22(1-62), #350148 φP27_Ea22(145-315), #358023 φP27_Ea22 (176-315), #350147 $φ24_B$_Ea22(159-301), #351679 $φ24_B$_Ea22(164-269), #393816 φP27_Ea22(235-315). Full-length N-terminal 6xHis-tagged φP27_Ea22(1-315) and $φ24_B$_Ea22(1-301) were made by inserting a suitable gene fragment into the *NdeI* and *BamHI* sites of pET28b (Novagen). For milligram quantities of purified protein, a 1.5-3.0 L culture of *E. coli* BL21(DE3) containing 50 μg/mL kanamycin was initially grown at 37 ˚C with shaking to an $A_{600}$ of 0.5 then chilled to either 16 ˚C (λ _Ea22) or 25 ˚C (φP27_Ea22, $φ24_B$_Ea22) prior to induction with 1 mM IPTG for 18 h (λ _Ea22) or 6 h (φP27_Ea22, $φ24_B$_Ea22). The cell pellet was harvested by centrifugation (4,000 *g*, Beckman JA-10 rotor), resuspended in T300 buffer (10 mM Tris, 300 mM NaCl, 0.05% $NaN_3$, pH 7.7) and lysed by French press and sonication. The soluble fraction was separated from the cell debris by centrifugation (28,000 *g*, Beckman JA-20 rotor) and applied to a 5 mL Nuvia nickel affinity column (BioRad). After the affinity column was washed with T300 + 10 mM imidazole and the protein of interest was eluted with T300 + 500 mM imidazole. The eluate was concentrated to 5 mL, treated with 20 mM DTT, and applied to a gel filtration column (HiPrep 16/60 S100, GE Life Sciences) that had been pre-equilibrated to 10 mM Tris, 100 mM NaCl, 0.05% $NaN_3$, pH 7.7. Suitable fractions from the column were assessed

for purity by SDS-PAGE, pooled and concentrated accordingly. Protein concentrations were estimated by $A_{280}$. For studies requiring protein concentrations of φP27_Ea22 > 2 mg/ml, the buffer NaCl concentration was increased to 0.5 M to improve protein solubility.

### 2.3.2 Nuclear magnetic resonance (NMR) spectroscopy

Proteins and protein fragments were uniformly $^{15}N$-labelled by following the protocol previously described except that M9 minimal media was used (6 g $Na_2HPO_4$, 3 g $KH_2PO_4$, 1 g $^{15}NH_4Cl$, 0.5 g NaCl, 10 g glucose in 1 L of water supplemented with 1 mM $CaCl_2$, 1 mM $MgSO_4$, 50 µg/ml kanamycin and a trace mineral mix). Samples were concentrated to 0.2 mM and supplemented with 10% $D_2O$. $^{1}H$,$^{15}N$-HSQC spectra were acquired at various temperatures on a 700 MHz Bruker Avance-III spectrometer with a nitrogen-cooled triple-resonance probe. Datasets were processed with NMRPipe [37] and interpreted with CCPN Analysis v2.4[38].

### 2.3.3 Circular dichroism (CD) spectroscopy

Protein samples were dialyzed to a 10 mM Tris, 50 mM NaCl, 0.05% $NaN_3$, 5 mM DTT, pH 7.7 and standardized to a 20 µM concentration for a 0.1 cm path length quartz cell. Spectra were acquired from 260-200 nm using a Jasco J810 instrument equipped with a six-position temperature-regulated cell holder. Thermostability was assessed by holding the wavelength constant at 222 nm and ramping temperature from 20-90 ˚C at a rate of 2 ˚C/min. Signals were converted from mdeg to mean residue ellipticity and plotted with Pro Fit (Quansoft).

### 2.3.4 Differential scanning calorimetry (DSC)

Protein samples were treated with 20 mM DTT, standardized to 1 mg/mL and then dialyzed to 10 mM Tris, 0.1 M NaCl, 0.05% $NaN_3$, pH 7.7. In cases where the solubility was limited, the NaCl concentration was adjusted to 0.5 M. Before analysis, samples centrifuged at 24,000 *g* for 5 min at 4˚C to remove any precipitates. Protein

and matched buffer were loaded into the sample and reference cells, respectively, and the temperature was increased from 20-100˚C at 1 ˚C/min at 3 atm pressure. Background subtracted data were fit to the simplest model (single or multistate) that conformed to a 99% confidence level.

### 2.3.5 Size-exclusion chromatography with multiangle laser scattering (SEC-MALS)

Protein samples were prepared in the same manner for DSC experiments. The chromatography unit of the instrument was configured with an Infinity-II HPLC (Agilent) and an AdvanceBio SEC300A column (Agilent). The chromatography system was equilibrated in the sample buffer for 18 h before the first injection for optimal stability. The detection section of the instrument was configured with a MiniDAWN TREOS MALS and OptiLab T-rEX refractive index components (Wyatt). Prior to the analysis, a 2 mg/ml BSA standard was injected to calibrate peak and retention time characteristics of the flow cells. Following calibration, a 20 µL protein sample was injected. Chromatograms were analyzed and molecular masses determined with ASTRA software (Wyatt).

### 2.3.6 Bioinformatics

Sequences were submitted to NetSurfP-2.0 [39] for secondary structure analysis. Coiled-coil prediction was performed with the programs, DeepCoil, MARCOIL, and PCOILS, all part of the HHpred suite [40].

### 2.3.7 Limited Proteolysis

A 20 µM solution of φP27_Ea22(1-315) in buffer containing 20 mM Tris, 150 mM NaCl, 10 mM DTT, pH 7.7 was digested with 200 nM of proteomics-grade trypsin (Sigma-Aldrich) for 30 min at 37˚C. The digest was stopped with either 2x SDS sample loading solution for electrophoretic assays or 1% formic acid (*v/v*) for LC-MS assays.

### 2.3.8 Mass spectrometry

Electrospray ionization (ESI) mass spectra were acquired on a 6538 UHD model quadrupole time-of-flight mass spectrometer coupled to a 1260 Infinity model high performance liquid chromatography system (Agilent Technologies). Approximately 1 pmol of protein mixture was separated with a 50 × 2 mm Jupiter 5 µm C4 300 Å reverse phase column using a flow rate of 0.25 mL/min. Following a 1 min wash with 0.1% (*v/v*) formic acid and 5% acetonitrile, a 15 min separating gradient to 60% acetonitrile was applied. MS data were processed and analyzed with MassHunter (Agilent).

#### 2.3.9. Bacterial survival assay

A bacterial culture was grown in LB at 30°C to $A_{600}$ = 0.2. From a 1 ml aliquot, bacteria were harvested by centrifugation, washed with 0.85% NaCl, and centrifuged again. In preparation for infection, the pellet was resuspended in 1 ml of TCM buffer (10 mM Tris-HCl, 10 mM $MgSO_4$, and 10 mM $CaCl_2$, pH 7.2). Following a 30 min incubation at 30°C, phage was added to m.o.i. of 1 and maintained for 15 min. Serial dilutions were made in 0.85% NaCl and plated. After an overnight incubation at 37°C, the percentage of surviving *E. coli* was calculated relative to a culture in which TCM buffer was added instead of phage. Bacteria, bacteriophages and plasmids for this assay are summarized in Table 2.1.

## 2.4 Results

### 2.4.1 Sequence analysis of Ea22 proteins

Given the maturity of the bacteriophage field, it is remarkable that there have been few explorations of the *exo-xis* region. One reason for the lack of interest may be due to the dispensability of *exo-xis* region for viral development. As shown in Figure 2.1, the number of open reading frames in the *exo-xis* region is variable among lambdoid phages; however, a core set of four genes is retained (*orf63*, *orf61*, *orf73*, and *ea22*) suggesting some functional significance. Before we embarked upon expression studies of Ea22 from λ and two representative Stx phages, φP27 and $\varphi 24_B$, the secondary structure of each protein was predicted and used to supplement a sequence alignment. Overall, the alignment suggests that Ea22 is partitioned into three regions (Fig. 1b). In

contrast to φ24$_B$ and φP27 Ea22, the N-terminal region of λ Ea22 appears to be truncated to a degree that a potential domain is lost. The central region of Ea22 is predicted to contain several long α-helical segments which suggests a coiled-coil structure. The C-terminal regions of the three Ea22 proteins are the most intriguing since they bear little, if any, resemblance to each other even at the secondary structure level.

**2.4.2 Structural and functional features of λ Ea22**

When lower induction temperatures were used (20 ˚C), we obtained milligram quantities of soluble 6xHis-tagged λ Ea22 (182 aa.) and several C-terminal fragments thereby providing a path towards a biochemical characterization of this unknown phage protein. The largest C-terminal fragment selected for the study, λ_Ea22(93-182), began just inside the coiled-coil region and was followed by four additional fragments that were designed to successively remove one predicted secondary structure at a time (Figure 2.2a). The two shortest C-terminal fragments, λ_Ea22(129-182) and λ_Ea22(119-182), expressed as inclusion bodies and could not be refolded from denaturant, leaving λ_Ea22(109-182) as the minimal soluble fragment with a predicted secondary structure of ß1ß2α1α2α3. Table 2.2 summarizes all expression conditions and solubilities of protein fragments presented in this study.

The far UV circular dichroism (CD) spectrum of full-length λ Ea22(1-182) demonstrated broad troughs at 208/222 nm suggesting that the protein was folded with hallmarks of α-helical secondary structure. Differential scanning calorimetry showed that full length λ Ea22(1-182) denatured with a single transition at 48 ˚C. A similar thermal denaturation midpoint was observed using circular dichroism spectroscopy. The reader is referred to Table 2 for a summary of the biophysical and biochemical parameters described in this report. The oligomeric state of λ Ea22(1-182) was determined by SEC-MALS, a method that combines analytical size exclusion chromatography (SEC) with multi-angle laser scattering (MALS). While full-length λ Ea22(1-182) was observed to be

tetrameric by SEC-MALS, the C-terminal fragment, λ_Ea22(119-182) was dimeric (Fig. 2b). In the discussion, we present a model of the Ea22 protein that reconciles these two observed oligomeric states. The calorimetry, circular dichroism, and SEC-MALS data for full length λ Ea22(1-182) is presented in Figure 2.3 and Figure 2.4.

A sample of full-length $^{15}$N-labeled λ Ea22 was assessed by nuclear magnetic resonance (NMR) methods. At room temperature (298 K), the majority of $^{1}$H-$^{15}$N resonances were extremely broadened. This observation is consistent with a 22.5 kDa Ea22 monomer that is exhibiting the solution characteristics of a 90 kDa tetramer. Line shapes improved when the temperature was raised to 310 K as a result of the protein being able to tumble faster in solution. In the $^{1}$H-$^{15}$N HSQC spectrum shown in Figure 2.2c, each resonance, or peak, is roughly attributed to one backbone amide $^{1}$H-$^{15}$N pair found in each amino acid except proline. Following this rule, approximately two hundred peaks were expected, yet only one-third of that predicted number was observed. This discrepancy can be attributed to the large apparent molecular weight of the protein and additional ms-μs time scale dynamics from unstructured regions and large domain movements. Considering these factors, we hypothesized that the observable peaks were localized to one domain that was separated enough from the rest of Ea22 for it take on solution characteristics of a small protein. From the secondary structure predictions, the C-terminus was the most plausible region for such a domain.

Of the five C-terminal protein fragments that were cloned, we prepared $^{15}$N-labeled NMR samples of the three that were soluble, being λ_Ea22(93-182), λ_Ea22(102-182), and λ_Ea22(109-182). The $^{1}$H-$^{15}$N HSQC spectrum of the λ_Ea22(109-182) protein fragment was comparable to the spectrum of the full-length λ Ea22 protein (Figure 2.2d), thereby confirming our hypothesis that the C-terminal region is tethered far enough from the coiled coil region to appear as if it is unrestrained.

A CD thermal melt assay of the three soluble C-terminal λ Ea22 protein fragments was performed by monitoring the change in ellipticity at 222 nm. As shown in Figure 2.2e, no change was observed up to 90 ˚C for λ_Ea22(93-182) and λ_Ea22(102-182) suggesting that these dimeric protein fragments were highly thermostable. In contrast, a $T_m$ of 82 ˚C was observed for λ_Ea22(109-182), suggesting amino acids between 102-108 make a minor contribution to the structure of the C-terminal domain.

Differential scanning calorimetry (DSC) at 3 atm pressure was performed to extend the CD thermal assay. From this investigation, we observed that the thermostability of λ_Ea22(93-182) and λ_Ea22(102-182) was greater than 100 ˚C. Thus, this first biochemical and biophysical survey of λ *ea22* suggests that it is more than a nonessential open reading frame and in fact, encodes a multidomain protein that is experiencing evolutionary pressure to maintain a stable fold.

Upon infection with λ phage, a greater percentage of bacteria survive when they express full-length λ Ea22 from a plasmid (Figure 2.2f). The observed percentage in this case was in excess of 100% due to the ability of lysogens to resist superinfection and continue dividing [29] . When this experiment was performed with the minimal thermostable C-terminal fragment, λ_Ea22(102-182), the proportion of bacteria surviving infection was similar to the empty plasmid control suggesting there was no effect.

#### 2.4.3 Structural and functional features of $\varphi 24_B$ Ea22

$\varphi 24_B$ is a model Stx phage for functional studies due to its rapid development in bacteria and ability to maintain titer during storage [32, 33] . Compared to the λ phage genome, $\varphi 24_B$ not only contains more *exo-xis* open reading frames, but also an Ea22 protein that differs in length and composition, particularly in its the C-terminal domain. These combined differences may contribute to the fitness of $\varphi 24_B$ and its bacterial host within the human intestine.

We began our biochemical survey by expressing full-length $\varphi 24_B$ Ea22; however, the protein was insoluble and could not be refolded either by slow or fast removal of denaturant. While a direct comparison between full-length $\varphi 24_B$ Ea22 and λ Ea22 could not be performed, we were successful at expressing a C-terminal fragment, $\varphi 24_B$_Ea22(135-261), extending from the coiled-coil region to the native C-terminus (Figure 2.5a). A $^{1}H$-$^{15}N$ HSQC NMR spectrum of isotopically labeled $\varphi 24_B$_Ea22(135-261) suggested the presence of a small, folded domain within the fragment (Figure 2.5b). While there was no obvious similarity in sequence or secondary structure with λ Ea22, the $\varphi 24_B$ Ea22 C-terminal fragment was also dimeric by SEC-MALS (Figure 2.5c) and demonstrated exceptional thermostability as no structural transitions were observed

by either CD or DSC methods performed at temperatures in excess of 100˚C. The $\varphi 24_B$ Ea22 C-terminal domain has the predicted secondary structure ß1ß2α1α2ß2ß3ß4ß5.

The same assay that tested the ability of λ Ea22 to affect the survival rate of infected *E. coli* was performed for full-length $\varphi 24_B$_Ea22(1-261) and the soluble C-terminal fragment $\varphi 24_B$_Ea22(135-261). Like the λ assay, survival was improved when $\varphi 24_B$_Ea22(1-261) was overexpressed (Figure 2.5d). However, unlike the λ assay, survival was also improved for the C-terminal fragment. It is possible that structural differences between the λ and $\varphi 24_B$ C-terminal fragments may have contributed to a different functional outcome during infection.

### 2.4.4 Oligomeric state and domain organization of φP27 Ea22

The *exo-xis* region of the $Stx^+$ phage φP27 has eleven additional genes that follow *ea22*, in contrast to four in $\varphi 24_B$ and two in λ. The two genes in λ include the putative transcription factor *ea85* and *orf55*, a gene with no known function.

The C-terminal domain of the φP27 Ea22 protein is 54 aa. larger than $\varphi 24_B$ and 133 aa. larger than λ. When full-length φP27_Ea22(1-315) was overexpressed, it partitioned between the soluble and insoluble fractions of the bacterial lysate. Lower induction temperatures improved the yield of soluble protein. Owing to the extensive coiled-coil domain that P27 Ea22 shares with λ and $\varphi 24_B$, the far-UV CD spectrum was characteristically α-helical. A CD-based thermal denaturation assay revealed a broad transition centered at 55 ˚C. When the thermostability of φP27_Ea22(1-315) was reassessed by differential scanning calorimetry, two transitions were observed around 50 ˚C suggesting that φP27 Ea22 is organized into two domains. Despite the sequence and predicted secondary structural differences between P27 and λ Ea22, a SEC-MALS analysis revealed that φP27_Ea22(1-315) is also a tetramer.

To delimit the domain boundaries of φP27 Ea22, a sample was digested with trypsin for varying periods. Over several trials and digestion periods, we observed that the protein was initially cleaved into a 27 kDa fragment and a C-terminal 9 kDa fragment. A 6 kDa fragment was also observed in some trials, albeit to a lesser extent (Figure

2.6a). Using LC-MS methods, the 9 kDa and 6 kDa fragments were mapped to 237-315 and 255-315, respectively in the φP27 Ea22 protein (Figure 2.6b). Given that the 9 kDa fragment φP27_Ea22(237-315) was the most stable, the predicted secondary structure of the φP27 Ea22 C-terminal domain is α1ß2α2ß2α3.

The N-terminal region of φP27 Ea22 is predicted to have the secondary structure α1ß2ß2 ß3α2. A fragment encompassing this region, φP27_Ea22(1-62), was expressed along with two C-terminal fragments, φP27_Ea22(145-315) and φP27_Ea22(176-315). These C-terminal protein fragments were insoluble and could not be refolded from denaturant. A third, minimal C-terminal protein φP27_Ea22(235-315) identified by mass spectrometry did not express.

### 2.4.5 A model of Ea22

A model of Ea22 is presented in Figure 2.7 that reconciles our biochemical investigations of Ea22 from phage λ, and the $Stx^+$ phages, $φ24_B$ and φP27. If the tetrameric coiled-coil region is placed centrally in an anti-parallel orientation, the C-terminal sequences are then juxtaposed at each end and are free to dimerize. A long enough tether may permit the C-terminal domain to tumble freely thereby explaining why $^{1}H$-$^{15}N$ HSQC NMR spectra of the C-terminal domain and full-length λ Ea22 were similar. Since the N-terminal region present in $φ24_B$ and φP27 could not be expressed, the propensity for it to dimerize like the C-terminal domain remains unknown.

Overall, the proposed architecture implies an adaptor-like function for Ea22. While Ea22 was not among the host-phage or phage-phage protein partnerships that were identified in yeast two-hybrid (Y2H) assays [41, 42] , Ea22 may still serve in a multiprotein complex that would be undetectable by a Y2H assay or alternatively bind nucleic acids. The proposed architecture permits the N-terminal and C-terminal regions to evolve independently ultimately leading to different partnerships possibly within the same host regulatory pathway.

## 2.5 Discussion

The *exo-xis* region of λ phage is not required for viral development, yet at least four genes are conserved among λ and Shiga toxin producing ($Stx^+$) phage found in clinically relevant *E. coli*. Only recently has the *exo-xis* region been studied at a molecular level of detail. In the case of *ea22*, the largest of the conserved *exo-xis* region genes, it was not even known if it encoded a protein until this study. The *exo-xis* region also presents a rare new opportunity to examine the vast genomic landscape of phages[43, 44] and to examine the protein sequence landscape that determines a given three-dimensional structure. For example, the *exo-xis* protein Ea8.5 [34], the capsid-tail interface structural protein gpU [45], and the major tail protein gpV [46] bear little sequence similarity to any other proteins except from closely related phage, yet the high-resolution structures reveal variations of a common fold. The *exo-xis* region also presents an opportunity to explore new phage-host relationships and the protein-protein and protein-nucleic acid interactions that underlie them.

A previous study had shown that *ea22* deletion mutants created a bias toward a sustained lytic developmental outcome by reducing the efficiency of lysogenization and increasing the number of phage progeny [31]. We extend these results with an assay in which bacteria expressing Ea22 were infected with either λ or p24B phage and then number of bacteria surviving were scored relative to an empty expression vector. In the case of full length Ea22, the relative survival was over 100% because bacteria persisted as lysogens long enough to make it through additional rounds of cell division. While the C-terminal domain of λ Ea22 had no effect survival, the C-terminal domain from the $Stx^+$ phage $\varphi 24_B$ did provide similar pro-survival benefits reinforcing our biochemical and biophysical data suggesting that the C-terminal domains are not only structurally different, but also functionally different. Following this idea, we speculate that the λ or $Stx^+$ C-terminal domain from Ea22 may interact with a set of host and phage proteins supporting lysogenization in different ways.

Early phage genes are expressed approximately 15-30 min post-infection and include the well characterized *N*, *cI* and *cro* gene products along with many *exo-xis* region members. In Shiga toxin producing phage such as $\varphi 24_B$ and φP27, *ea22* expression reached levels up to an order of magnitude higher than other early genes

[47]. In λ phage, high levels of *ea22* expression were also observed but not to the extent of its Shiga toxin producing counterparts. Thus, the amount of Ea22 produced during the earliest stage of infection could also have a direct outcome on host survival and suppression of lytic development.

The specific structural features that have have described for λ and $Stx^+$ phage Ea22 proteins may also be considered in terms of the EHEC host. Cattle, for example, do not have a suitable receptor for Shiga toxin therefore serve as a reservoir for EHEC. Within the intestinal tract of a cow, Ea22 may attenuate the response to stress and favor lysogenization in order for the bacterial population to establish itself. Likewise, in the human intestinal tract where bacterial survival is more challenging, Ea22 may slow down the lytic development cycle so bacteria are not simultaneously attacked from outside and within. Bacterial lysogens overexpressing λ *ea22* from a plasmid exhibit a delay in phage development when induced with hydrogen peroxide suggesting that oxidative stress pathways are involved [47]. When this experiment was repeated in $Stx^+$ phage, the effect was more pronounced suggesting that Ea22 may be different in $Stx^+$ phage or Ea22-mediated regulation of the response to oxidative stress is more complex.

Recently, a 20 bp small RNA (sRNA) was identified in Stx phage $\varphi 24_B$ with many of the same functional features as Ea22 [48]. Phage mutants lacking the region for the 80 bp precursor RNA did not lysogenize effectively, responded faster to SOS oxidative stress responses, and were more efficient at producing new progeny. One of the two possible targets of RNA was an anti-repressor, *d_ant*, that is found in other Stx phages and but notably absent in the λ genome. Since the molecular partners of Ea22 are unknown, it also remains to be determined if Ea22 interacts with this anti-repressor, or genes or gene products affected by the anti-repressor.

### 2.6 Conclusions

We have demonstrated that the *exo-xis* region protein Ea22 is a tetrameric functional, protein that affects host survival. Future proteomics and high-resolution structural studies are necessary to precisely establish the role of Ea22 and its suitability as a lead for combatting clinically important *E. coli* infections.

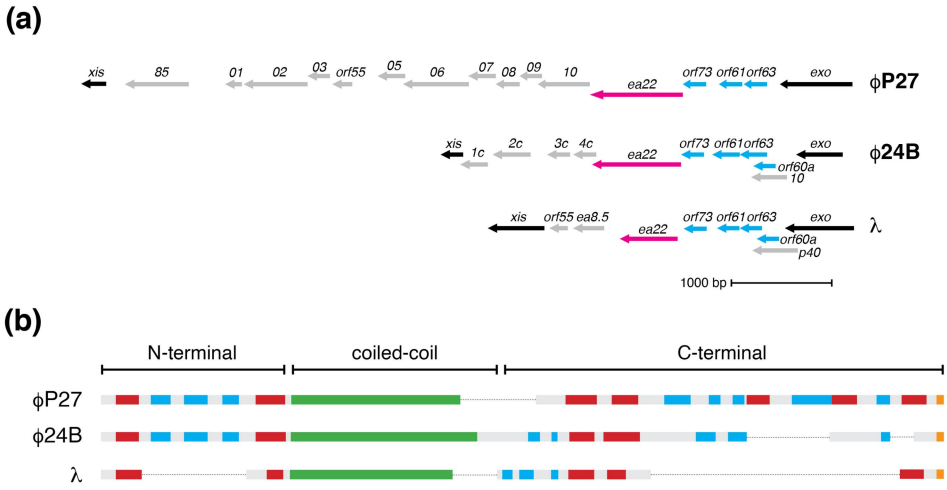



Figure 2.1. Gene map of the exo-xis regions from three phages.
(a) Ea22 is highlighted in magenta and other genes that are typically conserved among lambdoid phages are blue. For clarity, some ORFs of φP27, $\varphi 24_B$ and λ are shown without their prefixes of *vBEcoSP27_* and *vB_24B_*, and *lambda*. (b) Secondary structure prediction of Ea22 from λ phage and the Stx phages φP27 and $\varphi 24_B$. Predicted α-helices (red), ß-strands (blue), coiled-coil regions (green) and extended regions (grey) are indicated. The sequences are aligned according to sequence and secondary structure similarity. Ea22 sequences typically terminate with a +G or +GE motif (orange) of unknown significance. The secondary structure similarities suggest a tertiary structural level of organization consisting of N-terminal domain, central coiled-coil domain, and C-terminal domain.

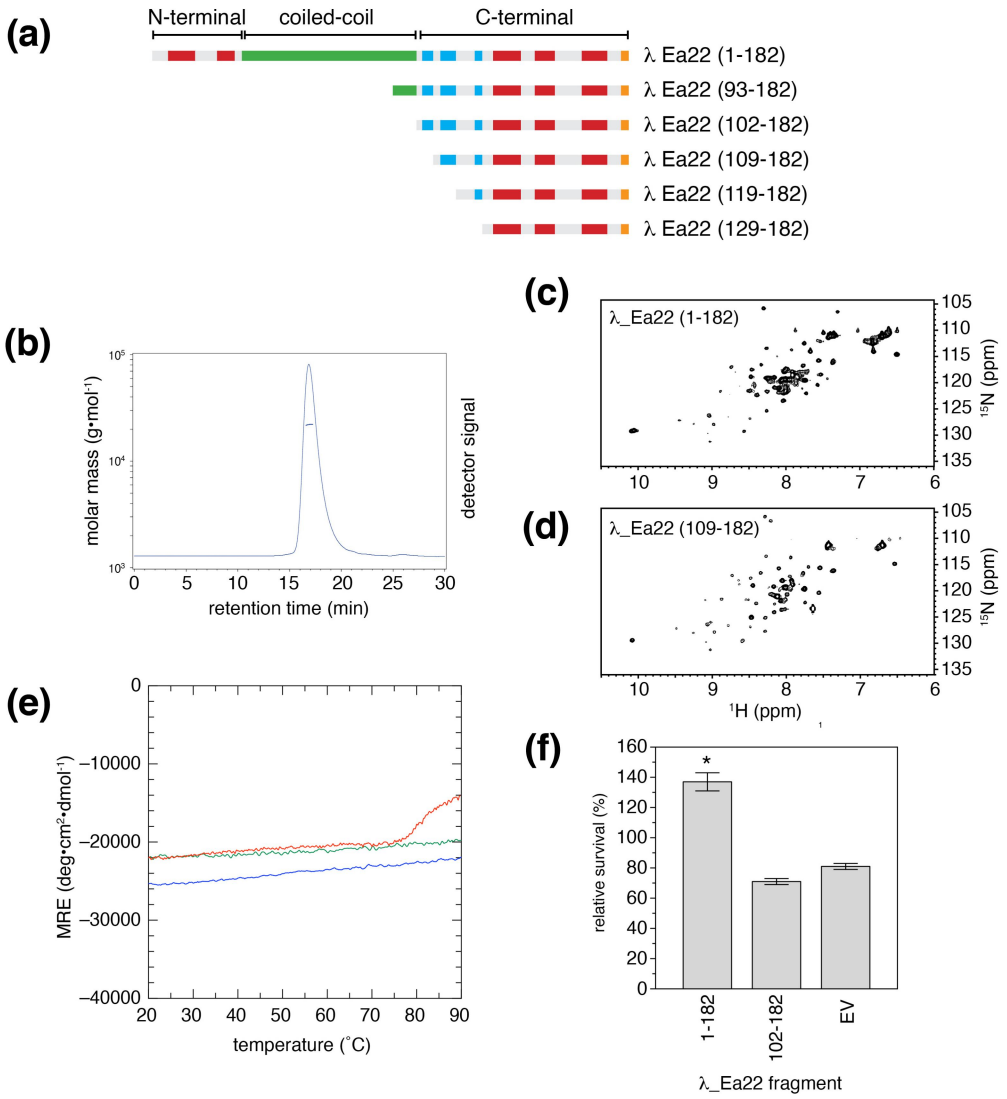



Figure 2.2. Structural characterization of λ Ea22 fragments.
(a) Various protein fragments extending from the latter part of the predicted coiled-coil region (green) to the native C-terminus were expressed. λ_Ea22(119-182) and λ_Ea22(129-182) were insoluble and not pursued further. (b) SEC-MALS analysis of λ_Ea22(93-182). One peak with a retention time of 16.9 min coincides with a molecular mass of 22.0 kDa suggesting that the native form of this protein fragment is a dimer (the calculated monomeric molecular mass from sequence is 10.7 Da). (c) $^{1}H$-$^{15}N$ HSQC NMR spectrum of $^{15}N$-labeled full-length λ Ea22. (d) $^{1}H$-$^{15}N$ HSQC NMR spectrum of the smallest soluble Ea22 C-terminal fragment, λ_Ea22(109-182). (e) CD thermograms of λ_Ea22(93-182, blue), λ_Ea22(102-182, green), and λ_Ea22(109-182, red) at a wavelength that is characteristic for α-helices. (f) Survival of E. coli MG1655 containing various λ Ea22 expression plasmids after infection with λ. Error bars represent the SD of five replicates. Relative survival is expressed against a control culture in which buffer was added in the place of phage particles. EV = empty vector control.

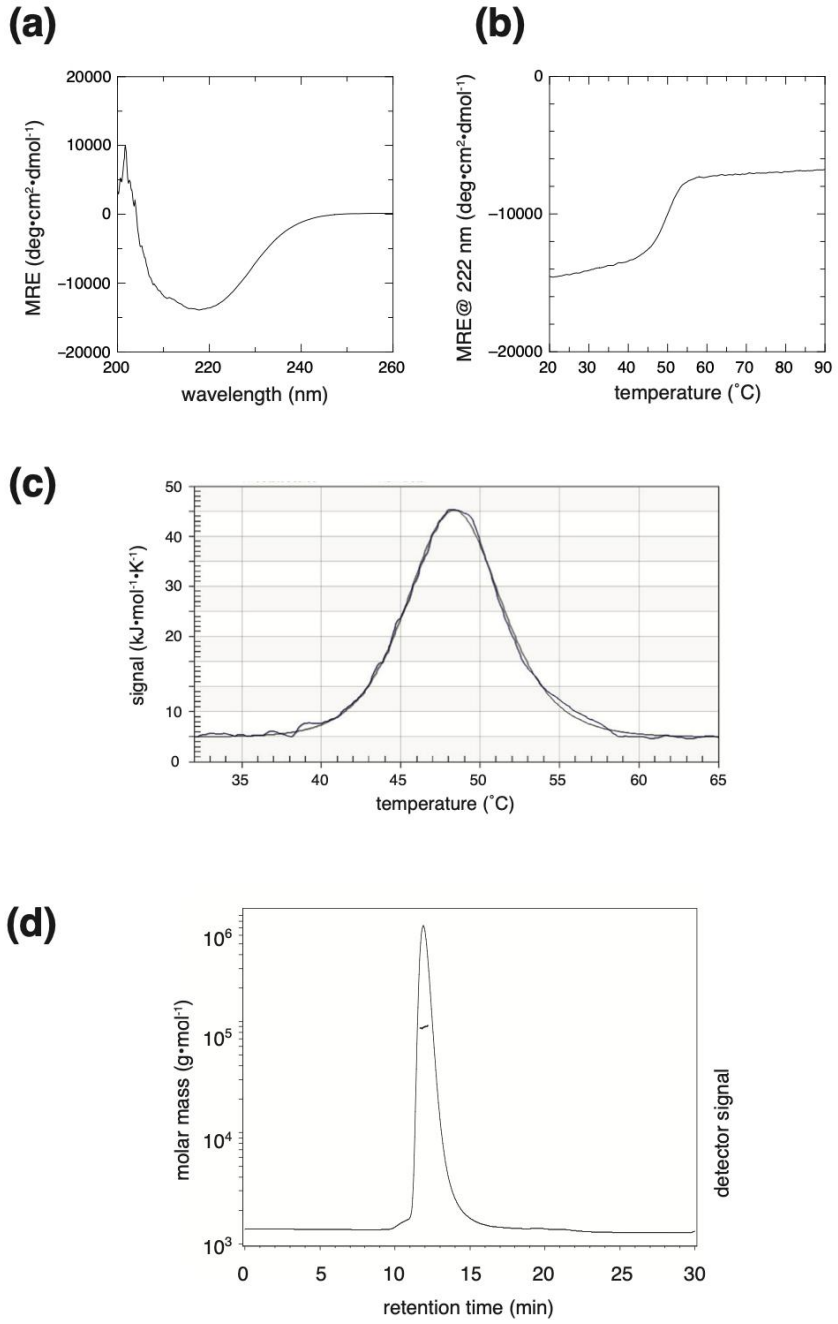



Figure 2.3. Biochemical features of full length λ Ea22.
(a,b) Far-UV circular dichroism spectrum and a thermostability assay monitored at 222 nm, a wavelength that is diagnostic for α-helical secondary structure (Tm = 44 ˚C). (c) Deconvoluted DSC thermogram showning one transition (Tm = 44 ˚C). (d) SEC-MALS assay. The observed peak coincides with a molecular mass of 88.8 kDa suggesting that λ Ea22 is a functional tetramer.

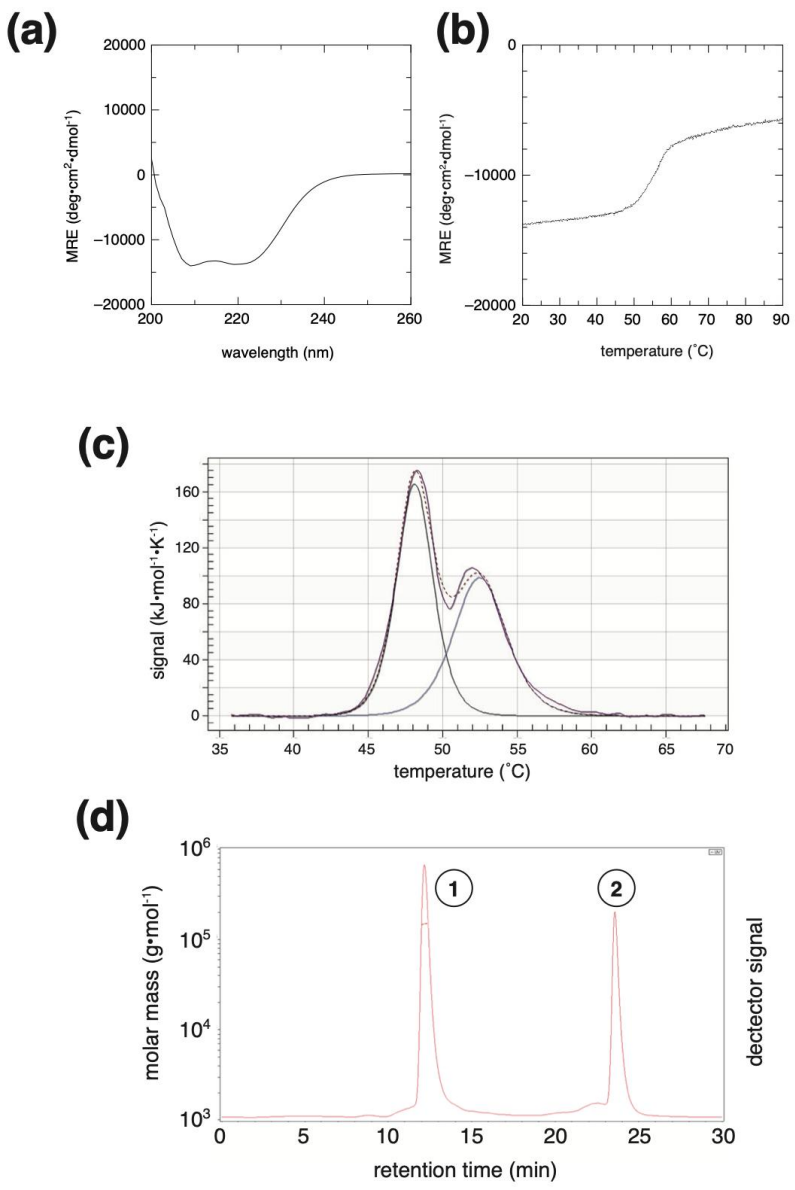



Figure 2.4. Biochemical features of full-length φP27 Ea22.
(a) Far UV CD spectrum (b) CD thermogram. (c) Deconvoluted DSC thermogram showing two transitions (d) SEC-MALS analysis. Two peaks were observed: Peak #1 with a retention time of 12.4 min coincides with a molecular mass of 150 kDa suggesting that the native form of this protein fragment is a tetramer (calculated monomeric molecular mass from sequence, 38.3 kDa). Peak #2 did not produce a light scattering signal and is likely a buffer component.

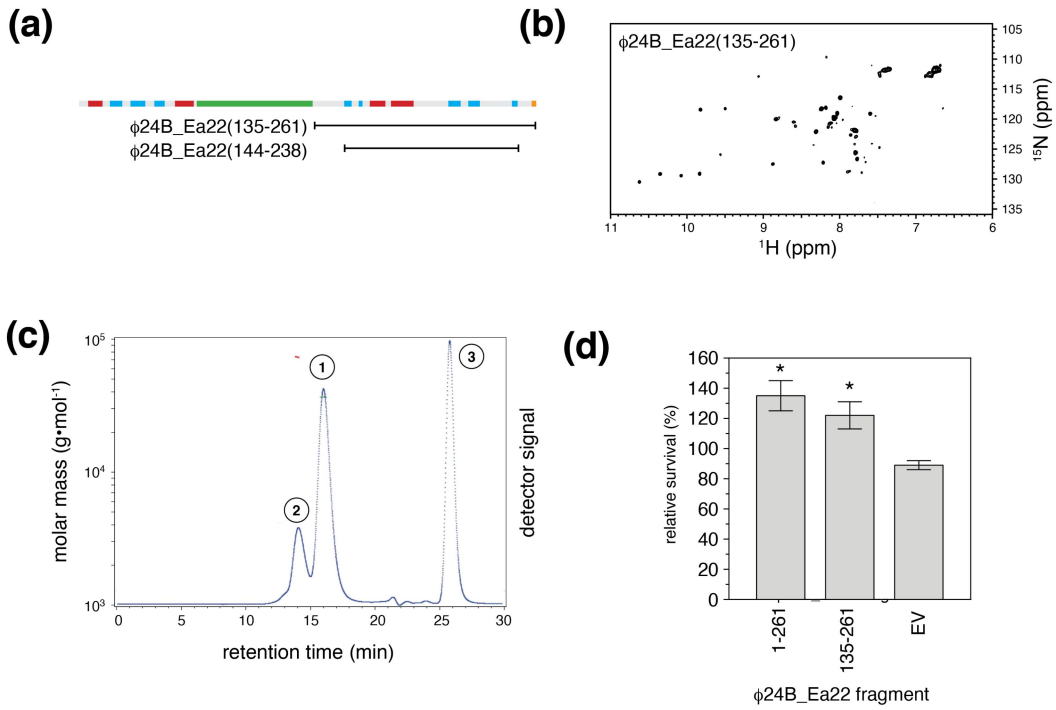



Figure 2.5. Structural and functional characterization of φ24B Ea22 fragments.
(a) A protein fragment φ24$_B$_Ea22(135-261) from the predicted coiled-coil region (green) to the native C-terminus expressed was soluble. A shorter fragment φs24$_B$_Ea22(144-238) constrained by the predicted secondary structures was insoluble. (b) $^{1}$H-$^{15}$N HSQC NMR spectrum of $^{15}$N-labeled φ24$_B$_Ea22(135-261). (c) SEC-MALS assay φ24$_B$_Ea22(135-261). Three peaks were observed. Peak #1 with a retention time of 16.2 min is the most prominent and coincides with a molecular mass of 37.2 kDa suggesting that the native form of this protein fragment is a dimer (the calculated monomeric molecular mass from sequence is 17.9 Da). Peak #2 with a retention time of 14.1 min coincides with a molecular mass of 73.1 kDa suggesting a possibly oxidized tetrameric species occurring as a dimer of dimers. Peak #3 did not produce a light scattering signal and is likely a buffer component. (d) Survival of *E. coli* MG1655 containing various φ24$_B$ Ea22 expression plasmids after infection with φ24$_B$. Error bars represent the SD of five replicates. Relative survival is expressed against a control culture in which buffer was added in the place of phage particles. EV = empty vector control.

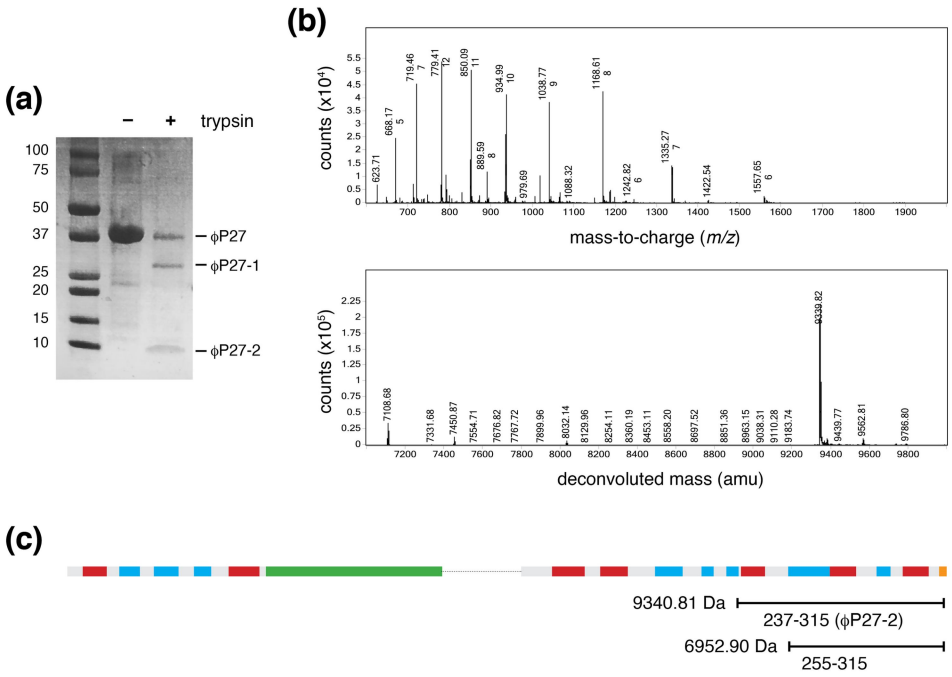



Figure 2.6. Limited proteolysis of φP27 Ea22.
(a) SDS-PAGE of purified full-length φP27 Ea22 protein before and after treatment with trypsin (b) Mass spectrum and deconvoluted spectrum of a peak obtained by reverse-phase chromatography at a retention time 10.62-11.03 min. The mass spectrum is consistent with the band labelled φP27-2. (c) Boundary of the best fit deconvoluted mass to the sequence of φP27-2 (observed: 9340.91 Da, expected: 9340.76 Da). In some longer protein digests, an additional C-terminal fragment was observed (observed: 6952.90 Da, expected: 6953.03 Da).

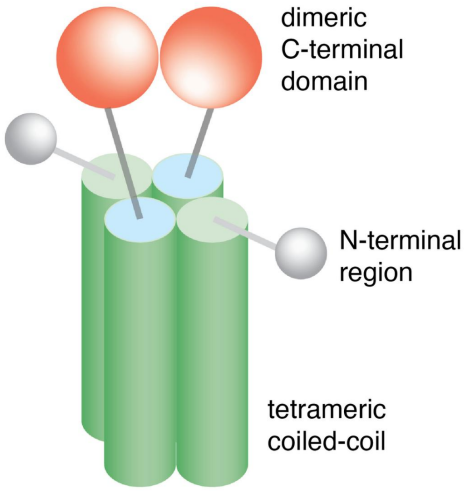



Figure 2.7. Possible architecture of Ea22.
The tetrameric coiled-coil region (green) is presented as an antiparallel four helix bundle placed centrally, leaving the C-terminal domain (red) free to dimerize on each end of the bundle. A long enough tether may allow the C-terminal ends to tumble freely. For clarity, only one C-terminal domain is shown. Since the N terminal region could not be expressed, the structure and oligomeric state of the N-terminal domain is unknown.

Table 2.1 Strains and bacteriophages

| Source | Genotype |
|---|---|
| *E. coli strains* | |
| MG1655 | F− λ− *ilvG rfb-50 rph-1* |
| MG1655 (λ) | MG1655 bearing λ prophage |
| MG1655 ($\varphi 24_B$) | MG1655 bearing $\varphi 24_B$ prophage |
| MG1655 (φP27) | MG1655 bearing φP27 prophage |
| | |
| Bacteriophages | |
| λ | carries a frameshift mutation relative to Ur-lambda |
| $\Phi 24_B$ | Δ*stx2*::*catGFP* |
| P27 | Δ*stx2*::*catGFP* |

Table 2.2 Ea22 protein fragments

| reference | source | fragment | solubility | comment |
|---|---|---|---|---|
| 134893 | λ | 1-182 | soluble | culture was grown at 16°C post-induction |
| 179655 | λ | 93-182 | soluble | culture was grown at 16°C post-induction |
| 179656 | λ | 102-182 | soluble | culture was grown at 16°C post-induction |
| 179657 | λ | 109-182 | soluble | culture was grown at 16°C post-induction |
| 179658 | λ | 119-182 | insouble | could not be refolded |
| pET28-ϕ24B-Ea22 | $\phi 24_B$ | 1-301 | soluble | culture was grown at 25°C post-induction |
| 350147 | $\phi 24_B$ | 159-301 | insoluble | could not be refolded |
| 351679 | $\phi 24_B$ | 164-269 | insoluble | could not be refolded |
| pET28-ϕP27-Ea22 | ϕP27 | 1-315 | soluble | culture was grown at 25°C post-induction |
| 350149 | ϕP27 | 1-62 | insoluble | could not be refolded |
| 350148 | ϕP27 | 145-315 | insoluble | could not be refolded |
| 358023 | ϕP27 | 176-315 | insoluble | could not be refolded |
| 393816 | ϕP27 | 235-315 | insoluble | could not be refolded |

Table 2.3 Biophysical analyses of Ea22 protein fragments

| Protein | Region | Oligomeric state | $T_m$ (CD, ˚C) | $T_m$ (DSC, ˚C) | ΔH (DSC, kJ/mol) |
|---|---|---|---|---|---|
| λ Ea22(1-182) | full length | tetramer | 48 | 48.4 ± 0.1 | 433 ± 3 |
| λ Ea22(93-182) | C-term domain | dimer | > 90 | > 100 | |
| λ Ea22(102-182) | C-term domain | dimer | > 90 | > 100 | |
| λ Ea22(109-182) | C-term domain | dimer | 82 | 90 | |
| $\varphi 24_B$ Ea22(135-261) | C-term domain | dimer | > 90 | > 100 | |
| φP27 Ea22(1-315) | full length | tetramer | 55 | 48.1 ± 0.1<br>52.5 ± 0.1 | 1048 ± 24<br>743 ± 23 |

# Chapter Three – Crystallization of the l Ea22 C-terminal domain

## 3.1 Introduction

Full length l Ea22 and f 27 Ea22 are both tetrameric in solution (total molecular weight ~80 kDa), NMR studies are not possible as signals diminish rapidly due to the slow tumbling rate of the tetramer. Since X-ray crystallography is not hindered by total molecular weight, an attempt was made to crystallize several Ea22 proteins. This chapter describes the ongoing work towards obtaining highly diffracting crystals suitable for data collection either at the York University X-ray diffractometer or the Canadian Light Source Synchrotron (Saskatoon, SK).

## 3.2 Materials and Methods

### 3.2.1 High throughput screening

Solutions of 10 mg/mL l Ea22, f 27 Ea22 and the l Ea22 C-terminal domain (600 µL in 10 mM Tris, 100 mM NaCl, pH 7.4) were mixed in equal amounts of one of 1536 precipitating solutions and overlaid with paraffin oil. The five 384-well plates containing the crystallization trial for each protein were incubated at 23 ˚C and imaged immediately after mixing (t=0) and then after one day and then each week to a maximum of six weeks. The images were inspected remotely using Macroscope J software. At the six-week mark, additional imaging was performed with UV light (to discern protein crystals from salt crystals) and SONICC laser imaging [49] (to discern poorly organized crystals from highly organized crystals).

### 3.2.2 In house screening

At York University, a series of plates were set up using similar protein solutions and precipitating solutions only at a much larger scale and the 140 nL protein-precipitant drops used by the high-throughput screen. Two methods [50] were pursued: (*i*) a hanging drop vapor diffusion method were 1 µL of protein is mixed with 1 µL of

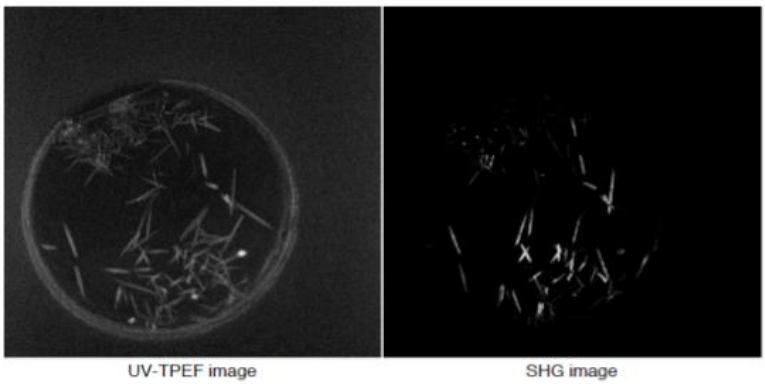



Figure 3.1. Representative crystals obtained by high throughput screening. The left panel depicts long needle like crystals imaged with UV light demonstrating the crystals are indeed protein. The right panel depicts SONICC laser imaging demonstrating that the crystals are likely ordered with a higher likelihood of producing high resolution diffraction patterns when subjected to X-rays. The specific crystallization condition show is 10 mg/mL Ea22 C-terminal fragment 0.1 M ammonium sulfate, 0.1 M TAPS pH 9, 40% PEG 8000 at 23 °C.

precipitant and the drop hung over 500 µL of precipitating solution in a sealed, closed environment provided by a specialized plastic plate. (*ii*) A similar strategy to the "micro-batch" high throughput screen where 1 µL of protein is mixed with 1 µL precipitant and overlaid with 15 µL paraffin oil to slowly equilibrate. A total of 48 conditions could be screened with one plastic cell hematology plate. Drops made by either method were inspected with a Nikon microscope equipped with a video camera and polarizing lens.

### 3.3 Results

As described in the previous chapter, the dimeric I Ea22 C-terminal fragment (~20 kDa total) produced good NMR spectra and was highly thermostable. It was not surprising that crystals were observed for a 10 mg/mL solution of this protein fragment at several conditions. The condition producing the largest crystals was 0.1 M ammonium sulfate, 0.1 M of TAPS pH 9, 40% PEG 8000 at 23 °C (Figure 3.1). This composition of this crystallization condition is in three components consisting of a buffering solution, a salt solution and a precipitating solution. In the HTSlab screens, the

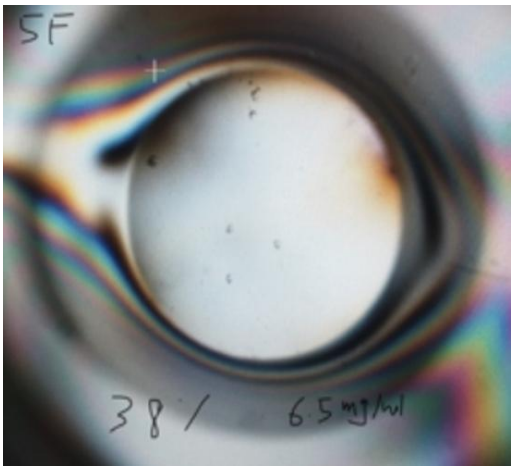



Figure 3.2. Micro-crystals of the l Ea22 C-terminal domain.
Specific conditions of the image: 4 days of incubation at room temperature with 6.5mg/mL protein mixed with a cocktail consisting of 0.1M of ammonium sulfate, 0.1M TAPS pH 9, and 38% PEG 8000.

typical precipitating components were various concentrations and molecular weights of polyethylene glycol (PEG) and various concentrations of ammonium sulfate. For this crystallization solution, ammonium sulfate serves as the salt rather than the precipitant.

A matrix of conditions near what was observed at the U Buffalo HTSlab was assayed at York University by varying either the method (hanging drop or micro-batch), the protein concentration or the concentration of the PEG 8000 precipitating component. As shown in Figure 3.2, micro-crystals were observed with the micro-batch methods at lower protein concentrations. While this result was encouraging, larger crystals were not obtained after further screening. Micro-seeding was also attempted. The micro-seeding method involved transferring one small crystals into a cocktail of protein and precipitating solution to stimulate further crystal growth [51].

### 3.4 Conclusions

X-ray crystallography offers a means to obtain high resolution structures of oligomeric proteins like Ea22. Further optimization is required to manufacture larger crystals that can be frozen and mounted on the X-ray diffractometer for a definitive

determination of crystal quality. Towards this objective, it may be beneficial to vary the drop size, temperature, and protein concentration further and explore other crystallization conditions that were identified by the high-throughput screen. Further optimization of the micro-seeding method is also possible.

# Chapter Four — Conclusions

### 4.1 Structure and function of l Ea22

I began this thesis work when relatively little was known about any of the *exo-xis* genes from phage l or any other related Shigatoxigenic phage. *Ea22* was selected for an investigation because (*i*) it was the largest gene of the group (and consequently, the largest protein), (*ii*) it demonstrates pro-lysogenic activity in several biological assays and (*iii*) is highly conserved suggesting a preservation of an important function in phage development.

Sequence comparison of related Ea22 proteins suggested a three-part, if not three-domain, organization. Initially, I expressed and purified the entire Ea22 protein and showed by SEC-MALS that it is tetrameric in solution. Next, I examined four C-terminal fragments of l Ea22 by NMR and stability assays to determine the minimal domain. The minimal domain was not only the most stable with a thermal denaturation point of > 100˚C, but it was also very soluble and expressed well in cultures at grown at room temperature. This protein formed the basis of a high throughout crystallographic survey performed at the University of Buffalo where several promising crystallization conditions were obtained. While I tried to optimize the crystals, Professor Donaldson worked concurrently on optimizing conditions for NMR spectroscopy.

After I left the laboratory to write up my thesis, Prof. Donaldson solved the structure of the dimeric l Ea22 C-terminal domain. The NMR structure of the l Ea22 C-terminal domain is shown in Figure 4.1. The structure of the Ea22 C-terminal domain expressed from f 24B phage strongly resembles Dmd as indicated in Figure 4.2, a protein from T4 phage that is acts as a viral counter-defense against a bacterial defense during late stages of infection [52]. Dmd works against RnlA, a ribonuclease that when activated, rapidly degrades viral mRNA and activates other defense gene expression programs in the bacterium. Due its activity, RnlA is tightly repressed by its natural anti-toxin partner, RnlB. As a counter-defense against the bacterial defense, Dmd substitute

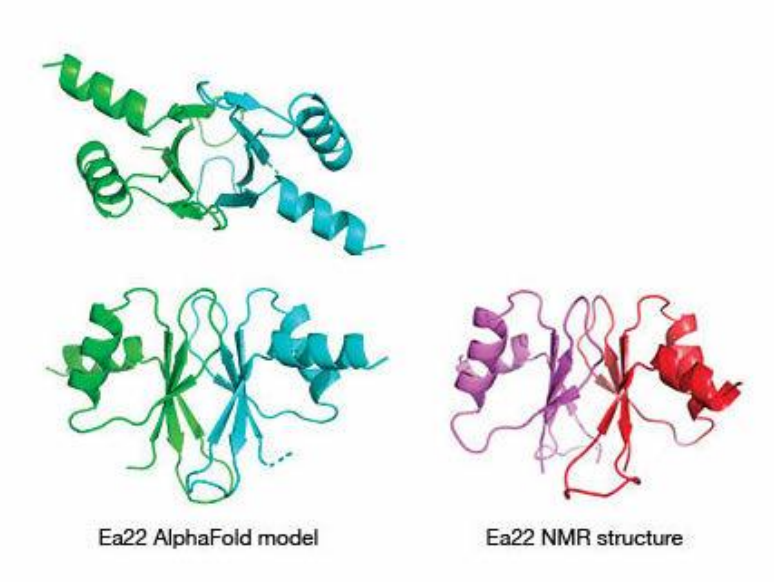
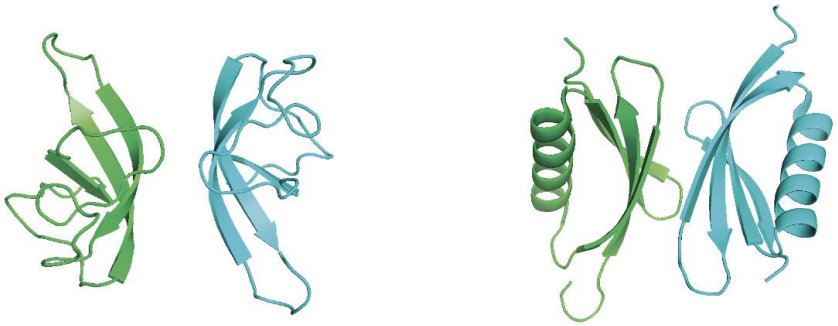



Figure 4.1 Ea22 C-terminal domain structure.
The left panel (in blue/cyan) depicts two orientations of the Ea22 model predicted by AlphaFold. The right panel (in purple/red) shows one orientation of the Ea22 NMR solution structure. There is a high degree of similarity between the AlphaFold and NMR structures that extends to the dimer interface.

for RnlB thereby ensuring the RnlA remains inactive. This toxin-antitoxin (TA) system is but one of many bacteria use against bacteriophages [53]. X-ray structures of RnlAB have helped determine its mechanism of action [54, 55]. Structure of T4 Dmd have also been determined both free and bound to RnlA [56, 57]. The monomers of Dmd and Ea22 are similar but in the dimer, the two monomers are in the opposite orientation. The same orientation of the monomers in l Ea22 is also predicted by Google AlphaFold as indicated by Figure 4.1 [58]. Looking deeper at the structure of the Ea22 C-terminal domain, three important ionic and hydrogen bonds are conserved between it and Dmd (Asp28, Gln35 and Glu38 in Dmd with to Arg 85, Arg 243 and Arg305 in RnlA). Some bacteria express an addition toxin-antitoxin pair on a placed known as LsoA/LsoB. The structure of LsoA and LsoB are known and Dmd can substitute for LsoB showing that there is a preservation of structure and function [57, 59]. In the Donaldson laboratory,



ϕ24B Ea22 C-term domain T4 phage Dmd

Figure 4.2 Structural prediction of the f 24B Ea22
AlphaFold was used to predict the dimeric state of the conserved f 24B Ea22 C-terminal domain. The fold has some similarities with the anti-toxin protein Dmd from T4 phage.

an attempt to co-express and co-purify LsoA and Ea22 was performed but preliminary results suggest that LsoA and Ea22 do not interact with each other. When milligram quantities of RnlA were made and mixed with Ea22 for an isothermal titration calorimetry study no binding was detected (Johnson laboratory, unpublished observations). Perhaps, not all of the critical contacts have been preserved. The possibility also remains that RnlA and LsoA are not the true biological targets of l Ea22. Towards resolving this issue, an affinity-purification-mass-spectrometry (AP-MS) study has been planned.

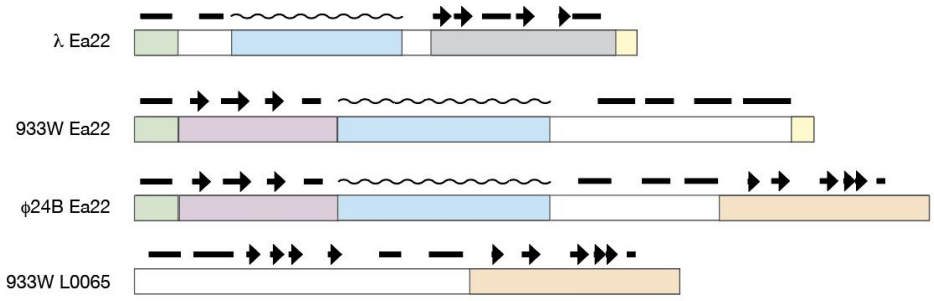



Figure 4.3 A mosaic organization of Ea22 proteins.
The sequences l Ea22 and two Shigatoxigenic Ea22 are compared. Each protein may be considered in terms of three or four conserved segments shown in different colours. l Ea22 has a different C-terminal segment/domain than its Shigatoxigenic counterparts. Furthermore, 933W Ea22 appears to be split between two genes (*ea22* and *L0065*).

### 4.2 The C-terminal domains of other Ea22 proteins

From sequence comparisons of l and Shigatoxigenic phages such as 933W, f 24B and f 27(Figure 4.3), it appears that the organization of the *ea22* gene is conserved and encodes an N-terminal segment, a coiled-coil central segment and a C-terminal segment. The C-terminal sequences shared by f 24B and f 27 Ea22 proteins are completely different than l Ea22 and may help these bacteriophages adapt to their pathogenic *E. coli* host. A closer inspection of the 933W and p24B genes also shows a possible genetic insertion event has split the N-terminal, coiled-coil and C-terminal segments across two separate genes in 933W. These observations were described in a recent publication [60].

### 4.3 The Ea22 protein from P27 phage

In addition to full-length l Ea22, I expressed and purified full-length Ea22 from phage f 27. From a SEC-MALS study, f 27 Ea22 was tetrameric like l Ea22. A sample

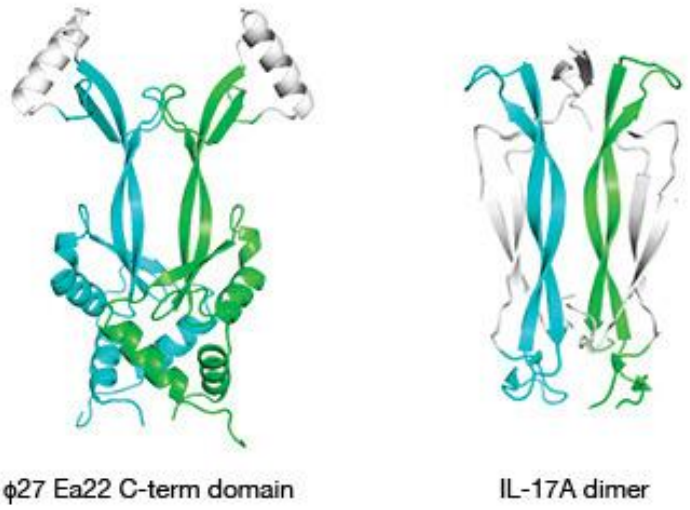



Figure 4.4 Structure prediction of f 27 Ea22 C-terminal domain.
Google AlphaFold was used to predict the dimeric form of the C-terminal domain. The predicted structure shares features with human cytokines such as IL-17 (PDB: 4RH9) shown here including a disulfide bond between the two monomers. Structurally dissimilar segments of each representative are coloured white.

of full-length protein was subjected to limited proteolysis for a varying amount of time to produce a series of metastable protein fragments. Several fragments were observed on a Coomassie-blue stained denaturing polyacrylamide gel suggesting that f 27 Ea22 has a multidomain organization. Further analysis of the gel fragments by mass spectrometry revealed a stable C-terminal domain. The predicted structure of this f 27 Ea22 C-terminal fragment as a dimer does not resemble an anti-toxin like I Ea22. In contrast, it resembles a cysteine-knot fold that is a hallmark of many cytokines like interleukin-17 (Figure 4.4). Unfortunately, efforts to produce the C-terminal fragments of f 27 Ea22 identified by limited proteolysis were unsuccessful. The proteins were expressed as inclusion bodies and could not be refolded from denaturant. As a result, no structural studies could proceed.

The last Ea22 protein that I investigated was from the Shigatoxigenic phage f 24B. The NMR data I presented in Chapter 2 suggests that there might be a folded domain in the C-terminal fragment, but no further work was done except to show that it was dimeric like the other C-terminal fragments described. An AlphaFold structure prediction of dimeric f 24B Ea22 is shown in Figure 4.4. This fold bears some resemblance to T4 phage Dmd, especially in the number and orientation of the b-strands. It is unknown if f 24B Ea22 is an anti-toxin although in several biological assays it exerts a strong effect on phage development [31].

### 4.4 Future directions

The *exo-xis* proteins represent a new platform for understanding aspects of phage development. As these proteins are further characterized by biochemical, structural and proteomics methods, they may reveal new ways of reducing the ill effects of food poisoning from pathogenic E. coli that naturally harbour Shigatoxigenic phages.

# Appendix

Sequences of proteins and protein fragments of Ea22 used in this work, in FASTA format. The header contains the clone #, generic name, length of protein, molecular weight, and absorbance of a 1 mg/mL solution at 280 nm. Each protein contains an N-terminal His-tag for affinity purification.

**>134893 ea22-v0 195aa 22.49kDa 1mg/mL=0.89**
KHHHHHHMSEINSQALREAAEQAMHDDWGFDADLFHELVTPSIVLELLDERERNQQYIKRRDQENEDIALTVGKLRVELETAKSKL
NEQREYYEGVISDGSKRIAKLESNEVREDGNQFLVVRHPGKTPVIKHCTGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDA
GEYLHMMSDAGIRIKGEIETAVG

**>137640 ea22-v1 101aa 11.54kDa 1mg/mL=0.85**
HHHHHHMGSKRIAKLESNEVREDGNQFLVVRHPGKTPVIKHCTGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDAGEYL
HMMSDAGIRIKGEIETAV

**>179655 ea22v2 93aa 10.68kDa 1mg/mL=0.93**
HHHHHHMSNEVREDGNQFLVVRHPGKTPVIKHCTGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDAGEYLHMMSDAGIRIK
GEIETAV

**>179656 ea22v3 86aa 9859kDa 1mg/ml=1.01**
HHHHHHMGNQFLVVRHPGKTPVIKHCTGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDAGEYLHMMSDAGIRIKGEIETAV

**>179657 ea22v4 76aa 8.71kDa 1mg/mL=1.14**
HHHHHHMGKTPVIKHCTGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDAGEYLHMMSDAGIRIKGEIETAV

**>179658 ea22v5 66aa 7.64kDa 1mg/mL=1.30**
HHHHHHMGDLEEFLRQLIEQDPLVTIDIITHRYYGVGGQWVQDAGEYLHMMSDAGIRIKGEIETAV

**>358023 p27c1(176-315) 147aa 17.34kDa 1mg/mL=1.46**
HHHHHHMSAPDSFGIIGENIRTQDNRITSDPMFCVYQKREIVVDADYDYDRIVWVDEDGNEANKLQSRRLELLHENFREPPEKWRR
VAVKDIDEFVTCCFTEQGCKDYLAVNGHNLRLPFIYVKSGFRNAEYIGIRNWLAGIRIKGE

**>350148 p27c2(145-315) 21.04kDa 178aa 1mg/mL=1.47**
HHHHHHMEKQCAEWERKALSNFEECAAMAERIEEMQTKSAPDSFGIIGENIRTQDNRITSDPMFCVYQKREIVVDADYDYDRIVWV
DEDGNEANKLQSRRLELLHENFREPPEKWRRVAVKDIDEFVTCCFTEQGCKDYLAVNGHNLRLPFIYVKSGFRNAEYIGIRNWLAG
IRIRGE

**>393816 p27c3(235-315) 97aa 11.42kDa 1mg/mL=1.36**
KHHHHHHMGSLVPRGSSRRLELLHENFREPPEKWRRVAVKDIDEFVTCCFTEQGCKDYLAVNGHNLRLPFIYVKSGFRNAEYIGIR
NWLAGIRIKGE

**>350149 phi27N(1-62) 69aa 7.68kDa 1mg/mL=1.12**
HHHHHHMVSEINYQALREKAEKATKGSYIVGHTSVNQHGNLTGVFVCQKWKGEPGGVIAECHVNCLVET

**>351679 phi24bc2(164-269) 114aa 13.30kDa 1mg/mL=2.00**
KHHHHHHMKLPHTQFEQIANLYEMQFDDGRTCAFHTDAQKAEQWLQACDGNRVQEYVKLERLQNALSGNSPVTPDGWISCSERMPD
TKTAVLVAVEFDRKGDWRMKWATYIPGH

**>350147 phi24bC1(159-301) 155aa 17.88kDa 1mg/mL=2.49**
HHHHHHMEAELVSQTYKLPHTQFEQIANLYEMQFDDGRTCAFHTDAQKAEQWLQACDGNRVQEYVKLERLQNALSGNSPVTPDGWI
SCSERMPDTKTAVLVAVEFDRKGDWRMKWATYIPGHPDANDGWIIPGASWKPSHWMPLPEPPQEVNRGE

# References


1. Kaper JB, O'Brien AD (2014) Overview and Historical Perspectives. Microbiol Spectr 2:.

2. Matthews L, Reeve R, Gally DL, et al (2013) Predicting the public health benefit of vaccinating cattle against Escherichia coli O157. Proc National Acad Sci 110:16265–16270.

3. Holme R (2003) Drinking water contamination in Walkerton, Ontario: positive resolutions from a tragic event. Water Sci Technol 47:1–6.

4. Buchholz U, Bernard H, Werber D, et al (2011) German Outbreak of Escherichia coli O104:H4 Associated with Sprouts. New Engl J Medicine 365:1763–1770.

5. Mayer CL, Leibowitz CS, Kurosawa S, Stearns-Kurosawa DJ (2012) Shiga Toxins and the Pathophysiology of Hemolytic Uremic Syndrome in Humans and Animals. Toxins 4:1261–1287.

6. Keen EC (2015) A century of phage research: Bacteriophages and the shaping of modern biology. Bioessays 37:6–9.

7. Hatfull GF, Hendrix RW (2011) Bacteriophages and their genomes. Curr Opin Virol 1:298–303.

8. Binnenkade L, Teichmann L, Thormann KM (2014) Iron Triggers λSo Prophage Induction and Release of Extracellular DNA in Shewanella oneidensis MR-1 Biofilms. Appl Environ Microb 80:5304–5316.

9. Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B, et al (2012) Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. Curr Protein Pept Sc 13:699–722.

10. Węgrzyn G, Licznerska K, Węgrzyn A (2012) Chapter 6 Phage λ—New Insights into Regulatory Circuits. Adv Virus Res 82:155–178.

11. Sergueev K, Yu D, Austin S, Court D (2001) Cell toxicity caused by products of the pL operon of bacteriophage lambda. Gene 272:227–235.

12. Smith DL, Rooks DJ, Fogg PC, et al (2012) Comparative genomics of Shiga toxin encoding bacteriophages. Bmc Genomics 13:311.

13. Elliott SJ, Sperandio V, Girón JA, et al (2000) The Locus of Enterocyte Effacement (LEE)-Encoded Regulator Controls Expression of Both LEE- and Non-LEE-Encoded Virulence Factors in Enteropathogenic and EnterohemorrhagicEscherichia coli. Infect Immun 68:6115–6126.

14. Cogliano MED, Pinto A, Goldstein J, et al (2018) Relevance of Bacteriophage 933W in the Development of Hemolytic Uremic Syndrome (HUS). Front Microbiol 9:3104.

15. Wagner PL, Acheson DWK, Waldor MK (2001) Human Neutrophils and Their Products Induce Shiga Toxin Production by Enterohemorrhagic Escherichia coli. Infect Immun 69:1934–1937.

16. Licznerska K, Dydecka A, Bloch S, et al (2016) The Role of the Exo-Xis Region in Oxidative Stress-Mediated Induction of Shiga Toxin-Converting Prophages. Oxid Med Cell Longev 2016:1–14.

17. Dydecka A, Bloch S, Necel A, et al (2020) The ea22 gene of lambdoid phages: preserved prolysogenic function despite of high sequence diversity. Virus Genes 56:266–277.

18. Silva CJ, Brandon DL, Skinner CB, He X (2017) Shiga toxins: A Review of Structure, Mechanism, and Detection. Springer International Publishing

19. Paton JC, Paton AW (1998) Pathogenesis and Diagnosis of Shiga Toxin-Producing Escherichia coli Infections. Clin Microbiol Rev 11:450–479.

20. Mauro SA, Koudelka GB (2011) Shiga Toxin: Expression, Distribution, and Its Role in the Environment. Toxins 3:608–625.

21. Hunt JM (2010) Shiga Toxin–Producing Escherichia coli (STEC). Clin Lab Med 30:21–45.

22. Herold S, Karch H, Schmidt H (2004) Shiga toxin-encoding bacteriophages – genomes in motion. Int J Med Microbiol 294:115–121.

23. Waldor MK, Friedman DI (2005) Phage regulatory circuits and virulence gene expression. Curr Opin Microbiol 8:459–465.

24. Smith DL, Rooks DJ, Fogg PC, et al (2012) Comparative genomics of Shiga toxin encoding bacteriophages. BMC Genomics 13:311.

25. Schmidt H (2001) Shiga-toxin-converting bacteriophages. Res Microbiol 152:687–695.

26. Allison HE, Sergeant MJ, James CE, et al (2003) Immunity Profiles of Wild-Type and Recombinant Shiga-Like Toxin-Encoding Bacteriophages and Characterization of Novel Double Lysogens. Infect Immun 71:3409–3418.

27. Sergueev K, Court D, Reaves L, Austin S (2002) E. coli Cell-cycle Regulation by Bacteriophage Lambda. J Mol Biol 324:297–307.

28. Kourilsky P, Knapp A (1975) Lysogenization by bacteriophage lambda III. - Multiplicity dependent phenomena occuring upon infection by lambda. Biochimie 56:1517–1523.

29. Bloch S, Nejman-Faleńczyk B, Dydecka A, et al (2014) Different Expression Patterns of Genes from the Exo-Xis Region of Bacteriophage λ and Shiga Toxin-Converting Bacteriophage Φ24B following Infection or Prophage Induction in Escherichia coli. PLoS ONE 9:e108233.

30. Łoś JM, Łoś M, Węgrzyn A, Węgrzyn G (2008) Role of the bacteriophage λ exo-xis region in the virus development. Folia Microbiol 53:443–450.

31. Dydecka A, Bloch S, Necel A, et al (2020) The ea22 gene of lambdoid phages: preserved prolysogenic function despite of high sequence diversity. Virus Genes 1–12.

32. Dydecka A, Nejman-Faleńczyk B, Bloch S, et al (2018) Roles of orf60a and orf61 in Development of Bacteriophages λ and Φ24B. Viruses 10:553.

33. Dydecka A, Bloch S, Rizvi A, et al (2017) Bad Phages in Good Bacteria: Role of the Mysterious orf63 of λ and Shiga Toxin-Converting Φ24B Bacteriophages. Front Microbiol 8:1618.

34. Kwan JJ, Smirnova E, Khazai S, et al (2013) The solution structures of two prophage homologues of the bacteriophage λ Ea8.5 protein reveal a newly discovered hybrid homeodomain/zinc-finger fold. Biochemistry 52:3612–4.

35. Licznerska K, Dydecka A, Bloch S, et al (2016) The Role of the Exo-Xis Region in Oxidative Stress-Mediated Induction of Shiga Toxin-Converting Prophages. Oxid Med Cell Longev 2016:1–14.

36. Recktenwald J, Schmidt H (2002) The Nucleotide Sequence of Shiga Toxin (Stx) 2e-Encoding Phage φP27 Is Not Related to Other Stx Phage Genomes, but the Modular Genetic Structure Is Conserved. Infect Immun 70:1896–1908.

37. Delaglio F, Grzesiek S, Vuister GW, et al (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293.

38. Fogh RH, Vranken WF, Boucher W, et al (2006) A nomenclature and data model to describe NMR experiments. J Biomol NMR 36:147–155.

39. Klausen MS, Jespersen MC, Nielsen H, et al (2019) NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. Proteins Struct Funct Bioinform 87:520–527.

40. Zimmermann L, Stephens A, Nam S-Z, et al (2018) A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol 430:2237–2243.

41. Blasche S, Wuchty S, Rajagopala SV, Uetz P (2013) The Protein Interaction Network of Bacteriophage Lambda with Its Host, Escherichia coli. J Virol 87:12745–12755.

42. Rajagopala SV, Casjens S, Uetz P (2011) The protein interaction map of bacteriophage lambda. BMC Microbiol 11:213.

43. Hatfull GF, Hendrix RW (2011) Bacteriophages and their genomes. Curr Opin Virol 1:298–303.

44. Hatfull GF (2015) Dark Matter of the Biosphere: The Amazing World of Bacteriophage Diversity. J Virol 89:8107 8110.

45. Edmonds L, Liu A, Kwan JJ, et al (2007) The NMR Structure of the gpU Tail-terminator Protein from Bacteriophage Lambda: Identification of Sites Contributing to Mg(II)-mediated Oligomerization and Biological Function. J Mol Biol 365:175–186.

46. Pell LG, Kanelis V, Donaldson LW, et al (2009) The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. Proc Natl Acad Sci USA 106:4160–5.

47. Bloch S, Nejman-Faleńczyk B, Łoś JM, et al (2013) Genes from the exo–xis region of λ and Shiga toxin-converting bacteriophages influence lysogenization and prophage induction. Arch Microbiol 195:693–703.

48. Nejman-Faleńczyk B, Bloch S, Licznerska K, et al (2015) A small, microRNA-size, ribonucleic acid regulating gene expression and development of Shiga toxin-converting bacteriophage Φ24B. Sci Rep 5:srep10080.

49. Guo F, Gualtieri E, Simpson G, Jiang W (2010) SONICC: A Novel Nonlinear Optical Detection Technique for 2D Celluar Crystallography. Biophys J 98:385a.

50. Gorrec F (2021) A beginner’s guide to macromolecular crystallization. Biochem 43:36–43.

51. Stewart PS (2014) “How to use ‘“random”’ microseeding before you get your first crystals.” Acta Crystallogr Sect Found Adv 70:C1756–C1756.

52. Wei Y, Gao Z, Otsuka Y, et al (2013) Structure–function studies of Escherichia coli RnlA reveal a novel toxin structure involved in bacteriophage resistance. Mol Microbiol 90:956–965.

53. Jurėnas D, Fraikin N, Goormaghtigh F, Melderen LV (2022) Biology and evolution of bacterial toxin–antitoxin systems. Nat Rev Microbiol 1–16.

54. Garcia-Rodriguez G, Charlier D, Wilmaerts D, et al (2021) Alternative dimerization is required for activity and inhibition of the HEPN ribonuclease RnlA. Nucleic Acids Res 49:7164–7178.

55. Garcia-Rodriguez G, Perez AT, Konijnenberg A, et al (2020) The Escherichia coli RnlA–RnlB toxin–antitoxin complex: production, characterization and crystallization. Acta Crystallogr Sect F 76:31–39.

56. Wei Y, Gao Z, Zhang H, Dong Y (2016) Structural characterizations of phage antitoxin Dmd and its interactions with bacterial toxin RnlA. Biochem Bioph Res Co 472:592–597.

57. Wan H, Otsuka Y, Gao Z, et al (2016) Structural insights into the inhibition mechanism of bacterial toxin LsoA by bacteriophage antitoxin Dmd. Mol Microbiol 101:757–769.

58. Jumper J, Evans R, Pritzel A, et al (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589.

59. Otsuka Y, Yonesaki T (2012) Dmd of bacteriophage T4 functions as an antitoxin against Escherichia coli LsoA and RnlA toxins. Mol Microbiol 83:669–681.

60. Donaldson LW (2021) Molecular Modeling the Proteins from the exo-xis Region of Lambda and Shigatoxigenic Bacteriophages. Antibiotics 10:1282.